Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Department of the Environment, Transport, Energy and Communication
DETEC

Swiss Federal Office of Energy SFOE
Section Geoinformation & Digital Innovation

Report from 10 April 2024

Smart INteroperability Architecture (SINA): the Decentralized Data Space in the Building Industry

Quelle: HSLU 2024

With the Support of swissenergy

Date: 10 April 2024

Town: Bern

Publisher:

Swiss Federal Office of Energy SFOE
CH-3003 Bern
www.bfe.admin.ch

Agent:

CSEM Neuchâtel
Rue Jaquet-Droz 1, 2002 Neuchâtel
www.csem.ch

Hochschule Luzern, Technik & Architektur
Technikumstrasse 21, 6048 Horw
www.hslu.ch/technik-architektur

Hochschule Luzern, Informatik
Campus Zug-Rotkreuz, Suurstoffi 1, 6343 Rotkreuz
www.hslu.ch/informatik

Co-financing:

Allthisfuture AG / WWZ AG, 6301 Zug, www.wwz.ch
arcade solutions ag, 6003 Luzern, www.arcade.ch
bonacasa AG, 7202 Oensingen, www.bonacasa.ch
esolva AG, 8570 Weinfelden, www.esolva.ch
Intellitec AG Stans, 6370 Stans, www.intelitec.ch
Plutinsus, Wolfram Willuhn, 8055 Zürich, www.plutinsus.com
Privera AG, 3073 Gümligen, www.privera.ch
St.Gallisch-Appenzellische Kraftwerke AG, 9001 St.Gallen, www.sak.ch
St.Galler Stadtwerke, 9001 St.Gallen, www.sgsw.ch
SIE SA, 1023 Crissier, www.sie.ch
Zug Estates AG, 6300 Zug, zugestates.ch
Zukunftsregion Argovia c/o Eniwa AG, 5033 Buchs, zukunftsregion-argovia.ch

Authors:

Said, Ahsaine, HSLU, said.ahsaine@hslu.ch
Jekaterina, Dmitrijeva, HSLU, jekaterina.dmitrijeva@hslu.ch
Tomasz, Gorecki, CSEM, tomasz.gorecki@csem.ch
Lynn, Grau, HSLU, lynn.grau@hslu.ch
Christoph, Imboden, HSLU, christoph.imboden@hslu.ch
Marco, Kunz, HSLU, marco.kunz@hslu.ch
Markus, Raschke, HSLU, markus.raschke@hslu.ch
Eugen, Rodel, HSLU, eugen.rodel@hslu.ch
Andreas, Rumsch, HSLU, andreas.rumsch@hslu.ch
Yves, Stauffer, CSEM, yves.stauffer@csem.ch
Olivier, Steiger, HSLU, olivier.steiger@hslu.ch
Ursula, Sury, HSLU, ursula.sury@hslu.ch

SFOE head of domain: Dr. Matthias Galus, Head of Section Geoinformation & Digital Innovation, matthias.galus@bfe.admin.ch.ch

SFOE programme manager: Dr. Matthias Galus, Head of Section Geoinformation & Digital Innovation matthias.galus@bfe.admin.ch.ch

Project lead: Christoph Imboden, Head of the Competence Center for Business Engineering HSLU, christoph.imboden@hslu.ch

SFOE contract number: SH/8100398-02-01-35

The authors of this report bear the entire responsibility for the content and for the conclusions drawn therefrom.

Swiss Federal Office of Energy SFOE
Pulverstrasse 13, CH-3063 Ittigen; postal address: Swiss Federal Office of Energy SFOE, CH-3003 Bern
Phone +41 58 462 56 11 · Fax +41 58 463 25 00 · contact@bfe.admin.ch · www.bfe.admin.ch

Zusammenfassung

Die Notwendigkeit, die ambitionierten Dekarbonisierungsziele auch vor dem Hintergrund geopolitischer Unwägbarkeiten zu erreichen, zwingt den Schweizer Energiesektor zu einer massiven Transformation. Digitalisierung leistet hierzu einen grossen Beitrag, indem die unausweichliche Komplexität besser gemanagt werden kann und Innovationen transformativ wirken können. Dabei spielen digitale Daten, ihre Verfügbarkeit und Zugänglichkeit eine eminent wichtige Rolle. Dies insbesondere vor dem Hintergrund, dass die Systeme der Strom-, Gas- und Wärmeversorgung mit Mobilität und Gebäuden zusammenwachsen.

Potential von Energiemanagementsystemen (HEMS) und digitalen Daten

Daten über heutige Systemgrenzen hinaus zugänglich zu machen, auszutauschen und zu nutzen bietet mit Blick auf die Anwendung von künstlicher Intelligenz und weiteren Innovationen grosses Potential. Die Europäische Kommission hat dieses Potential für Europa schon länger erkannt und fördert deswegen die Entwicklung von sogenannten vertrauenswürdigen Datenräumen - auch im Energiesektor. Erst kürzlich entschied der Bundesrat, Datenräume in der Schweiz ebenfalls zu fördern[1]. Ein kollaborativer Datenaustausch und die Datennutzung ermöglichen beispielsweise Stromeinsparungen von 6-9% alleine durch die Sensibilisierung über Messdaten des Verbrauches[2]. Die vorliegenden Untersuchungen kommen erstmals zu übergreifenden Schätzungen des Potentials von (Home) Energiemanagementsystemen (HEMS). Unter der Annahme, dass 30% aller Haushalte Zugang zu HEMS haben, können in der Schweiz jährlich mindestens 2 TWh an Energie oder 240’000 Tonnen an CO_2 eingespart werden (Abbildung A, Szenario Selektiv).

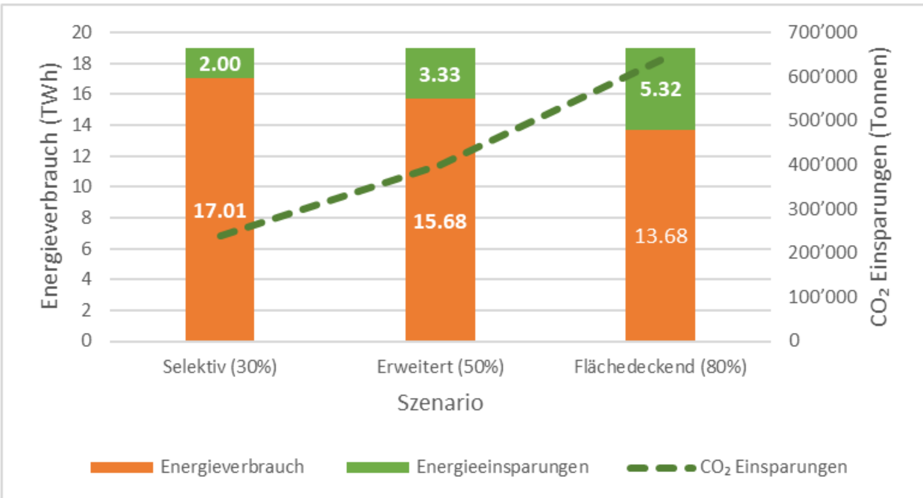

Abbildung A: Schätzung der Energie- und CO_2-Einsparungen für drei Szenarien bezüglich der Benutzung von HEMS in wachsender Anzahl (%) aller Haushalte in der Schweiz

[1] Förderung vertrauenswürdiger Datenräume und der digitalen Selbstbestimmung (admin.ch)

[2] Digitale Energieberatung führt zu Einsparungen von 9% (enerlytica.com)

HEMS ermöglichen neben Einsparungen auch die optimale Abstimmung von erneuerbarer Stromproduktion mit Elektromobilität und der Wärmeversorgung über Wärmepumpen. Das Bundesamt für Energie (BFE) hat die strategische Bedeutung von HEMS für die digital unterstützte Systemtransformation erkannt und unterstützt ihre Verbreitung durch die Reduktion von Komplexität für Nachfrager mit einen Marktüberblick zu HEMS[3].

Datenräume zum Austausch digitaler Daten über System- und Plattformgrenzen hinweg

Über regulatorische Instrumente wie die Etablierung einer nationalen Datenplattform (Datahub) zum Austausch von digitalen Messdaten und ihren Aggregaten unter den regulierten Akteuren im Strom- und Gassektor sollen zeitnah Verfügbarkeit und Zugang zu digitalen Daten gestärkt werden. Der Datenaustausch zwischen vielen weiteren Akteuren insbesondere der Privatwirtschaft bleibt jedoch eine Herausforderung. Der Bedarf einer robusten, interoperablen und kosteneffizienten digitalen Infrastruktur, die es ermöglicht, Daten über Plattform- und Applikationsgrenzen hinweg zu verknüpfen steigt jedoch vehement.

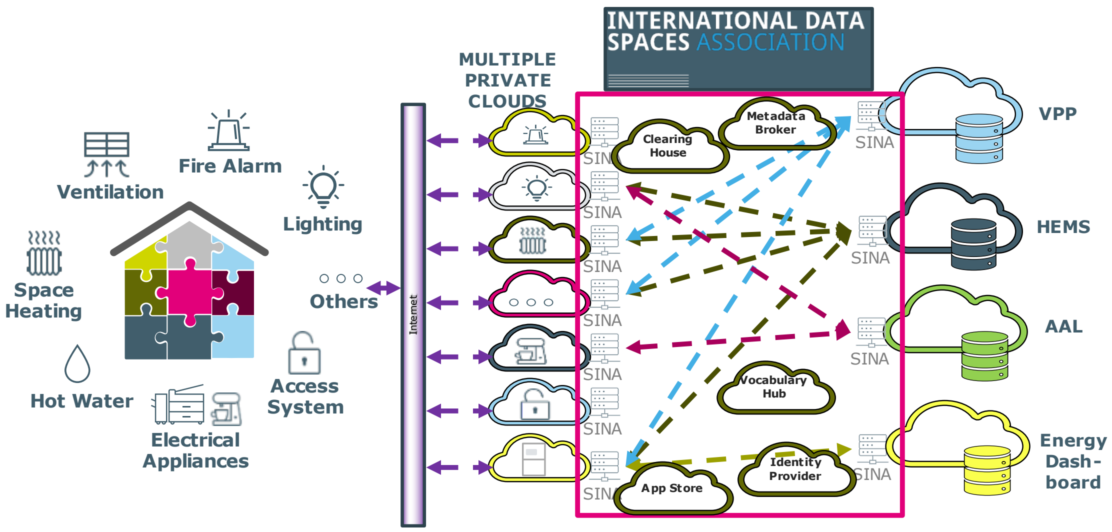

Abbildung B: Durch die Anbindung dezentraler Datenspeicher an den Datenraum werden die Daten verschiedenen Nutzern zur Verfügung gestellt

Der hier konkretisierte Ansatz der Smart INteroperability Architecture (SINA) versteht sich als eine Möglichkeit, einen vertrauenswürdigen Datenraum im Energiesektor aufzubauen und so regulierte Akteure und den Datahub mit weiteren dezentralen Datenspeichern der Privatwirtschaft sicher und vertrauenswürdig zu verknüpfen. SINA bietet einen strukturierten technischen und organisatorischen Rahmen für den Zugang, die Verarbeitung und die anschliessende Nutzung von dezental gehaltenen Daten über die Grenzen der regulierten Sektoren hinweg in die Privatwirtschaft. Der Ansatz bietet eine offene Inf-

[3] Marktübersicht Energiemanagementsysteme (EMS) (admin.ch)

rastruktur, welche Akteuren einen gleichberechtigten Datenzugang auf der Grundlage eines gemeinsamen Kodex[4] mit Vereinbarungen, Regeln und Normen ermöglicht. Dabei verzichtet der Datenraum gänzlich auf eine zentrale Datenhaltung und die Installation zusätzlicher peripherer Hardware. Stattdessen setzt er auf standardisierte, digitale Schnittstellen (Application Programming Interfaces API) zwischen den dezentralen Datenspeichern verschiedener Akteure. Exemplarisch wird in Abbildung B dargestellt, wie private Datenspeicher (z.B. Cloudspeicher) verschiedener Akteure durch einen Konnektor in einem Datenraum integriert sind.

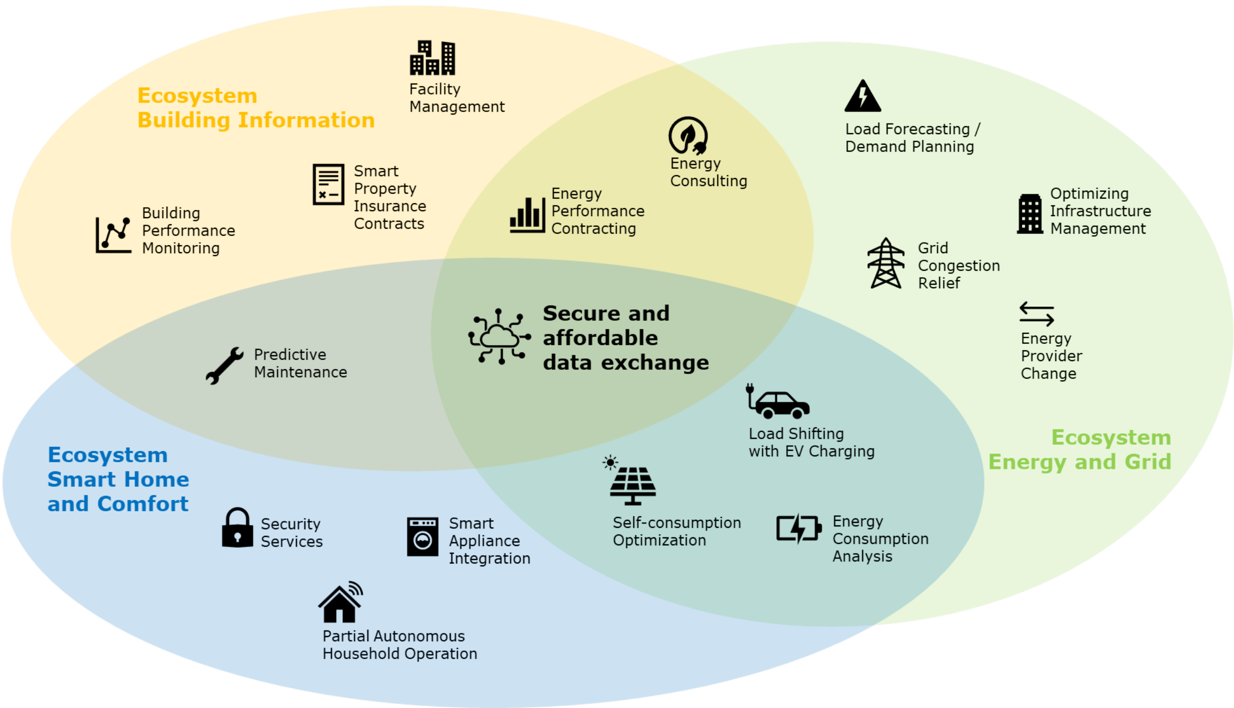

Abbildung C: Die drei Ökosysteme Building Information, Smart Home and Comfort, sowie Energy and Grid und die wichtigsten der dazugehörigen, identifizierten Anwendungsfälle

Weiter werden im Rahmen von SINA konzeptionelle, ökonomische und ökologische Grundlagen für die Ausgestaltung eines vertrauenswürdigen Datenraumes im Energiesektor am Beispiel ausgewählter Anwendungsfälle im Gebäudebereich geschaffen. Die Ausarbeitung und Bewertung verschiedener Anwendungsfälle in den Ökosystemen Building Information, Smart Home and Comfort, sowie Energy and Grid skizziert das Potential des Datenraums Energie (Abbildung C). Die Anwendungsfälle wurden in Zusammenarbeit mit Experten aus der Industrie identifiziert. Wie sich zeigt, erscheinen viele Anwendungsfälle schon heute realisierbar.

[4] Verhaltenskodex für den Betrieb von vertrauenswürdigen Datenräumen (admin.ch)

Die wichtigsten konzeptionell-technischen Grundsätze zur Schaffung eines Datenraumes im Energiesektor orientieren sich an den Prinzipien der International Data Space Association (IDSA), die auf Europäischem Niveau bereits viele Vorarbeiten geleistet hat. Im Rahmen von SINA werden technische und organisatorische Aspekte vertieft analysiert und für einen Anwendungsbereich angewandt. Dabei wird erstmalig die Bedeutung von bestehenden Ontologien näher beleuchtet, die notwendig sind für ein effizientes Datenmanagement und eine interoperable Nutzung der Daten zwischen verschiedenen Domänen, bspw. zwischen Gebäude und Mobilität.

Die Analyse verbleibt nicht auf theoretischem Niveau. Vielmehr wird anhand einer angewandten Machbarkeitsstudie demonstriert, wie ein vertrauenswürdiger Datenraum, unter der Berücksichtigung aktueller Standards, technisch mit einem Peer-to-Peer (P2P) Kommunikationsmodell umgesetzt und realisiert werden kann. Darüber hinaus wird die Verknüpfung des Datenraumes und einer Energieanwendung demonstriert, welche intelligente, Blockchain-basierte Verträge und Transaktionen umfasst.

Die Arbeiten demonstrieren, dass es möglich ist unter Zuhilfenahme der verfügbaren technischen Grundlagen der IDSA einen vertrauenswürdigen Datenraum aufzubauen und zu betreiben. Der hier aufgebaute Datenraum realisiert a) den Nachweis zur Beseitigung redundanter Datenspeicherung; b) die Wahrung der Datensouveränität durch die Befähigung der Dateneigentümer, Zugriffsberechtigungen autonom zu vergeben und zu entziehen; c) die Sicherstellung der überprüfbaren Authentizität der Teilnehmeridentität; und d) die Dezentralisierung und Orchestrierung des Datenmanagements. Dabei kommen Ontologien zur Anwendung, um eine barrierefreie Nutzung der Daten zu ermöglichen. Ontologien wie SAREF oder EEBus sind strukturierte Vokabulare, welche mittels eines Vocabulary Hub in den Datenraum integriert werden, um eine sektorübergreifende Dateninteroperabilität zu ermöglichen.

Aktuelle Herausforderungen von vertrauenswürdigen Datenräumen

Zugleich werden bei dieser ersten Realisierung eines Datenraumes diverse Herausforderungen identifiziert. Ontologien werden oft durch «Communities» gepflegt, werden laufend weiterentwickelt, sind oft schwer verständlich bzw. komplex und liegen nicht immer in digital verwendbarer Form vor. Für den Energiesektor existiert eine Vielzahl an Ontologien, die nicht zentral verwaltet werden. Die Softwarekomponente für einen Vocabulary Hub steht zwar zur Verfügung, jedoch noch in einer frühen Version. Ausserdem ist die Dokumentation dazu noch spärlich, was die Integration in einen Datenraum erschwert.

Der Aufwand für Hersteller - beispielsweise von Geräten wie Wechselrichtern - ihre privaten Datenspeicher durch einen Konnektor zu integrieren ist noch hoch, weil keine Referenzimplementation vorliegt. Zudem ist die Verwaltung von Zertifikaten mit Aufwand verbunden. Auch die Lösungsfindung bei Echtzeitdatenbedarf ist noch nicht klar.

In vielen Projekten, meist auf Europäischer Ebene, werden zurzeit Datenräume realisiert. Dabei ist festzustellen, dass meist für jedes Projekt ein neuer Datenraum aufgebaut und nicht auf einen bestehenden zurückgegriffen wird. Die Europäische Union hat in der europäischen Datenstrategie gemeinsame Datenräume (Common Data Spaces) für diverse Domänen vorgeschlagen. In einem gemeinsamen Datenraum sollen alle Daten, welche im Zusammenhang mit einer Domäne wie z.B. Energie oder Mobilität stehen, ausgetauscht werden können. Die gemeinsamen Datenräume sind allerdings noch zu wenig weit fortgeschritten, so dass sie in den Projekten im Moment noch keine brauchbare Alternative darstellen. Hier besteht die Herausforderung, diese gemeinsamen Datenräume weiter voranzutreiben und bekannt zu machen.

Rechtliche und organisatorische Überlegungen zu Datenräumen

Bei der Implementierung von Datenräumen müssen verschiedene Gesetze eingehalten werden, insbesondere bezüglich datenrechtlicher Belange, wie die General Data Protection Regulation (GDPR) der EU, oder das Eidgenössische Datenschutzgesetz (DSG). Dies erhöht die Komplexität und den Aufwand. Das Framework der IDSA bietet hier Lösungen, da es mit den Datenschutzbestimmungen der Europäischen Union konform ist und fortlaufend an die Änderungen des rechtlichen Rahmens angepasst wird. Damit vereinfacht es den Aufbau und Betrieb eines Datenraumes wesentlich. Dennoch muss in der Schweiz auf Spezifika des DSG geachtet werden.

Die Untersuchung der organisatorischen Aspekte des Datenraumes verdeutlicht die Notwendigkeit der Etablierung eines unterstützenden Ökosystems. Eine vertrauenswürdige non-profit Organisation, beispielsweise ein Verein, kann eine konsensorientierte Verwaltung und Weiterentwicklung des Datenraumes sicherstellen. Eine gemeinsame, breit abgestützte Governance ist von entscheidender Bedeutung, um das Vertrauen und die Bereitschaft der in der Digitalwirtschaft beteiligten Akteure zu gewinnen. Als neutrales Gremium hat der Verein die Aufgabe, Betrieb und Entwicklung des Datenraums Energie zu überwachen, Ontologien zu pflegen, weiterzuentwickeln und Dateninteroperabilität zwischen verschiedenen Sektoren oder Domänen sicherzustellen.

Ausblick

Die in diesem Projekt gewonnenen Erkenntnisse ermöglichen es erstmalig, einen Datenraum im Energiesektor auf Basis digitaler Infrastruktur und Standards der IDSA in der Schweiz aufzubauen. Es sind jedoch koordinierte Anstrengungen erforderlich, die die Weiterentwicklung des Frameworks von IDSA bezüglich der Bedürfnisse der Schweizer Akteure sicherstellen. Skalierbare Pilotprojekte im Energiesektor sind notwendig, um die identifizierten Herausforderungen praxisorientiert und pragmatisch zu lösen. Nicht zuletzt sind Datenräume ein neuartiges Konstrukt, dessen Konzept noch wenig bekannt ist. Konzept und Vorteile von Datenräumen müssen deshalb stärker kommuniziert werden. Die Schaffung von Datenräumen wie SINA kann für die weitere Entwicklung der Schweizer Datenökonomie, für die Befähigung der Endverbraucher und für die Transformation des gesamten Energiesektors wertvolle Beiträge leisten und sollte weiterverfolgt werden.

Résumé

La nécessité d'atteindre les objectifs ambitieux de décarbonisation, même dans un contexte d'incertitudes géopolitiques, oblige le secteur énergétique suisse à se transformer massivement. La numérisation y contribue largement, car elle permet de mieux gérer une complexité inévitable et les innovations peuvent avoir un effet transformateur. Dans ce contexte, les données numériques, leur disponibilité et leur accessibilité jouent un rôle éminemment important. Les données numériques sont un élément essentiel de la stratégie de l'entreprise, en particulier dans le contexte de la convergence des systèmes d'approvisionnement en électricité, en gaz et en chaleur avec la mobilité et les bâtiments.

Potentiel des systèmes de gestion de l'énergie (SGE) et des données numériques

Rendre les données accessibles, les échanger et les utiliser au-delà des limites actuelles des systèmes offre un grand potentiel dans l'optique de l'application de l'intelligence artificielle et d'autres innovations. La Commission européenne a reconnu depuis longtemps ce potentiel pour l'Europe et encourage donc le développement d'espaces de données dits de confiance - également dans le secteur de l'énergie. Ce n'est que récemment que le Conseil fédéral a décidé de promouvoir les espaces de données en Suisse[5]. L'échange et l'utilisation collaboratifs des données permettent par exemple de réaliser des économies d'électricité de 6 à 9% par la seule sensibilisation aux données de mesure de la consommation[6]. Les présentes études aboutissent pour la première fois à des estimations globales du potentiel des systèmes de gestion de l'énergie (à domicile) (SGE). En supposant que 30% des ménages aient accès à des SGE, il est possible d'économiser chaque année en Suisse au moins 2 TWh d'énergie ou 240'000 tonnes de CO_2 (Figure A, scénario sélectif).

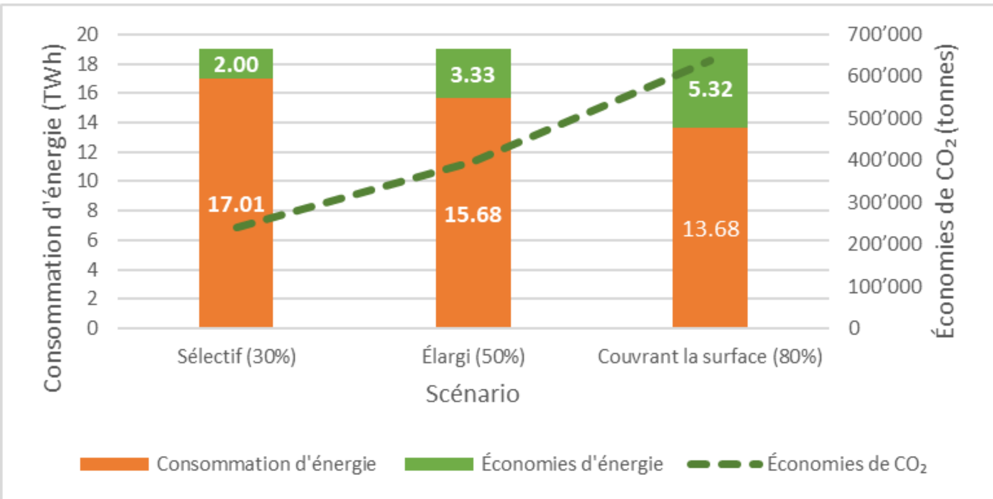

Figure A: Estimation des économies d'énergie et de CO_2 pour trois scénarios concernant l'utilisation de SGE dans un nombre croissant (%) de tous les ménages en Suisse.

5 Promouvoir des espaces de données fiables et l'autodétermination numérique (admin.ch)

6 Le conseil énergétique numérique permet d'économiser 9% (en allemand) (enerlytica.com)

Outre les économies, les SGE permettent également de coordonner de manière optimale la production d'électricité renouvelable avec la mobilité électrique et l'approvisionnement en chaleur via des pompes à chaleur. L'Office fédéral de l'énergie (OFEN) a reconnu l'importance stratégique des SGE pour la transformation du système assistée par le numérique et soutient leur diffusion en réduisant la complexité pour les demandeurs grâce à un aperçu du marché des SGE[7].

Espaces de données pour l'échange de données numériques au-delà des frontières des systèmes et des plateformes

Des instruments réglementaires tels que la mise en place d'une plateforme nationale de données (Datahub) pour l'échange de données de mesure numériques et de leurs agrégats entre les acteurs réglementés du secteur de l'électricité et du gaz doivent permettre de renforcer rapidement la disponibilité et l'accès aux données numériques. L'échange de données entre de nombreux autres acteurs, en particulier dans le secteur privé, reste toutefois un défi. Le besoin d'une infrastructure numérique robuste, interopérable et rentable, permettant de relier les données au-delà des frontières des plateformes et des applications, augmente fortement.

L'approche de la Smart INteroperability Architecture (SINA) concrétisée ici se comprend comme une possibilité de mettre en place un espace de données fiable dans le secteur de l'énergie et de relier ainsi de manière sûre et fiable les acteurs régulés et le hub de données à d'autres stockages de données décentralisés du secteur privé. SINA offre un cadre technique et organisationnel structuré pour l'accès, le traitement et l'utilisation ultérieure de données détenues de manière décentralisée, au-delà des frontières des secteurs réglementés, vers le secteur privé.

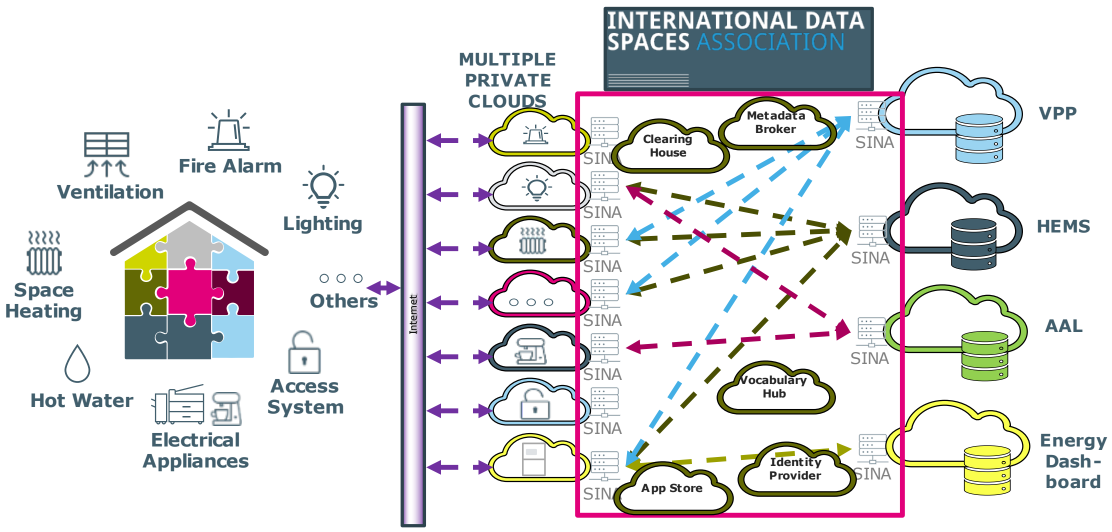

Figure B: Grâce à la connexion des entrepôts de données décentralisés à l'espace de données, les données sont mises à la disposition de différents utilisateurs.

7 Aperçu du marché des systèmes de gestion de l'énergie (SGE) (admin.ch)

L'approche offre une infrastructure ouverte qui permet aux acteurs d'accéder aux données sur un pied d'égalité, sur la base d'un code commun[8] comprenant des accords, des règles et des normes. L'espace de données renonce entièrement à un stockage centralisé des données et à l'installation de matériel périphérique supplémentaire. Au lieu de cela, il mise sur des interfaces numériques standardisées (Application Programming Interfaces API) entre les mémoires de données décentralisées de différents acteurs. La Figure B illustre de manière exemplaire la manière dont les stockages de données privés (par ex. stockage dans le nuage) de différents acteurs sont intégrés dans un espace de données par le biais d'un connecteur.

Dans le cadre de SINA, des bases conceptuelles, économiques et écologiques sont en outre créées pour la conception d'un espace de données fiable dans le secteur de l'énergie, à l'exemple de cas d'application sélectionnés dans le domaine du bâtiment. L'élaboration et l'évaluation de différents cas d'application dans les écosystèmes Building Information, Smart Home and Comfort, ainsi que Energy and Grid esquissent le potentiel de l'espace de données énergie (Figure C). Les cas d'utilisation ont été identifiés en collaboration avec des experts de l'industrie. Il s'avère que de nombreux cas d'application semblent déjà réalisables aujourd'hui.

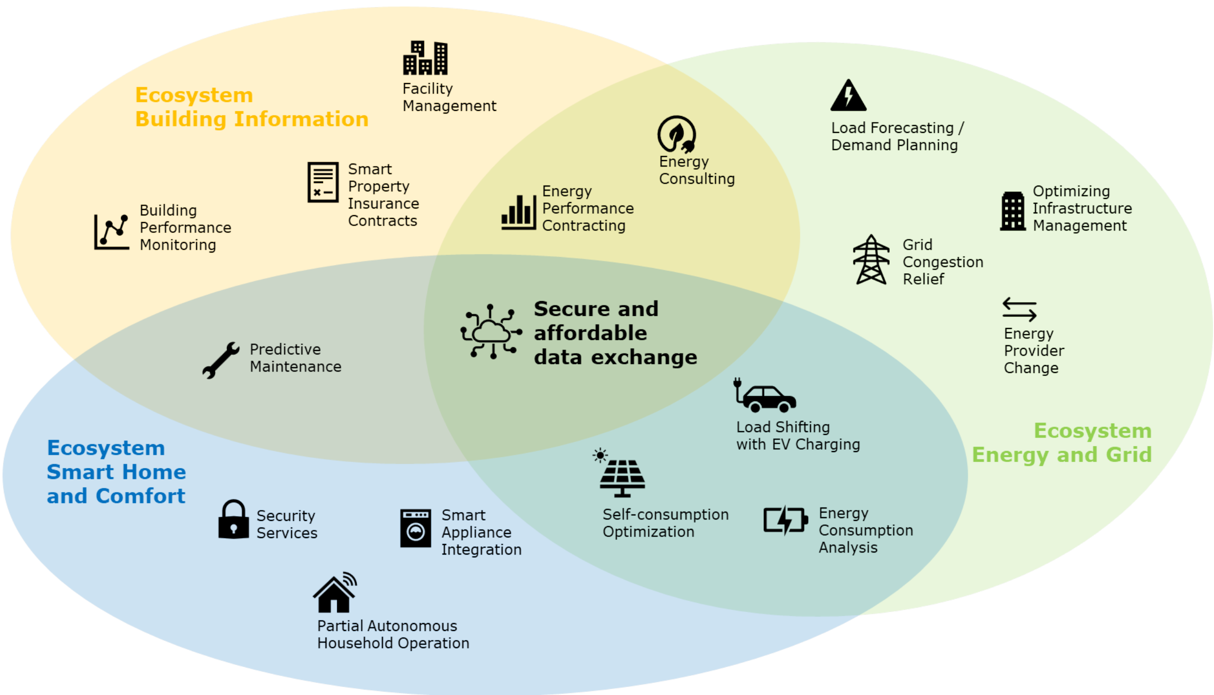

Figure C: Les trois écosystèmes Building Information, Smart Home and Comfort, et Energy and Grid et les principaux cas d'utilisation identifiés.

8 Code de conduite pour l'exploitation des espaces de données fiables (admin.ch)

Les principes conceptuels et techniques les plus importants pour la création d'un espace de données dans le secteur de l'énergie s'inspirent des principes de l'International Data Space Association (IDSA), qui a déjà réalisé de nombreux travaux préparatoires au niveau européen. Dans le cadre de SINA, les aspects techniques et organisationnels sont analysés de manière approfondie et appliqués à un domaine d'application. Pour la première fois, l'importance des ontologies existantes, nécessaires à une gestion efficace des données et à une utilisation interopérable des données entre différents domaines, par exemple entre le bâtiment et la mobilité, est mise en lumière.

L'analyse ne reste pas à un niveau théorique. Au contraire, une étude de faisabilité appliquée démontre comment un espace de données fiable peut être mis en œuvre et réalisé techniquement avec un modèle de communication peer-to-peer (P2P), en tenant compte des normes actuelles. En outre, le lien entre l'espace de données et une application énergétique comprenant des contrats et des transactions intelligents basés sur la blockchain sera démontré.

Les travaux démontrent qu'il est possible de mettre en place et d'exploiter un espace de données fiable en utilisant les bases techniques disponibles de l'IDSA. L'espace de données mis en place ici réalise a) la preuve de l'élimination du stockage redondant des données ; b) la préservation de la souveraineté des données en permettant aux propriétaires des données d'attribuer et de retirer les droits d'accès de manière autonome ; c) la garantie de l'authenticité vérifiable de l'identité des participants ; et d) la décentralisation et l'orchestration de la gestion des données. Des ontologies sont utilisées afin de permettre une utilisation sans barrières des données. Les ontologies telles que SAREF ou EEBus sont des vocabulaires structurés qui sont intégrés dans l'espace de données au moyen d'un Vocabulary Hub afin de permettre l'interopérabilité des données dans tous les secteurs.

Les défis actuels des espaces de données de confiance

Parallèlement, divers défis sont identifiés lors de cette première réalisation d'un espace de données. Les ontologies sont souvent gérées par des «communautés», sont développées en permanence, sont souvent difficiles à comprendre ou complexes et ne sont pas toujours disponibles sous une forme utilisable numériquement. Pour le secteur de l'énergie, il existe une multitude d'ontologies qui ne sont pas gérées de manière centralisée. Le composant logiciel pour un Vocabulary Hub est certes disponible, mais encore dans une version précoce. En outre, la documentation est encore peu abondante, ce qui rend l'intégration dans un espace de données difficile.

Pour les fabricants - par exemple d'appareils tels que des onduleurs -, l'effort d'intégration de leurs espaces de données privés par un connecteur est encore élevé, car il n'existe pas d'implémentation de référence. De plus, la gestion des certificats est liée à des dépenses. La recherche de solutions en cas de besoin de données en temps réel n'est pas encore claire non plus.

Des espaces de données sont actuellement réalisés dans de nombreux projets, le plus souvent au niveau européen. On constate que la plupart du temps, un nouvel espace de données est créé pour chaque projet et qu'il n'est pas fait appel à un espace existant. Dans sa stratégie européenne en matière de données, l'Union européenne a proposé des espaces de données communs (Common Data Spaces) pour différents domaines. Dans un espace de données commun, toutes les données liées à un domaine, comme par exemple l'énergie ou la mobilité, doivent pouvoir être échangées. Les espaces de données communs sont toutefois encore trop peu avancés, de sorte qu'ils ne constituent pas encore une alternative valable dans les projets pour le moment. Le défi consiste ici à faire progresser ces espaces de données communs et à les faire connaître.

Considérations juridiques et organisationnelles sur les espaces de données

Lors de la mise en œuvre d'espaces de données, différentes lois doivent être respectées, notamment en ce qui concerne les questions de droit des données, comme le Règlement général sur la protection des données (RGPD) de l'UE, ou la loi fédérale sur la protection des données (LPD). Cela augmente la complexité et la charge de travail. Le cadre de l'IDSA offre des solutions à ce problème, car il est conforme aux dispositions de l'Union européenne en matière de protection des données et est continuellement adapté aux modifications du cadre juridique. Il simplifie donc considérablement la mise en place et l'exploitation d'un espace de données. Néanmoins, en Suisse, il faut tenir compte des spécificités de la LPD.

L'examen des aspects organisationnels de l'espace de données met en évidence la nécessité de mettre en place un écosystème de soutien. Une organisation à but non lucratif digne de confiance, par exemple une association, peut garantir une gestion et un développement de l'espace de données basés sur le consensus. Une gouvernance commune et largement soutenue est essentielle pour gagner la confiance et la volonté des acteurs impliqués dans l'économie numérique. En tant qu'organe neutre, l'association a pour mission de surveiller le fonctionnement et le développement de l'espace de données Énergie, de maintenir et de développer les ontologies et d'assurer l'interopérabilité des données entre différents secteurs ou domaines.

Perspectives

Les connaissances acquises dans le cadre de ce projet permettent pour la première fois de mettre en place un espace de données dans le secteur de l'énergie en Suisse sur la base de l'infrastructure numérique et des normes de l'IDSA. Cependant, des efforts coordonnés sont nécessaires pour assurer le développement du cadre de l'IDSA par rapport aux besoins des acteurs suisses. Des projets pilotes évolutifs dans le secteur de l'énergie sont nécessaires pour résoudre les défis identifiés de manière pratique et pragmatique. Enfin, les espaces de données sont une construction récente dont le concept est encore peu connu. Le concept et les avantages des espaces de données doivent donc être mieux communiqués. La création d'espaces de données tels que SINA peut apporter une contribution précieuse à la poursuite du développement de l'économie suisse des données, à l'autonomisation des consommateurs finaux et à la transformation de l'ensemble du secteur énergétique, et devrait être poursuivie.

Summary

The need to achieve the ambitious decarbonization targets, even against the backdrop of geopolitical uncertainties, is forcing the Swiss energy sector to undergo a massive transformation. Digitalization is making a major contribution to this by enabling the inevitable complexity to be better managed and innovations to have a transformative effect. Digital data, its availability and accessibility play an eminently important role here. This is particularly true because electricity, gas and heating supply systems are converging with mobility and buildings.

Potential of energy management systems (HEMS) and digital data

Making data accessible, exchanging, and utilizing it beyond current system boundaries offers great potential with regard to the application of artificial intelligence and other innovations. The European Commission has long recognized this potential for Europe and is therefore promoting the development of so-called trusted data spaces - including in the energy sector. The Federal Council made a decision to also promote data spaces in Switzerland[9]. Collaborative data exchange and data utilization were found to enable, for example, electricity savings of 6-9% simply by raising awareness of consumption measurement data[10]. These enquiries are the first to provide comprehensive estimates of the energy saving potential of (home) energy management systems (HEMS). Assuming that 30% of all households have access to HEMS, at least 2 TWh of energy or 240,000 tons of CO_2 can be saved annually in Switzerland (Figure A, selective scenario).

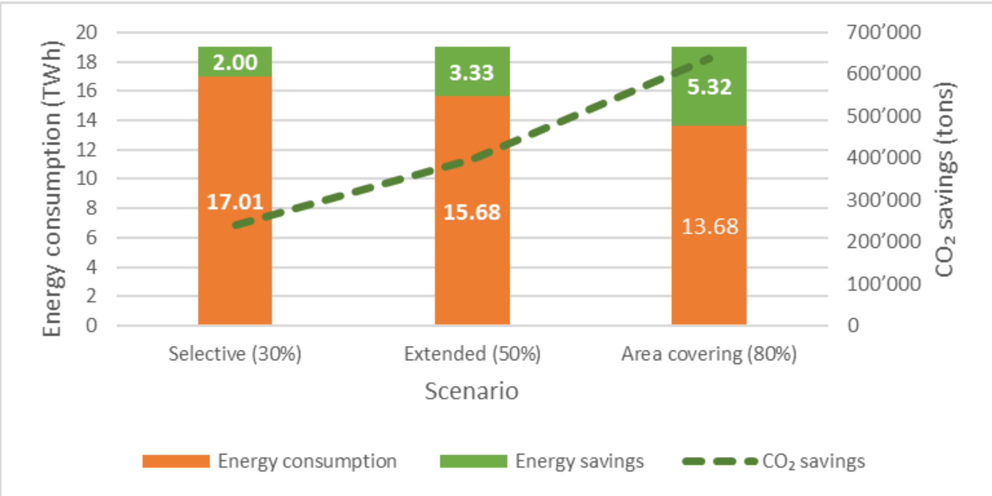

Figure A: Estimation of energy and CO_2 savings for three scenarios regarding the use of HEMS in increasing numbers (%) of all households in Switzerland

9 Promotion of trustworthy data spaces and digital self-determination (admin.ch)

10 Digital energy consulting leads to savings of 9% (in German) (enerlytica.com)

In addition to savings, HEMS also make it possible to optimize the coordination of renewable electricity production with electromobility and heat supply via heat pumps. The Swiss Federal Office of Energy (SFOE) has recognized the strategic importance of HEMS for digitally supported system transformation and supports their dissemination by reducing complexity for consumers through providing a market overview of HEMS[11].

Data spaces for exchanging digital data across system and platform boundaries

Regulatory instruments such as the establishment of a national data platform (data hub) for the exchange of digital measurement data and its aggregates among the regulated players in the electricity and gas sector are intended to increase the availability of and access to digital data in the near future. Still, the exchange of data between many other players, particularly in the private sector, remains a challenge. Nevertheless, the need for a robust, interoperable, and cost-efficient digital infrastructure that enables data to be linked across platform and application boundaries is growing rapidly.

The Smart INteroperability Architecture (SINA) approach described here is intended as a way of establishing a trustworthy data space in the energy sector and thus linking regulated players and the data hub with other decentralized data repositories in the private sector in a secure and trustworthy manner. SINA provides a structured technical and organizational framework for the access, processing, and subsequent use of decentralized data across the boundaries of regulated sectors into the private sector.

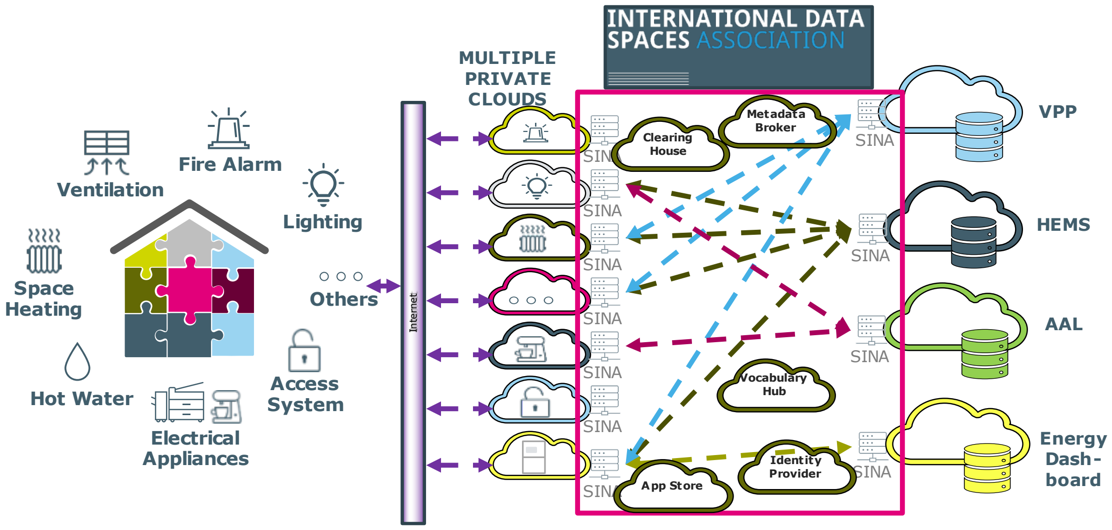

Figure B: By connecting decentralized data storage to the data space, the data is made available to different users

---

11 Market overview of energy management systems (EMS) (admin.ch)

The approach provides an open infrastructure that enables actors to access data on an equal footing based on a common code of conduct[12], rules and standards. The data space completely eliminates the need for a centralized data storage and the installation of additional peripheral hardware. Instead, it relies on standardized, digital interfaces (Application Programming Interfaces API) between the decentralized data repositories of various players. Figure B shows an example of how private data storage (e.g., cloud storage) used by various actors is integrated into a data space via a connector.

Furthermore, SINA will create conceptual, economic, and ecological foundations for the design of a trustworthy data space in the energy sector using the example of selected use cases in the building sector. The development and evaluation of various use cases in the Building Information, Smart Home and Comfort, and Energy and Grid ecosystems outlines the potential of the energy data space (Figure C). The use cases were identified in collaboration with experts from the industry. As can be seen, many use cases already appear feasible today.

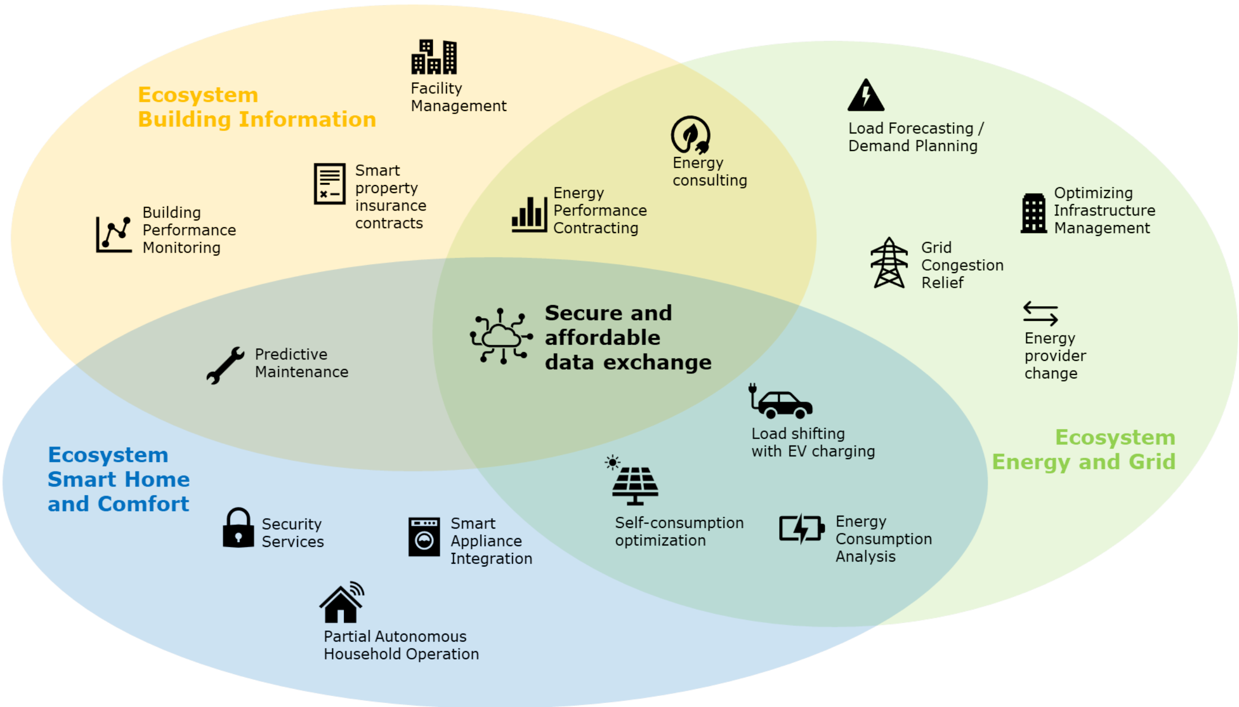

Figure C: The three ecosystems Building Information, Smart Home and Comfort, and Energy and Grid and the most important of the associated, identified use cases

The most important conceptual and technical principles for creating a data space in the energy sector are based on the principles of the International Data Space Association (IDSA), which has already carried out a great deal of preliminary work at European level. Within the framework of SINA, technical and organizational aspects are analyzed in depth and are utilized in the application area. For the first time,

12 Code of conduct for operating trustworthy data spaces (admin.ch)

the importance of existing ontologies, which are necessary for efficient data management and the interoperable use of data between different domains, e.g., between buildings and mobility, are examined in more detail.

The analysis does not remain at a theoretical level. Rather, an applied feasibility study is used to demonstrate how a data space can be technically implemented and realized with a peer-to-peer (P2P) communication model, taking current standards into account. In addition, the linking of the data space and an energy application is demonstrated, which includes intelligent, blockchain-based contracts and transactions.

The work demonstrates that it is possible to set up and operate a trustworthy data space using the available technical foundations of the IDSA. The data space built here enables a) the proof of elimination of redundant data storage; b) the preservation of data sovereignty by enabling data owners to autonomously grant and revoke access permissions; c) the assurance of verifiable authenticity of participant identity; and d) the decentralization and orchestration of data management. Ontologies are used to enable barrier-free use of the data. Ontologies such as SAREF or EEBus are structured vocabularies that are integrated into the data space using a vocabulary hub to enable cross-sector data interoperability.

Current challenges of trustworthy data spaces

Various challenges are identified during this first realization of a data space. Ontologies are often maintained by «communities», are constantly being further developed, are often difficult to understand or complex and are not always available in a digitally usable form. There are many ontologies for the energy sector that are not managed centrally. Although the software component for a vocabulary hub is available, it is still in an early stage of its development. In addition, the documentation is still sparse, which makes integration into a data space more difficult.

The effort needed from manufacturers - for example of devices such as inverters - to integrate their private data storage devices via a connector is still high because there is no reference implementation. In addition, the administration of certificates is time-consuming. There is also no clear pathway to finding a solution for real-time data requirements.

Data spaces are currently being implemented in many projects, mostly at European level. It should be noted that, rather than using an existing one, a new data space is usually set up for each project. In its European data strategy, the European Union has proposed common data spaces for various domains. All data related to a domain such as energy or mobility should be able to be exchanged in a common data space. However, common data spaces are not yet sufficiently advanced, meaning that they are not yet a viable alternative in projects. The challenge here is to further promote and publicize these shared data spaces.

Legal and organizational considerations for data spaces

When implementing data spaces, various laws must be complied with, particularly with regard to data protection issues such as the EU's General Data Protection Regulation (GDPR) or the Swiss Federal Act on Data Protection (FADP). This increases complexity and calls for more effort. The IDSA framework offers solutions here, as it is compliant with the data protection regulations of the European Union and is continuously adapted to changes in the legal framework. This makes it much easier to set up and operate a data space. Nevertheless, the specifics of the FADP must be taken into account in Switzerland.

The examination of the organizational aspects of the data space illustrates the need to establish a supporting ecosystem. A trustworthy non-profit organization, such as an association, can ensure consensus-based management and further development of the data space. Joint, broad-based governance is crucial in order to gain the trust and willingness of the players involved in the digital economy. As a neutral body, the association has the task of monitoring the operation and development of the Energy Data Space, maintaining and further developing ontologies and ensuring data interoperability between different sectors or domains.

Outlook

The knowledge gained in this project makes it possible for the first time to establish a data space in the energy sector based on IDSA's digital infrastructure and standards in Switzerland. However, coordinated efforts are required to ensure the further development of the IDSA framework to meet the needs of Swiss stakeholders. Scalable pilot projects in the energy sector are necessary to solve the identified challenges in a practical and pragmatic manner. Finally, data spaces are a new construct whose concept is still little known. The concept and benefits of data spaces therefore need to be communicated more effectively. The creation of data spaces such as SINA can make valuable contributions to the further development of the Swiss data economy, to the empowerment of end users and to the transformation of the entire energy sector and should be pursued further.

Content

Table of figures

Table of tables

List of abbreviations

ABAC	Attribute-Based Access Control
API	Application Programming Interface
AuthN	Authentication
AuthZ	Authorization
BCT	Blockchain Technology
BIM	Building Information Management
BDVA	Big Data Value Association
BLC	Broadband Loop Carrier
CA	Certificate Authority
CIM	Common Information Model
COSEM	Companion Specification for Energy Metering
CRIM	Conceptual Representation of the Information Model
CSR	Certificate Signing Requests
CSV	Comma-Separated Values
DA-E	Data Act
DAPS	Dynamic Attribute Provisioning Service
DAT	Dynamic Attribute Tokens
DEP	Data Exchange Program
DGA	Data Governance Act
DBMS	Data Base Management System
DLMS	Device Language Message Specification
DLT	Distributed Ledger Technology
DSA	Digital Services Act
DRIM	Declarative Representation of the Information Model
DCAT	Data Catalog Vocabulary
DMA	Digital Markets Act
EDC	Eclipse Dataspace Connector
EF	Eclipse Foundation
EEPSA	Energy Efficiency Prediction Semantic Assistant
EN	European Standards
ETSI	European Telecommunications Standards Institute
EMHAS	Energy Management for Home Assistant
EU	European Union
EV	Electric Vehicle
FADP	Federal Act on Data Protection
GDPR	General Data Protection Regulation
GEE	Gebäude-Elektroengineering
HEMS	Home Energy Management Systems
HTTP	Hypertext Transfer Protocol

HSLU	Lucerne University of Applied Sciences and Arts
IAM	Identity and Access Management
IEC	International Electrotechnical Commission
IDS	International Data Space
IDSA	International Data Space Association
IDS-RAM	International Data Space Association's Reference Architecture Model
IDC	International Dataspace Connector
IRI	Internationalized Resource Identifier
IoT	Internet of Things
IP	Internet Protocol
JSON	JavaScript Object Notation
MVDS	Minimum Viable Data Space
MQTT	Message Queuing Telemetry Transport
OBIS	Object Identification System
OCPP	Open Charge Point Protocol
ODRL	Open Digital Rights Language
OEM	Original Equipment Manufacturer
OWL	Web Ontology Language
ParIS	Participants Information Service
PV	Photovoltaics
P2P	Peer-to-Peer
PRIM	Programmatic Representation of the Information Model
RAM	Reference Architecture Model
RBAC	Role-Based Access Control
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
RES	Renewable Energy Sources
SGready	Smart Grid Ready
SINA	Smart INteroperability Architecture
SITRA	The Finnish Innovation Fund
SunSpec	SunSpec Alliance
TCP	Transport Control Protocol
TPM	Trusted Platform Modules
UDP	User Datagram Protocol
UI	User Interface
UML	Unified Modelling Language
URI	Uniform Resource Identifier
USEF	Universal Smart Energy Framework
VPP	Virtual Power Plant
vHEMS	Virtual Household Energy Management System

W3C	World Wide Web Consortium
XML (.xsd)	Extensible Markup Language

1 Introduction

The energy sector is undergoing a profound transformation, driven by factors such as the energy market liberalization, the growing decentralization and integration of renewable energies, and the imperative to achieve ambitious decarbonization targets. Recent geopolitical events, particularly the Russian war in Ukraine, have heightened the urgency to accelerate the shift toward sustainable energy sources and reduce dependence on fossil fuels. In navigating these challenges, digitalization has emerged as a crucial enabler for the energy sector's metamorphosis.

Digitalization is imperative for the comprehensive transformation of the energy sector. However, despite its pivotal role, some critical aspects remain insufficiently addressed: the centrality of data, data infrastructures, and equitable data exchange relationships among industry stakeholders. For digitalization to realize its potential and expedite the energy system's transformation, a concerted focus on the role and dynamics of data in the coming years is paramount.

Data exchange in the energy sector is primarily driven by regulatory requirements, often linked to specific processes (such as the switching provider process). While this ensures the flow of information necessary for the existing system's operation, it falls short of harnessing the full potential of data for the industry's future. A shift is needed towards creating frameworks that incentivize and facilitate data exchange beyond regulatory mandates to unlock this potential.

At the heart of this paradigm shift is the need for a robust digital infrastructure that facilitates data exchange and incentivizes it. Various options for implementing such an infrastructure have been discussed in the industry, ranging from data platforms and data trustees to data rooms and data marketplaces. However, the complexity and diversity of use cases demand tailored technical concepts and governance principles. An adequate data infrastructure must meet stringent criteria to gain acceptance from building industry players. These criteria encompass data protection, security, sovereignty, and interoperability. The multifaceted nature of these requirements emphasizes the intricate balance that must be struck to foster a collaborative and secure environment for data exchange.

Creating a suitable data infrastructure necessitates more than entrepreneurial initiatives alone. Establishing the essential trust within the energy sector for data exchange is a multifaceted challenge. Publicly funded projects, exemplified by initiatives like the European «GAIA-X» and the Swiss SINA project, demonstrate promising avenues for building cross-sector data infrastructure. The transparency, co-design opportunities, and clear principles embedded in these initiatives foster an environment conducive to trust-building among stakeholders (Knüsel & Richard, 2022).

In March 2022, the Swiss Federal Council embarked on a pivotal mission: to cultivate a sustainable data society that facilitates more precise and efficient data sharing, and steers Switzerland towards greater social well-being, economic growth, and innovation. Nevertheless, as with any significant transformation, the path to this data-driven future is laden with challenges that demand careful consideration and strategic resolution.

One fundamental challenge arises from the concentration of data within a handful of dominant players across an expanding spectrum of sectors. Within these sectors, the custodians of data often lack the incentive to share these invaluable resources. The consequence is that a potential wealth of insights and opportunities remains untapped, locked behind closed digital doors.

Another obstacle to fully realizing Switzerland's digital potential lies in the underutilization of data sharing by private and public service providers. Factors such as a lack of expertise, limited resources, apprehensions about weakening market positioning, and the web of administrative, technical, and legal complexities hamper the harnessing of data's transformative power. Consequently, substantial reservoirs of actionable insights remain latent, eluding the grasp of those who could benefit most.

Amidst these challenges, distrust looms over how data is harnessed. A significant segment of the population harbors concerns about data manipulation, misuse, the erosion of privacy, and the absence of tangible incentives for data sharing. This mistrust further complicates the endeavor to create a thriving digital ecosystem built on the principles of transparency and empowerment.

1.1 The path forward: Digital self-determination and trustworthy data spaces

In navigating the complex landscape of data utilization, the concept of a «data space» emerges as a fundamental cornerstone. This innovative concept can be defined as a unique data relationship forged among trusted partners, each adhering to a stringent set of high standards and rules governing the storage and sharing of their data. A defining feature of data spaces is the decentralized nature of data storage, where data remains at its source and is only shared when deemed necessary, facilitated by semantic interoperability (Pettenpohl et al., 2022).

This definition aligns with ongoing initiatives within Europe and across the global landscape, aiming to lay the foundations for data spaces tailored to specific sectors such as energy, mobility, manufacturing, and more. Data spaces are central to the European strategy for data, envisioning a unified European data space that functions as a robust single market for both personal and non-personal data. The objective is to ensure the security of data, including sensitive business data, while fostering easy access to high-quality industrial data, thereby catalyzing growth, and creating substantial value (European Commission, 2019). The European strategy for data envisions the development of an initial set of nine sectoral data spaces to foster a unified European data ecosystem (Figure 1).

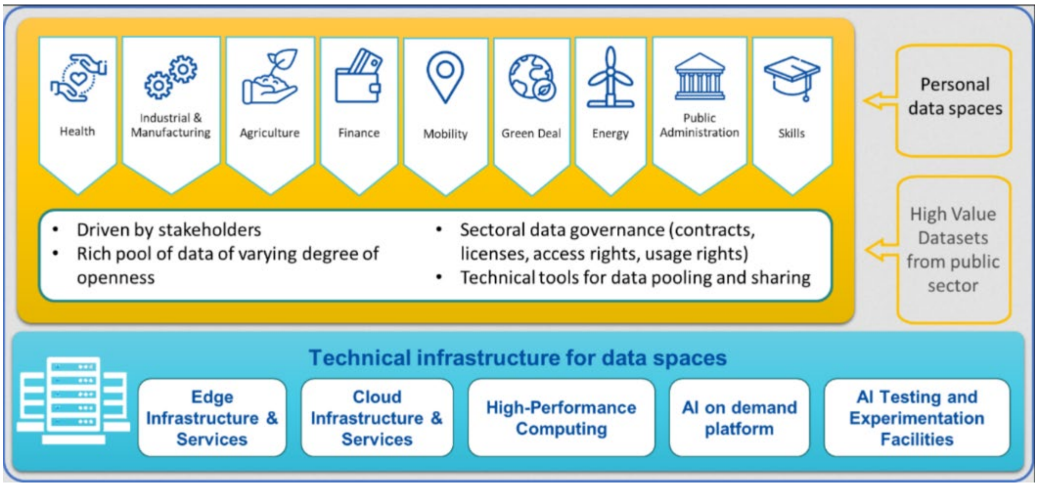

Figure 1 Sectors intended by the European Commission to develop a data space. Source: European Commission (2019)

These data spaces are designed to address critical sector-specific needs and opportunities, including enhancing industrial competitiveness, advancing sustainability goals through the Green Deal, promoting intelligent mobility systems, supporting healthcare advancements, driving financial innovation and transparency, enabling secure and customer-centric energy data sharing, improving agricultural sector sustainability, enhancing public administration transparency, and addressing skills mismatches in the labor market. These data spaces are conceived as foundational pillars for facilitating data-driven growth, innovation, and cross-sector collaboration within the European Union and beyond.

The Big Data Value Association (BDVA) has embraced an open-innovation model, fostering collaboration between academia and industry to conceptualize the Data Sharing Value Wheel, a pivotal framework depicted in Figure 2 .

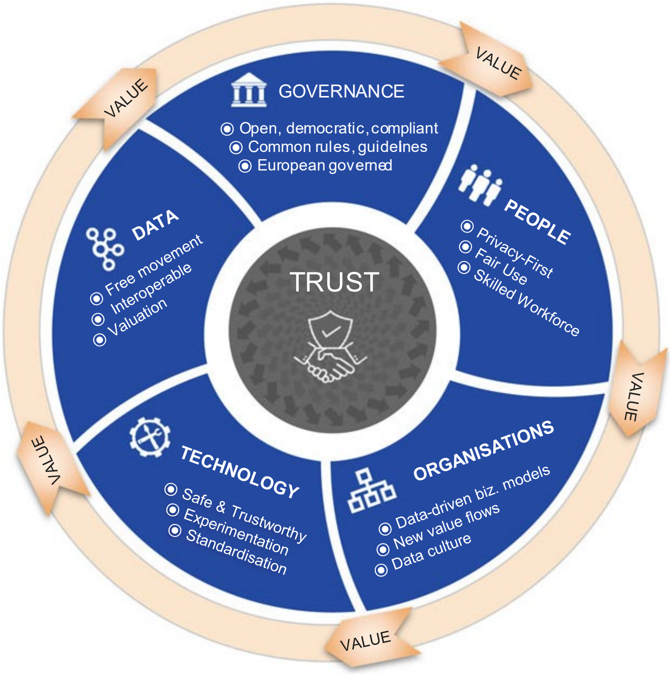

Figure 2 Data Sharing Value Wheel. Source: Šestak & Copot (2023)

At its core, the success of broad data-sharing endeavors orbits a central tenet: trust. Trust must be instilled in multiple facets of the data-sharing landscape, encompassing the reliability of the data itself, the algorithms processing it, the governing entities responsible for data spaces, the enabling technologies underpinning the ecosystem, and the diverse array of users, both organizations and individuals, who serve as data producers, consumers, and intermediaries. Achieving the requisite levels of trust necessitates the alignment of five key pillars, each with its own set of prerequisites (Šestak & Copot, 2023):

– Organizations: A comprehensive shift in organizational strategies, including businesses, research institutions, and government entities, must cultivate a data-centric culture where data takes center stage in its value proposition. This shift should explore innovative data-driven business models and harness the emerging data value streams.

- Data: As the heralded «5th European fundamental freedom», the free flow of data relies upon organizational data strategies that inherently incorporate data-sharing methodologies, such as interoperability. Apparent and standardized guidelines are pivotal in ascertaining the market value of data assets.
- Technology: To catalyze the maturation of dependable data technologies, data access mechanisms, and algorithms - touching upon aspects of privacy, interoperability, security, and quality - a need arises for secure experimentation environments. Additionally, standardization initiatives should adapt to ensure swift responses to emerging standards and the identification of novel ones.
- People: In data sharing, individual privacy must be safeguarded, and the equitable compensation of shared personal data must be assured. To drive data-sharing endeavors in Europe, the workforce should undergo appropriate reskilling and upskilling initiatives to meet the dynamic demands of the labor market.
- Governance: Establishing a European-governed data-sharing space can inspire trust by adhering to advanced European rules, guidelines, and regulations, thereby promoting European values. Participation in this governance framework should be open to all, underpinned by transparent and equitable rules of engagement.

1.2 Recommendation by Swiss Data Alliance

The Swiss Data Alliance emphasizes that to ensure its qualitative integration into emerging European data spaces, Switzerland must strategically position itself along the value chain as both a contributor and beneficiary within the data economy (Swiss Data Alliance, 2023). This necessitates prompt and comprehensive actions across various stakeholders, further elaborated in the following sections.

The entities within the data economy, encompassing companies and public authorities as data producers and consumers should expeditiously assess and implement the potential offered by data spaces and marketplaces. This entails exploring avenues such as developing valuable data sets, providing infrastructure for data spaces or marketplaces, and fostering innovative analytical competences. Such initiatives are pivotal for Switzerland to engage in the evolving landscape of data utilization actively.

Simultaneously, legislators are urged to formulate an adept legal framework conducive to these activities, emphasizing sustainable models and principles for forward-looking data governance. Recognizing the pivotal role of data governance in shaping the future data economy, this initiative requires collaboration and support from civil society and the Swiss public. Switzerland, with its well-established culture of debate, liberal tradition, and straightforward legislative practices, is well-positioned to assume a leadership role in this arena.

Furthermore, leveraging its reputation as a trustworthy and neutral partner, Switzerland can potentially excel in providing infrastructures that facilitate secure sector-specific data exchange. Politicians and the administration should encourage the responsible bodies, including authorities and companies, to collaborate with EU countries in establishing and operating sector-specific data spaces. The temporal aspect is crucial, necessitating prioritized evaluations and actions to capitalize on Switzerland's potential as a first mover in co-designing evolving data structures. Failing to do so risks leaving Switzerland's advantage in contributing to and shaping the emerging data landscape untapped.

1.3 A focus on the building industry

While the European Commission and the Swiss Federal Council's vision encompass a wide array of critical sectors, this report will put particular emphasis on the building industry as a sector of paramount importance. Within the building industry, sharing high-quality data holds immense potential to accelerate Switzerland's pursuit of ambitious sustainability goals, including achieving net-zero emissions. Data spaces, tailored to the building industry's unique needs, can revolutionize operations, enabling precise planning of construction projects, meticulous tracking of energy consumption, and most notably, empowering consumers to shape the sustainability profile of their built environments actively.

This section delves into the digitalization challenges of Switzerland's energy transition and their far-reaching implications. One of the primary challenges is the sheer abundance of data, generated by the proliferation of sensors and smart devices. This demands efficient management, analysis, and protection of data. Moreover, Switzerland's stringent data protection regulations, in accordance with the Swiss Federal Data Protection Act and European GDPR, impose meticulous handling of personal and sensitive data. Interoperability issues arise as digitalization necessitates seamless data exchange among stakeholders, calling for standardized formats and interfaces. Establishing trust and transparency is imperative for fostering stakeholder confidence in energy-related data collection and usage. Further, the advancement of digitalization exposes the energy sector to heightened cybersecurity risks, including the looming threat of attacks on critical infrastructure. Therefore, Switzerland's energy transition relies on enabling both digitalization and sustainability, offering promising prospects and formidable challenges.

Digitalization is one of the significant trends in today’s technology, accompanying a shift where digital services that utilize data create new business opportunities in nearly all sectors of the economy (Neugebauer, 2019). Beverungen et al. (2022) suggest that data spaces could act as a new form of digital platforms complementing technological developments, such as Cloud Computing, the Internet of Things (IoT), Big Data, and Artificial Intelligence (AI).

The current developments in the energy sector require more sustainable energy transitions, effective data sharing, using the integration of renewable energy and optimization of energy, among many other drivers, with, for example, IoT and smart energy systems or efficient buildings (Motlagh et al., 2020; Rusitschka & Curry, 2016; Trzaska et al., 2021). The sector deals with complexity and intensive data communication to operate the systems reliably with these dynamic transitions. Therefore, Berkhout et al. (2022) agree that a data space infrastructure can enable the digitalization of the energy transition by making data readily available to increase efficiency during asset and system operation.

In the face of these formidable digitalization challenges, the effective management and integration of data emerge as pivotal factors in Switzerland's energy transition. Leveraging data offers a multitude of advantages in tackling these challenges head-on. Firstly, it enables data-driven decision-making, a cornerstone for optimizing the Swiss landscape's energy production, distribution, and consumption. Secondly, rigorous data management practices ensure adherence to Swiss and European data protection regulations, thus nurturing a climate of trust among consumers. Thirdly, data integration initiatives are pivotal in bolstering interoperability by advocating for standardized data exchange protocols, thereby enhancing overall efficiency in the energy sector. Moreover, cultivating trust and transparency through transparent data practices empowers stakeholders and encourages active consumer participation in the ongoing energy transition. Lastly, robust cybersecurity measures within data management strategies are essential to safeguard critical energy infrastructure against the escalating digital threats posed by cyberattacks.

On the technical side, the current state of the art heavily relies on hardware gateways that must be physically installed on site. Interfaces to the different devices («from PV to the fridge») and data points (e.g., sensors, device controls, gateways) need to be implemented because:

– Application Programming Interfaces (APIs) vary along the devices.
– APIs are often not suited (e.g., lack required performance) for the targeted business case.
– APIs are sometimes even missing.

Gateway-cantered infrastructures require good technical skills to deal with different standards, license agreements and sufficient resources. The total cost of ownership is high as hardware and software must be installed, deployed, and operated separately on-premises for many business cases. An overview of the conventional setup of such a gateway-centered system is depicted in Figure 3.

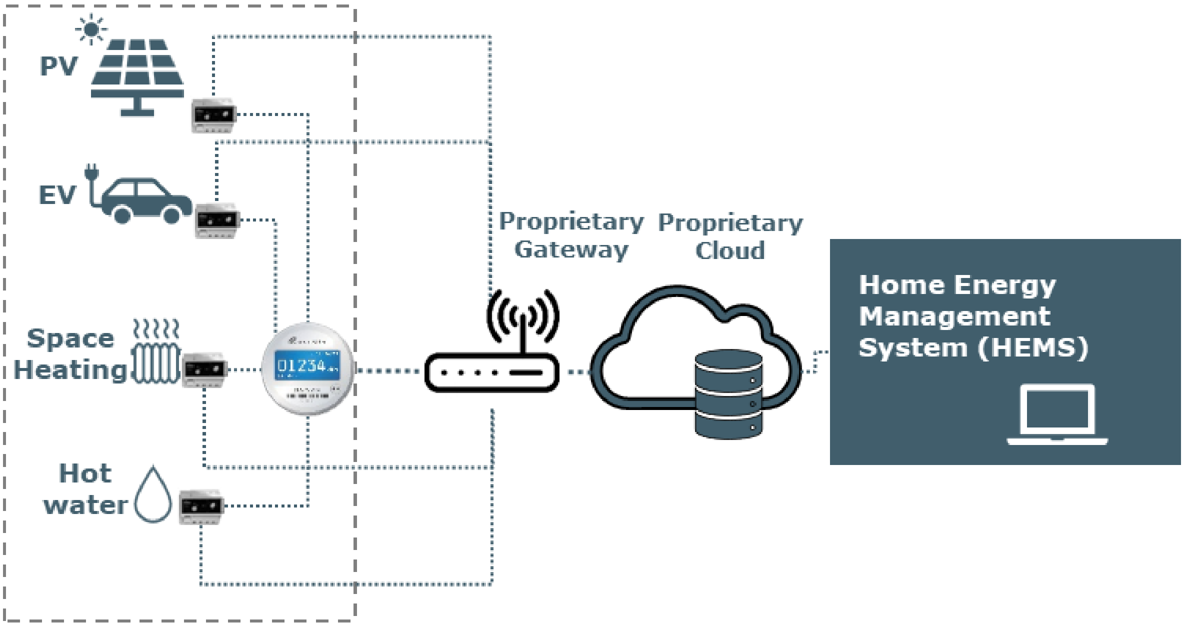

Figure 3 Conventional set-up with the application of home energy management system (HEMS)

IoT business cases lag behind expectations. Most business cases involve data usage and sharing between more than two stakeholders. This drives complexity. Time-intensive negotiations increase transaction costs to a level at which the economic viability is questioned. Alternatively, some companies try to bypass the negotiations by installing expensive infrastructures such as gateways to get access to the required data (Figure 3). Then, questions of technical interoperability between the individual devices and the gateway arise, which are manifold and complex. Small service providers, in particular, lack technical skills and resources for implementation.

1.4 Need for action: Leveraging data space for solutions in the building industry

Amid the digitalization challenges encountered on Switzerland's path towards a sustainable energy transition, the concept of a well-structured data space emerges as a compelling solution to address these issues comprehensively. A data space, in essence, is a platform that facilitates the organized collection, sharing, and utilization of data while ensuring strict adherence to data privacy and security regulations. Figure 4 shows the concept of such a data space solution where the data exchange between various actors and devices is illustrated. The technical concept and building blocks of such a data space solution is further described in section 2.

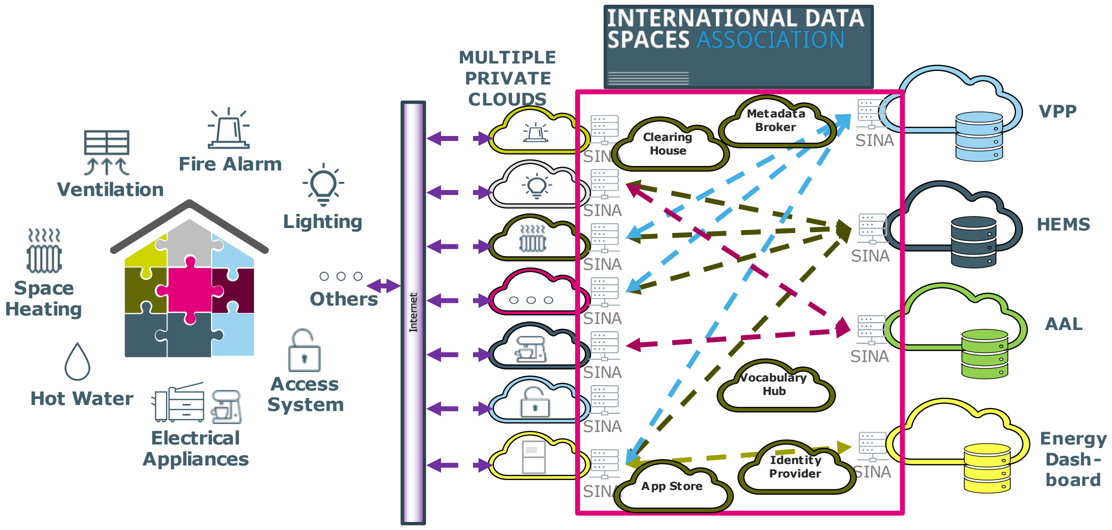

Figure 4 Concept of the SINA data space solution

Leveraging a data space can play a pivotal role in mitigating the challenges mentioned above in the following areas:

a. Data sovereignty: Within data spaces, data providers retain control over their data even during data exchanges, ensuring full sovereignty for the data owner which is a crucial element for successful data-driven business models. Maintaining data sovereignty relies on the necessity of usage control, which governs data access and sharing within the confines of data spaces.

b. Interoperability and standardization: Data spaces promote interoperability by facilitating standardized data exchange formats and interfaces. By enabling seamless communication between different energy systems, stakeholders can collaborate more effectively, resulting in improved efficiency, reduced redundancy, and streamlined operations throughout the energy sector.

c. Data protection compliance: Data spaces are built with data privacy and security at their core. They provide a structured framework ensuring compliance with stringent Swiss and European data protection regulations, instilling confidence among consumers that their personal and sensitive data is handled responsibly. This compliance fosters trust and safeguards against legal and reputational risks.

d. Transparency and consumer engagement: Data spaces encourage transparent data practices, enhancing transparency and trust among stakeholders. By providing consumers with access to their energy usage data, they can actively engage in energy management, make informed choices, and contribute to the broader sustainability goals of Switzerland's energy transition.

e. Enhanced cybersecurity: Data spaces incorporate robust cybersecurity measures to protect critical energy infrastructure from digital threats. These security measures, including encryption, access controls, and real-time threat detection, fortify the energy sector's resilience against cyberattacks, ensuring an uninterrupted energy supply.

f. Data optimization: A well-designed data space efficiently collects and processes diverse data types from various energy sources, providing a comprehensive view of the energy ecosystem. This, in turn, enables data-driven decision-making, facilitating the optimization of energy production, distribution, and consumption in Switzerland. By aggregating and analyzing data in real-time, stakeholders can identify inefficiencies, predict trends, and make informed choices to enhance energy efficiency and reduce waste.

Establishing a well-structured data space is imperative for Switzerland's energy transition, as it offers a holistic solution to tackle the energy sector's digitalization challenges. By harnessing the power of data within a secure and collaborative environment, Switzerland can optimize energy resources, comply with data protection regulations, enhance interoperability, foster trust, and fortify its cybersecurity posture. Implementing a data space aligns seamlessly with the nation's commitment to sustainable energy practices and paves the way for a more efficient and resilient energy future. In order to exploit potential use cases for such a data space, this project focuses on three ecosystems which were derived to cluster them. As it could be seen in Figure 1 many more sectors where data spaces are applicable could be exploited. The three ecosystems which were further analyzed within SINA were: Building Information, Smart Home and Comfort, and Energy and Grid. Each ecosystem significantly benefits from the potential solutions and advancements these integrated data spaces enable.

Building information ecosystem: Realizing building performance improvements with data solutions

Data spaces within the Building Information Ecosystem present opportunities for improved collaboration among various stakeholders in the building process, such as architects, engineers, contractors, and facility managers. Therefore, processes can be streamlined and optimized over the whole lifecycle of a building and interfaces between the different stages such as planning, construction, operation and dismantling can be improved. This can lead to more efficient, sustainable, and resilient management of buildings.

Smart home and comfort ecosystem: Enhancing management and comfort through data-driven solutions

The use cases in this ecosystem, facilitated or improved by data spaces, introduce novel solutions in the area of smart living and smart home. As technology continues to advance, the integration of data spaces could play a pivotal role in shaping the future of smart homes, creating intelligent environments that are responsive, efficient, and tailored to individual preferences. The ongoing evolution of these ecosystems holds great promise for improved energy efficiency, enhanced security, and increase the interoperability of various smart devices and the interface between buildings and grids.

Energy and grid ecosystem: Optimizing stability and performance through data-enabled solutions

In the Energy and Grid Ecosystem, data spaces can offer potential solutions for solving arising challenges in a future smart grid with a large number of distributed energy resources as well as electrification in sectors such as mobility and heating. Therefore, an efficient use of the flexibility potential and the implementation of suitable peak shaving, balancing and grid congestion management strategies will be essential to reduce the need for costly infrastructure investment. Data spaces can contribute to enabling the seamless integration and interoperability between the various devices and actors guaranteeing a secure and standardized data exchange and reducing barriers for innovative solutions to exploit the flexibility potential.

1.5 Project purpose and objectives

The primary purpose of the SINA project is to revolutionize data sharing within the building ecosystem, fostering innovation in IoT business applications. By establishing a robust data environment, SINA aims to enable rapid, cost-effective, and secure data exchange, democratizing the building industry's sensor and data landscape.

The objectives for this initial SINA project are to:

- Establish a dedicated test environment for Proof of Concept (PoC).
- Implement core conceptual aspects based on International Data Space Association's (IDSA) principles for data sharing and interoperability.
- Demonstrate key aspects of the data space concept from a technical, organizational, legal and business perspective.
- Explore peer-to-peer (P2P) communication models and simulate and test energy data communication scenarios within the local energy network.
- Address evolving needs of the energy sector in the digital era.

The SINA project aligns with strategic objectives defined by DETEC & FDFA (2022). Through its multifaceted objectives, SINA aims to drive technological advancements and sustainability within the building industry. In the chapters that follow, we delve deeper into the intricacies of decentralizing data spaces within the building industry, exploring potential solutions and strategies to address the evolving needs of the energy sector in the digital era. The subsequent chapter introduces the core conceptual aspects of SINA, built upon the IDSA's principles for data sharing and interoperability.

2 Technical conceptual aspects

This chapter describes the concept of a SINA solution. Integral to the SINA project are specific requirements:

- Data is not redundantly stored; it remains in its original location.
- Data owners retain control over determining access permissions for their data.
- Participants can trust the identity authenticity of others.
- No centralized entity manages data.

These requirements align with the concept of data spaces. Consequently, the implementation prioritizes identifying available software components for building a data space.

This chapter delves into the technical implementation of data spaces, encompassing a general architecture overview, a specific description of the software components utilized, and a detailed account of the data space implemented within the SINA context. It therefore showcases the viability and functionality of SINA’s concept through a Proof of Concept.

2.1 Data space approach

A data space is a data integration concept based on four pillars (Otto, 2022):

- Data is left where it is created or managed, which results in a distributed data architecture and a distributed data infrastructure.
- There is no common database schema (data integration on the semantic level) required; therefore, vocabularies are needed to achieve semantic interoperability between the distributed data.
- Data spaces allow for data redundancies and «co-existence» of data.
- Data spaces can be nested, i.e., several data spaces can overlap but do not have to be disjoint and participants can be part of multiple data spaces.

In addition, the concept of data spaces comprises data sovereignty and data traceability. Data providers want to determine who can do what with their data, and they want transparency about what happens to their data when they share it. This results in traceability of the transactions. Further, data recipients must be able to trust that the data providers are who they say they are. A trust anchor is therefore needed, which can be achieved through a certification process. A digital certificate is required to document this trust anchor.

Data spaces, as stand-alone data integration entities, may not inherently deliver direct benefits. Data integration's actual value emerges through its application utilization, facilitating innovative business models. This perspective considers all levels, including the application layer, leading to the following insights:

- Application: Imagine, for instance, an innovative web-based energy management service. Such a service necessitates the collaboration of multiple stakeholders, as no single entity possesses access to all requisite data.

- Data space: This is where data spaces come into play, providing a platform where all necessary data can be consolidated and made accessible. Importantly, data spaces accomplish this while respecting data sovereignty, ensuring that data providers maintain control over their information. This approach also fosters interoperable data usage.
- Software architecture: The foundation of data spaces lies in a distributed software architecture. Connectors play a pivotal role in facilitating the provision and utilization of data while upholding data sovereignty. These connectors serve as the distributed data endpoints. Additional components are essential to enable these endpoints to discover each other and engage effectively. Broker services act as intermediaries, connecting data providers with data consumers and facilitating data exchange, ensuring a harmonious match between data supply and demand. Clearinghouse services oversee the successful execution of data transactions without delving into the data itself. The components mentioned here are described in Figure 5. In summary, data spaces unlock the potential for innovative business models by bridging the gap between data integration and application-driven value creation. Through a distributed software architecture and the careful orchestration of connectors, broker services, and clearinghouse services, data spaces empower organizations to harness data while respecting the sovereignty of its providers, thus catalyzing a new era of data-driven innovation.

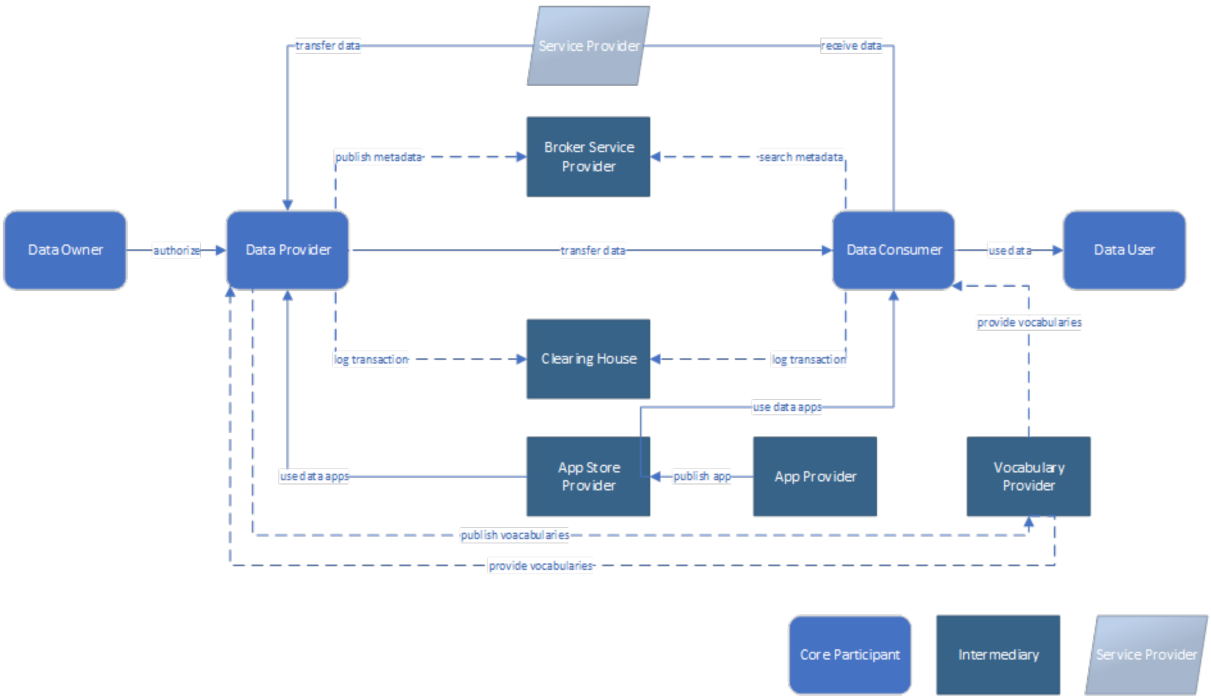

Figure 5 Basic interactions for data exchange and data sharing in the international data spaces

2.2 Evaluation of data space software

The SINA project intends to demonstrate that a data space can be realized. Existing software is to be used as far as possible for this purpose. An evaluation revealed that such software is available from the Eclipse Foundation (EF) and the IDSA. For the project, the evaluation includes the Eclipse Data space Connector (EDC) from EF and the International Data space Connector (IDC) from IDSA. Both fulfil the requirements of International Data Space (IDS). The criteria used for the selection include: the community behind the software, the estimated effort required to build a data space, the maturity of the software, the existence and quality of documentation, supported data formats, the estimated effort required for deployment and the existence of a minimum viable data space. On the results of the comparison with the criteria applied:

– Community: While IDSA lists all members (>150) on its website, the EF website only lists a few partner companies. Some companies are members of both organizations. Overall, support for IDSA appears to be higher.
– Implementation costs: At the start of the project, the implementation costs were estimated based on the available instructions. The estimate was in favor of EF's software.
– Maturity: The maturity of the available software was assessed by analyzing the existing code, particularly the change history of the commitments. The software from IDSA appeared more mature in the comparison. In addition, an initial test installation of a data space with two connectors using IDSA's software was successful, while EF's software could not be made to work.
– Documentation: Very detailed documentation is available for IDSA's software, whereas EF's is less comprehensive.
– Data formats: IDSA's IDC only supports access to local files and to data sources for which access is available via an API. The EDC also supports storage providers and streaming.
– Deployment: The deployment effort is similar for both connectors and is carried out via Docker containers.
– Minimum Viable Data Space: Both organizations provide a minimum viable data space (MVDS). IDSA's MVDS is more extensive and contains more components of a data space than EF's MVDS. In addition, IDSA's roadmap is documented and shows where the MVDS will develop. EF lacks this vision.

As a result of this comparison, the software from IDSA was chosen. The software is open source and managed on GitHub. For each building block there is a responsible project team which manages the repositories. Most of the projects are under Apache 2.0 license.

2.3 Reference architecture model

The International Data Spaces Association offers a software architecture for data spaces with the above-mentioned characteristics. IDSA created a Reference Architecture Model (RAM) that meets all the requirements of SINA. Figure 6 provides an overview of the RAM. The IDSA reference architecture model describes five levels and three perspectives which are described in more detail in Appendix A1. The RAM is the foundation of the demonstrator described in the next section.

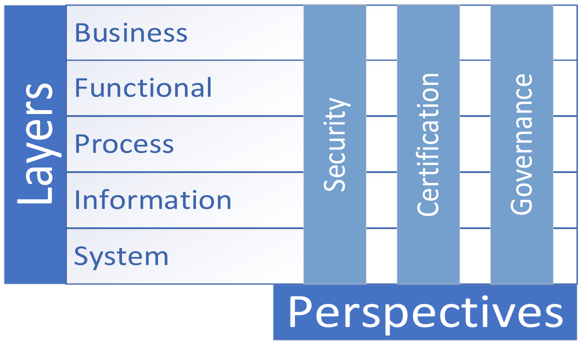

Figure 6 The five layers and three perspectives of the reference architecture model of international data spaces

2.4 Demonstrator

In this section the demonstrator that was built within the SINA project is described.

2.4.1 IDSA's minimum viable data space

The minimum viable data space (MVDS) comprises a minimum set of components for a working data space (Figure 7). The identity provider in MVDS comprises the Certificate Authority (CA) and the Dynamic Attribute Provisioning Service (DAPS). The CA issues the requested certificates for the components in the data space. The two connectors, one acting as a data provider and the other as a data consumer, are exchanging data. The providing connector can register the data set description at the broker; the consuming connector queries the available data sets. A use case for such an MVDS is testing components that a software provider has developed. Nevertheless, it can also be used to become familiar with data spaces.

The SINA data space is built using the MVDS as a base. In the first step, the data space is configured as a testbed that is running as Docker containers on the same Docker host. Consequently, the certificates can be obtained from the integrated CA as self-signed certificates. Even though the data space runs in an internal network, the connectors can access data from any source that provides an API, also from outside of the internal network.

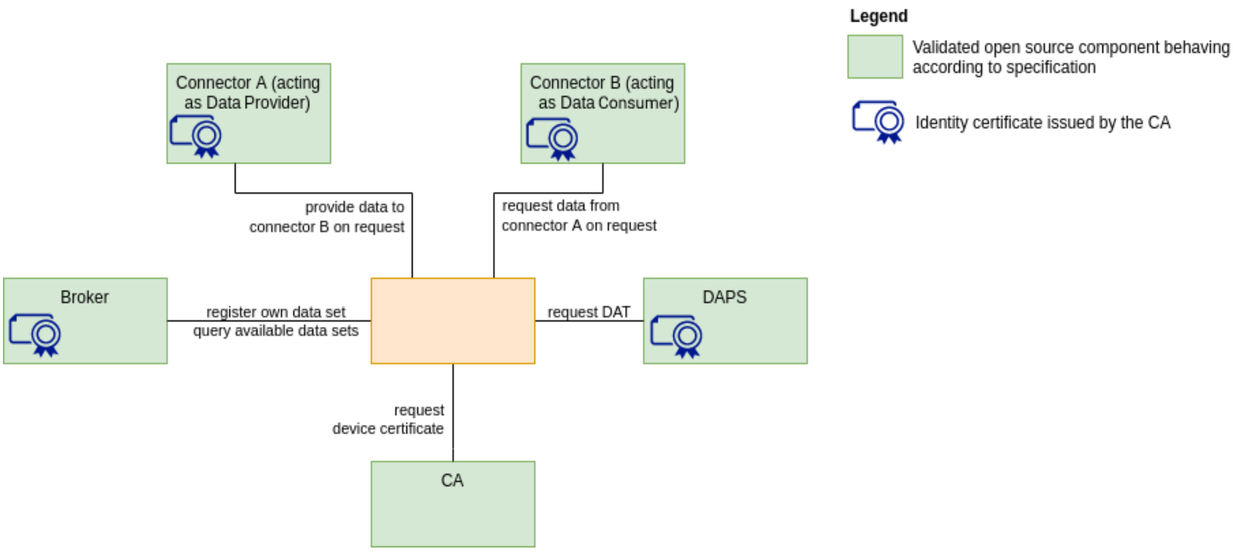

Figure 7 The components of a minimum viable data space with the component under test in the center. CA and DAPS form the Identity Provider. Source: IDSA (2022b)

The components of the demonstrator data space are distributed to a set of containers, which include, besides the mentioned components above, also databases and containers for configuration User Interfaces (UI). All the containers run on the same hardware (Figure 8).

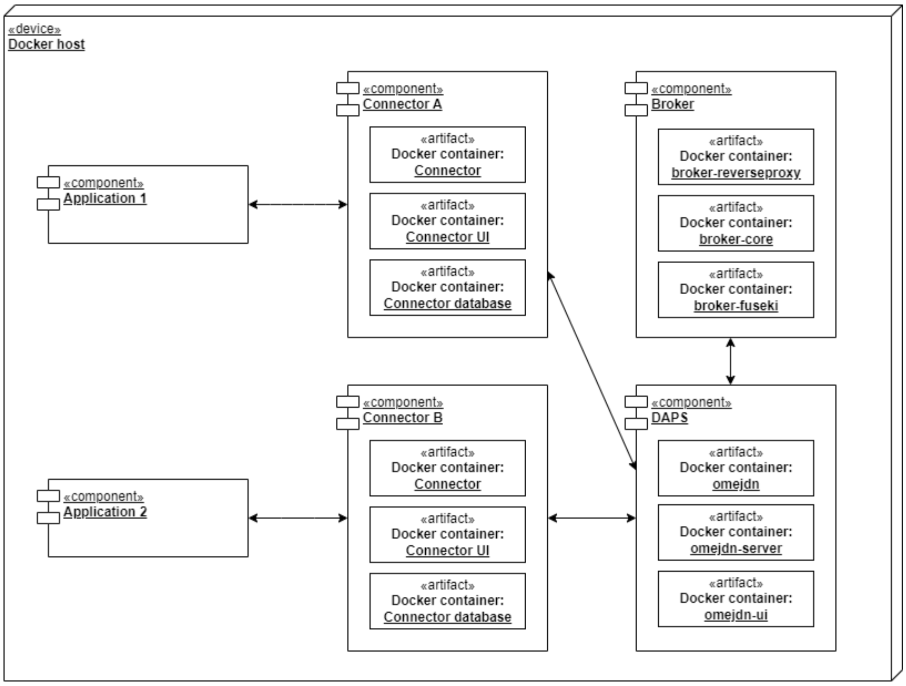

Figure 8 General deployment diagram for the minimum viable data space that is used as a base for the demonstration of the SINA data space running in a local network

After proving the data space in a simple realization as a testbed, in a second step the MVDS is exposed to the public. The deployment changes to containers running on different servers with public IP-addresses (Figure 9).

This step requires the proper handling of certificates that must be signed by an acknowledged certification authority. A public data space does not work with self-signed certificates. The handling of the certificates is one of the challenging issues. It showed that the replacement of the self-signed certificates by official-signed certificates is hard to reach with the MVDS because it is built mainly for testing components.

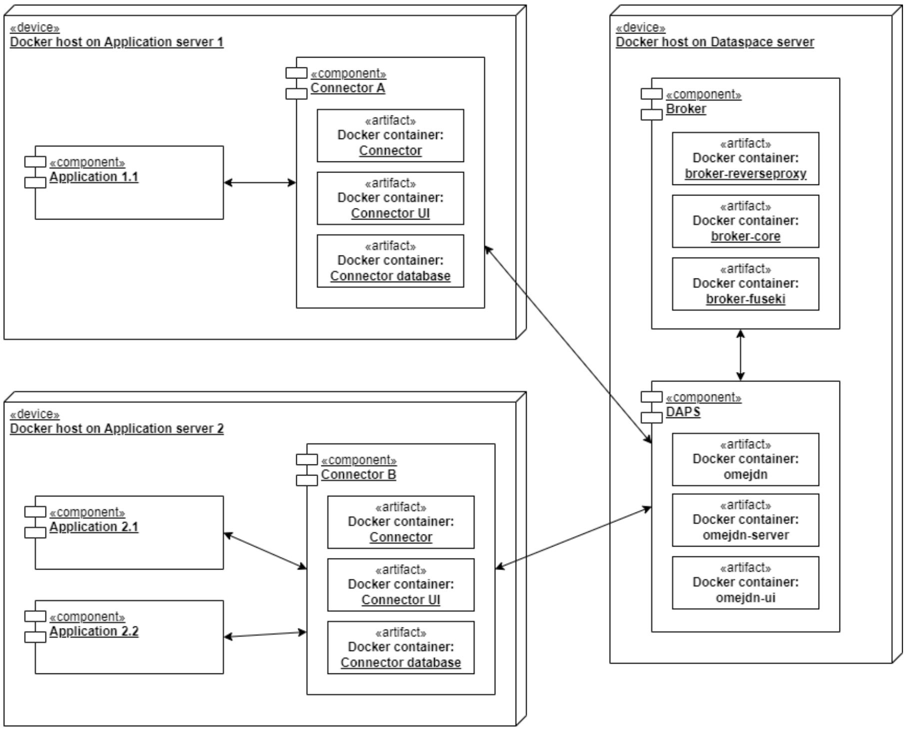

Figure 9 General deployment diagram for a data space that can be reached from public internet

2.4.2 Realization of SINA data space as a demonstrator

As a use case with the demonstrator of the SINA data space, the exchange of data about available energy from a PV system is shown (Figure 10). In this example one connector of the demonstrator connects to a data source from a vendor of solutions for optimized self-consumption via the credentials of a known user. The same connector also connects to a weather service to obtain actual weather information. The second connector retrieves data from the first connector and processes the data in an app. In the demonstrator, this app is not running in the context of the connector but on a separate server.

With the demonstrator it is shown how an owner of a PV system can offer his surplus energy on a peer-to-peer energy trading platform. Not only the implementation aspects are demonstrated (see section 6) but it is also shown how data is offered and that data exchange has to follow a data usage policy defined by the user. A data usage policy can restrict the duration of availability of data, e.g., that data is only available during a certain time span.

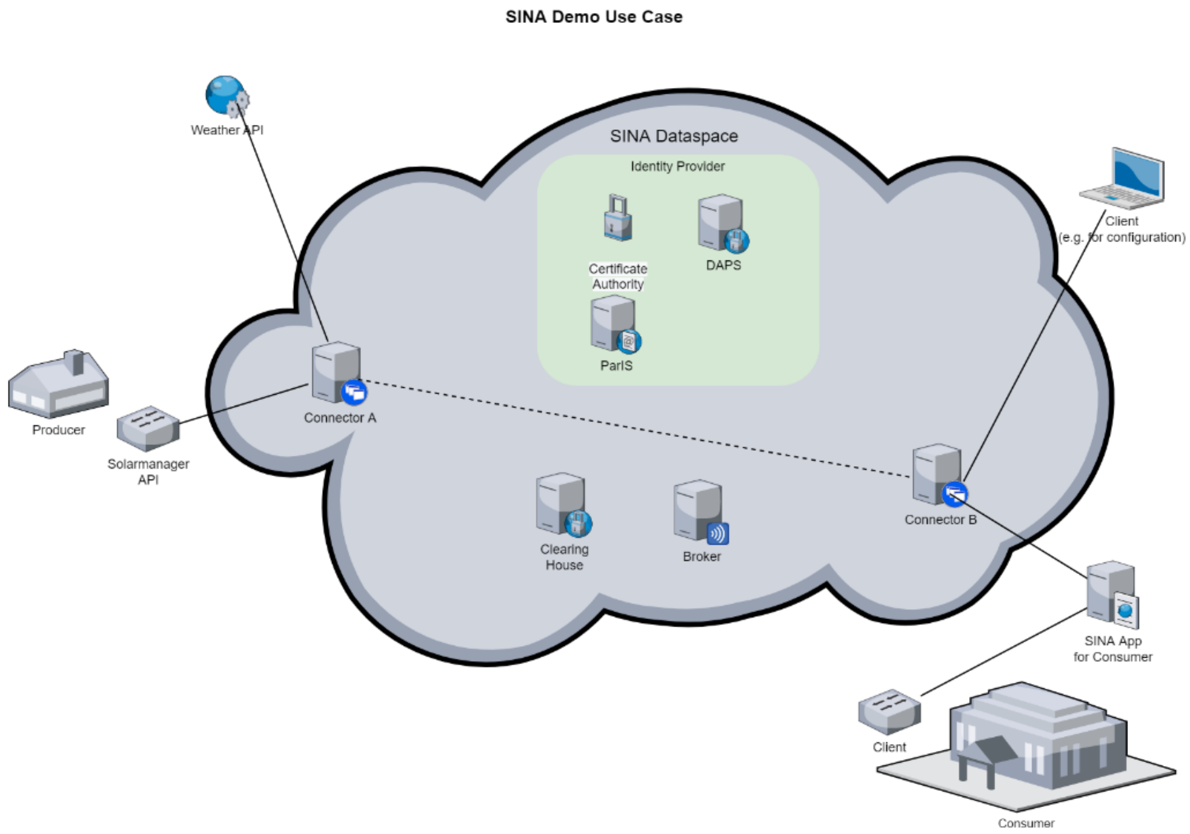

Figure 10 Symbolic diagram of the participants in the SINA Demo Use Case

A main challenge to build such a data space is the implementation of the process of data exchange. As shown in the reference architecture model (see Appendix A1) the process comprises of several steps from negotiation of the data offering and then exchanging the data, always respecting the requirements to security and sovereignty (Figure 11 to Figure 14). To obtain efficient handling, these processes must be automated. This is not done in the distribution of the MVDS coming from IDSA but must be implemented in the apps using the data space. As an alternative, a software library offering base data space services could be implemented and then used by several apps.

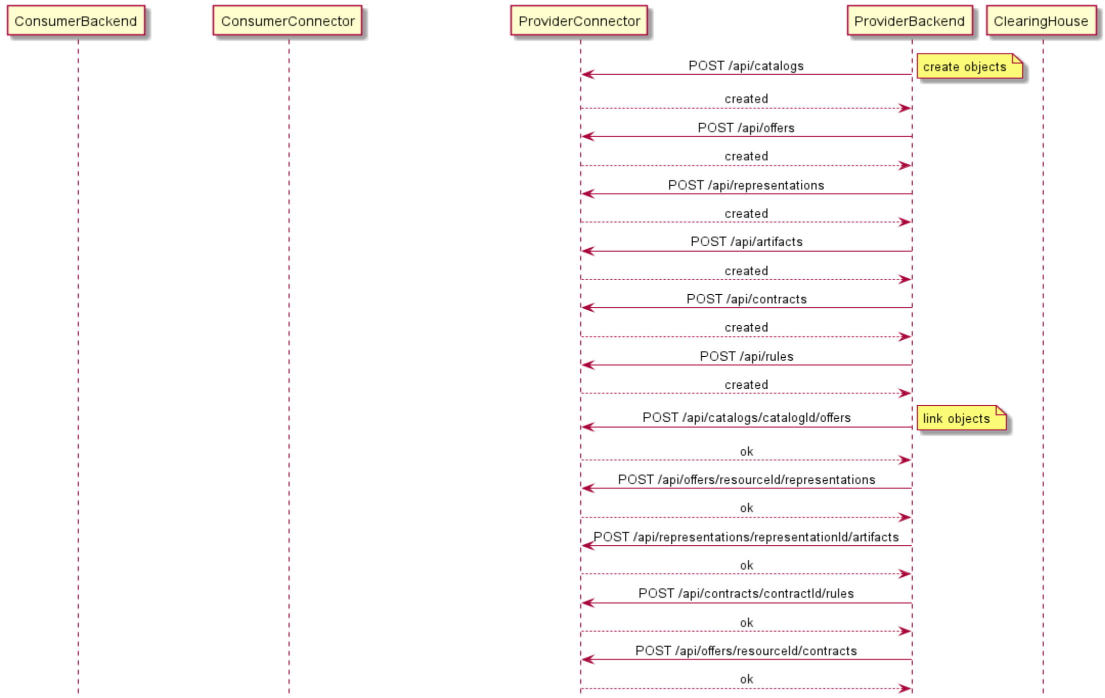

Figure 11 Preparation work needed by data provider for the exchange of data. Source: Fraunhofer ISST (2021)

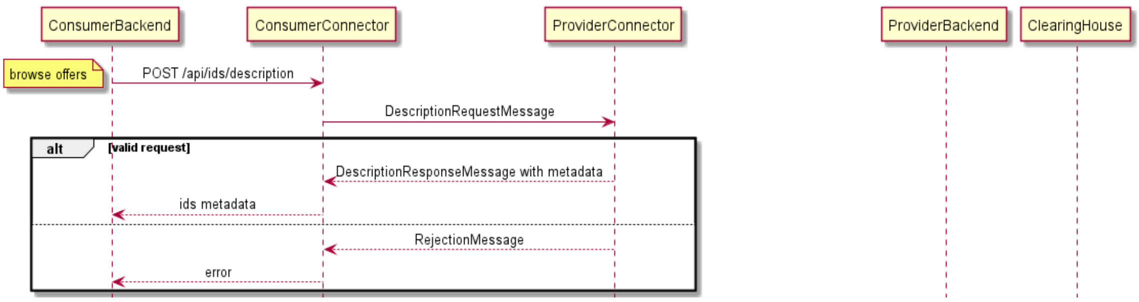

Figure 12 Process for the consumer to browse for offers from a provider. Source: Fraunhofer ISST (2021)

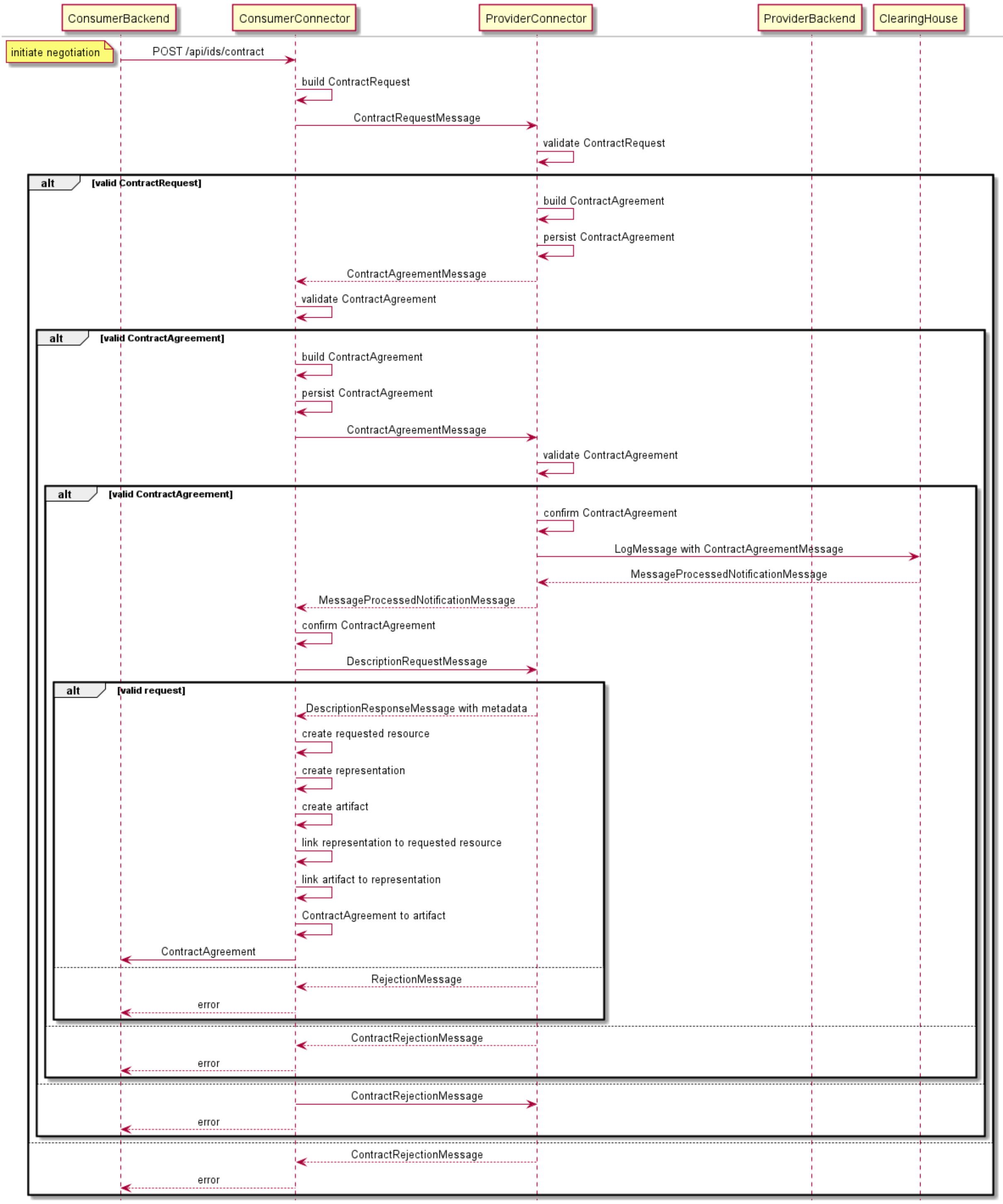

Figure 13 Contract negotiation process for consumer and provider. Source: Fraunhofer ISST (2021)

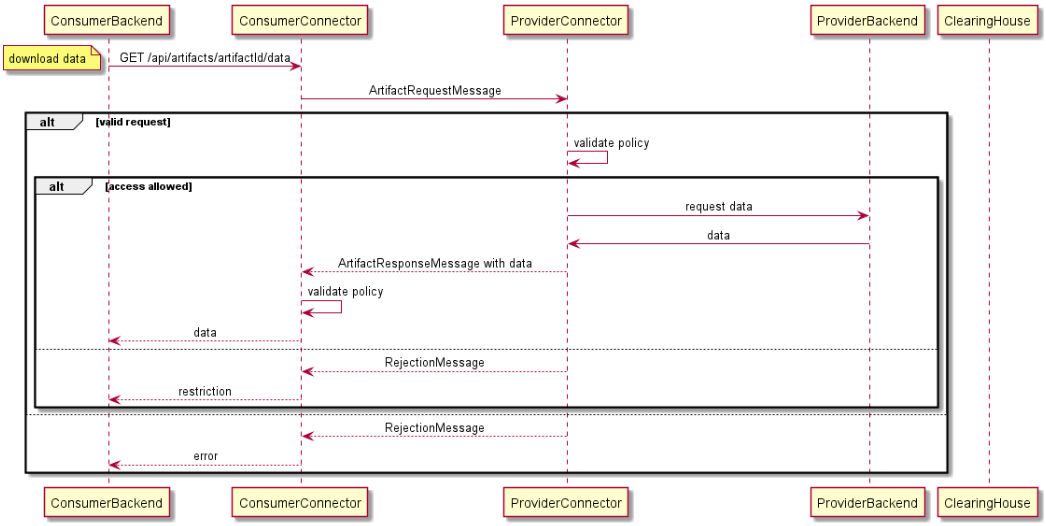

Figure 14 Data exchange process which is only possible when a contract is accepted from both sides. Source: Fraunhofer ISST (2021)

The main work in implementing a data space is the communication between backends and connectors on the providing and consuming side. Only the communication between the connectors is part of the implementation of the open-source software from IDSA. The work done in the project SINA is a base for a reference implementation that can help other companies to implement their own integration into a data space. In the demonstrator an UI shows for the data provider, the data consumer and the data space, represented by the metadata broker, the available offers (see Figure 15).

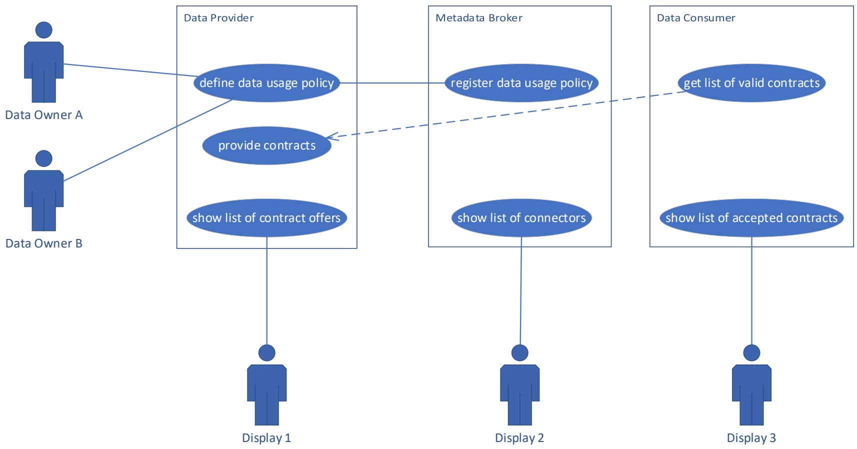

Figure 15 The use case diagram showing the perspective of offers and contracts and how they are registered and listed in the demonstrator

Availability of offers depends on the conditions of the data usage policy defined by the data owners. And only offers where the conditions of the data usage policy are fulfilled, listed, and shown on the UI. The data consumer only can get the data when the requirements of the data usage policy are met. This is shown in the demonstrator with a list of accepted offers and contracts.

3 Ontological conceptual aspects

Energy management applications require combining data from multiple sources and devices, which typically are based on manufacturer-specific conventions and information models. To make sense of this data, they need to be interpreted or in other words translated into a common representation that is clearly defined. Ontologies aim at providing a framework to enable this translation. In the following chapter an overview of protocols and standards is provided to identify requirements generally valid for ontologies used in future within SINA. Thus, it forms the conceptual blueprint for following up projects.

Data spaces may leverage existing, generic information models for data spaces while also necessitating domain-specific information models. Typically, providers specializing in vocabulary supply these domain-specific models, as outlined in the role model below. The adoption of a shared semantic model, utilized by both data service providers and consumers, presents substantial advantages in simplifying interconnection and collaboration complexities. Consequently, data spaces incorporate semantic management, achievable through a vocabulary hub (Figure 16). This hub is responsible for the management, registration, and publication of vocabularies, such as ontologies, reference data models, or metadata elements. Additionally, semantic transformation applications are employed to facilitate user-friendly mappings between semantic models.

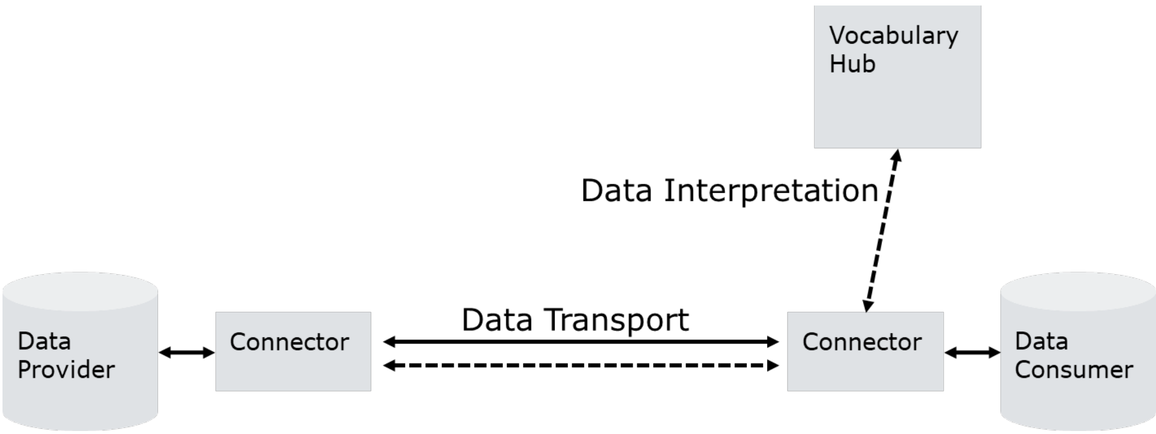

Figure 16 Representation of interaction between data space connector and vocabulary hub

To better understand the meaning of ontologies, the two following definitions are considered:

«Ontology: formal specification of a conceptualization, used to explicitly capture the semantics of a certain reality» (Daniele et al., 2015).

«In computer science and information science, an ontology encompasses a representation, formal naming, and definition of the categories, properties, and relations between the concepts, data, and entities, many, or all domains of discourse. More simply, an ontology is a way of showing the properties of a subject area and how they are related, by defining a set of concepts and categories that represent the subject» (Jacquette, 2002).

The main goals of using ontologies are:

– Help humans to interpret data: ontologies work with concepts and relationships in ways that are close to the way humans perceive interlinkages.
– Enable interoperability: by allowing distinct devices to exchange information and perform tasks together automatically.
– Automate reasoning about data: by having the essential relationships between concepts built into them, they can enable to implement e.g., data storage technologies such as semantic graph databases that use ontologies as their semantic schema. Semantic graph databases allow to store complex data with the relationships in the data included, without having the database being explicitly pre-structured, like would be required for SQL databases.

Therefore, the term ontology can sometimes be (partly) substituted with other terms such as data or information model, semantic model or vocabulary. In particular, in the context of data spaces, the term used by IDSA is vocabulary. There are however differences between an ontology and a data model. While an ontology defines all relevant concepts for a concept with their definitions, a data model is a representation of the structure of data, defining how data is organized within an information system. Typically, the data model will rely on the subset of the ontology that is relevant to the application. For example, the SAREF4ENER ontology (see next sections) defines all relevant concepts for energy including a power sequence, a load control event data, different temporal entities such delays, start times, time uncertainties, and so on. On the other hand, Sunspec is an information model describing what and how data can be retrieved from solar inverters.

Essentially, an ontology is a normative description of objects and their interrelations. In many cases, this description relies on a modelling language, e.g., OWL, UML. OWL is the preferred language used to describe several modern ontologies as it has gained wide acceptance in the web community. In other cases, the ontology is described in a reference document, meaning that it is not available in a machine-readable format.

In addition to that generic description, an ontology can sometimes have one or several reference implementations which allow to map the objects and relations into a given programming language, for example Python or Java. In particular, when an ontology or data model is one of the building blocks of a given software, it may then be directly reflected into the source code of that software as specific classes, attributes, etc. (for example the OpenHab home automation software relies on a data model described by openHAB (n.d.)).

3.1 Overview of existing ontologies

It is important to notice that different kinds of ontologies have been created for different purposes. Most ontologies are domain-specific, but they may be built on top of overarching ones that for example specify generic concepts such as devices, data point, time series, etc. Even among domain-specific ontologies, some are designed to be applicable in multiple use cases (e.g., SAREF4ENER) whereas others are designed for a single use case, for example to be used in pair with a given communication protocol for a specific application (e.g., OPEN ADR data model). In the latter case, the ontology may not be specified as an independent element but in conjunction with other elements that serve to be used in that application.

In Figure 17 an overview of existing ontologies that were found relevant to the SINA project is provided. Each ontology is positioned with respect to its intended field of use.

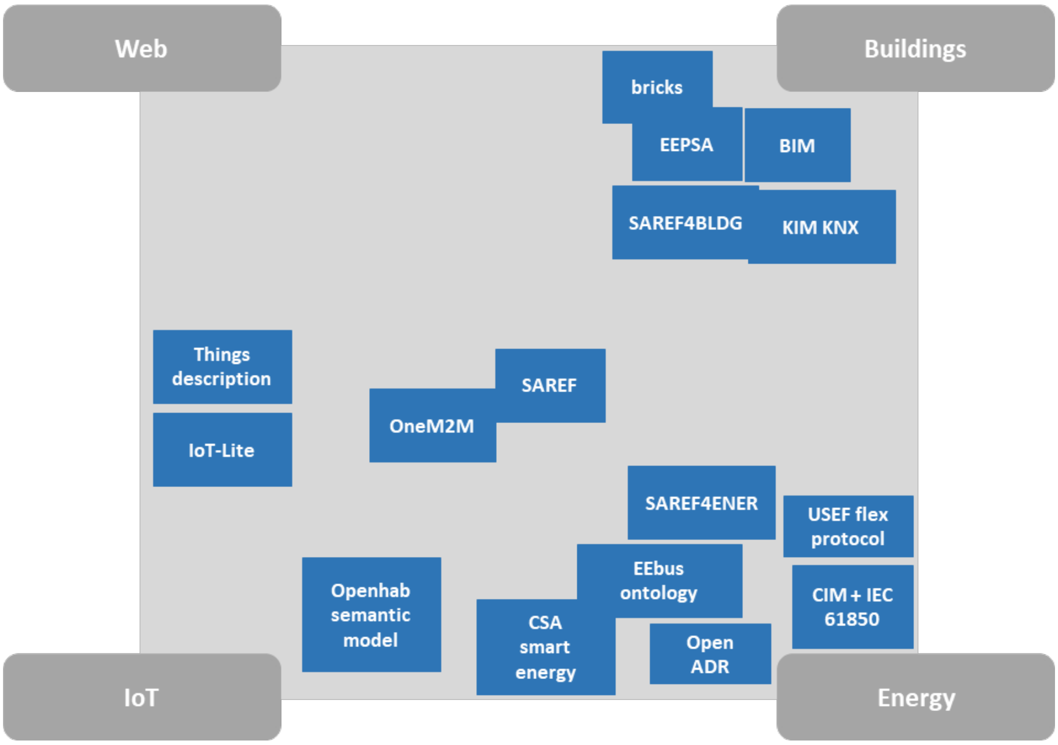

Figure 17 Map of relevant ontologies for SINA

Note that these categories are not strict, there certainly is a wide overlap between categories such as IoT and energy, especially if considering for example home energy management systems (HEMS) (Table 1). Appendix A2 provides a comprehensive overview of these relevant ontologies with associated online resources.

Table 1 Overview of relevant ontologies

Ontology / Data model	Authors	Field / Specific scope	Modelling language
CIM	IEC	Energy / Power grids, energy markets	UML
IEC61850 data models	IEC	Energy / Distribution substations, DERs	UML/XML

Ontology / Data model	Authors	Field / Specific scope	Modelling language
SAREF4ENER	ETSI	Energy / All use cases	OWL
EEBus ontology	EEBus association	Energy / Home and industrial energy management	OWL
OpenADR information model	OpenADR alliance	Energy / Demand Response events and tariffs	XML for messages
USEF flex protocol	USEF	Energy / Communication between flex aggregator and flex providers	XML for message
CSA Smart energy	CSA	Energy / Energy management and metering	NA
OpenHab semantic model	OpenHAB Community and foundation	Home automation	NA, directly embedded in OpenHab software
IoT Lite	W3C	IoT / all use cases	OWL
Things description	W3C	IoT / all use cases	Json based representation format
SAREF	ETSI	All / NA	OWL
OneM2M	oneM2M partnership	IoT / all use cases	OWL
Bricks	Open source – community driven	Building / building description	OWL
EEPSA	Academia	Building / Energy efficiency and comfort	OWL
BIM ontology	ISO	Building / Life cycle management	NA

Ontology / Data model	Authors	Field / Specific scope	Modelling language
SAREF4B LDG	ETSI	Building / all use case	OWL

3.2 Ontologies in the domain of energy and grid

To the best of our knowledge, there is no ontology that is specifically tailored for VPP applications. Naturally, ontologies that are for the energy domain are relevant, most prominently SAREF4ENER. Additionally, some parts of the common information model (CIM) standard are relevant: the IEC standard 62325-301 Framework for energy market communications - Part 301: Common information model (CIM) extensions for markets covers. According to IEC 62325-301:2018 (IEC, 2018):

«It specifies the common information model (CIM) for energy market communications. The CIM facilitates integration by defining a common language (i.e. semantics) based on the CIM to enable these applications or systems to access public data and exchange information independent of how such information is represented internally. The object classes represented in the CIM are abstract in nature and may be used in a wide variety of applications. The use of the CIM goes far beyond its application in a market management system. This new edition of IEC 62325-301 contains support for demand-side communication within a wholesale market. The IEC 62325-301 additions include support for demand-side resource registration and enrolment of a market participating resource as well as support for deployment and performance evaluation of demand side resources. A new package has been included in this edition of IEC 62325-301 to support environmental (weather) data».

As VPPs are designed for interfacing small prosumers towards the larger energy markets, the aspect related to demand-side resources are certainly relevant.

3.3 Ontologies in the domain of buildings

Within the building domain existing standards and ontologies can be leveraged in several main directions, which are illustrated in Figure 18:

– Communication and device-specific interactions with various kinds of equipment.
– Communication with smart meters.
– Communication with flexibility buyers, typically aggregators.

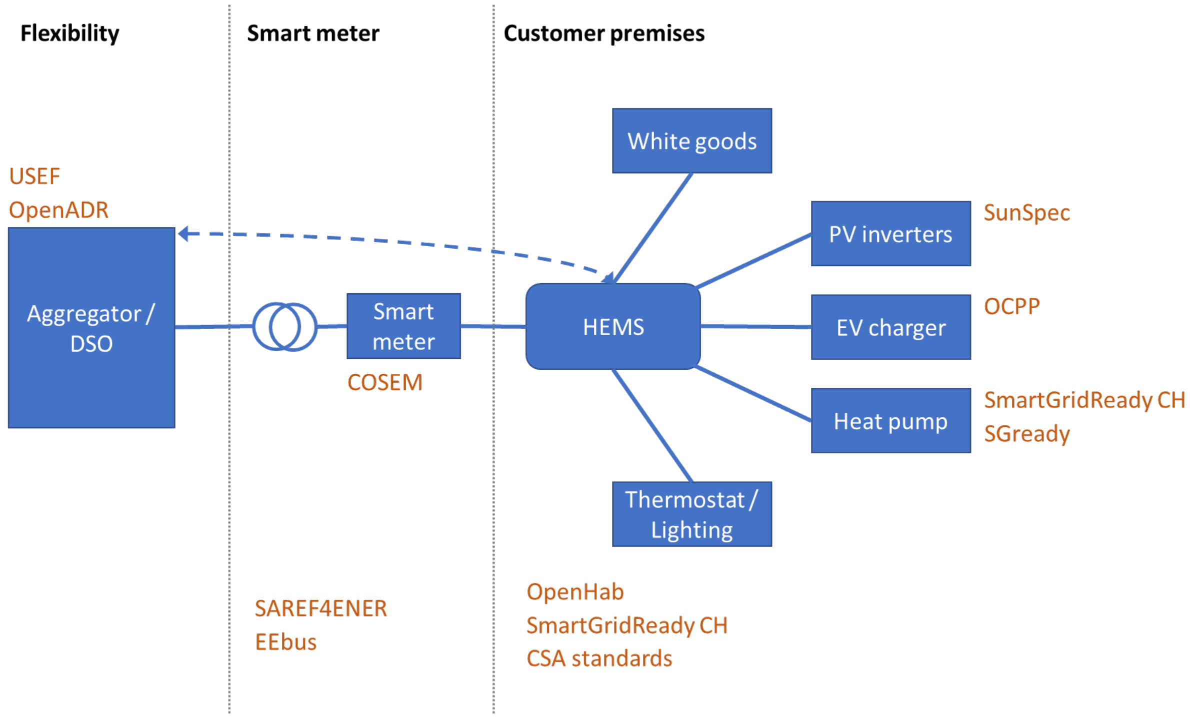

Figure 18 HEMS main scope

3.3.1 Communication and device-specific interactions

One of the base features shared by all HEMS is to collect energy-related data from various types of devices. Maintaining the support communication with devices such as heat pumps, electric car chargers, batteries, inverters, etc. remains one of the main hurdles for companies that create HEMS. In existing products, some companies deliberately support only specific (often proprietary) equipment (e.g., Lox-one) which limits the pool of potential customers and diminishes the benefits provided by the HEMS when some equipment is not supported by them. In other cases, HEMS products attempt to support as much equipment as possible, but this results in lower reliability and generates significant costs to main-tain and support the ever-increasing numbers of protocols, devices, etc. Using ontologies can help to bring more uniformity to the semantic model of different models, for example by helping to specify what data are expected to be measured by a given type of equipment or what commands should a device be capable of receiving. The following standards are relevant for communication with a specific category of device:

– SunSpec for communication with solar inverters.
– The Open Charge Point Protocol (OCPP) for communication with charging stations.
– SGready is a very simple protocol that has seen relatively wide adoption for heat pump energy management. It, however, only offers very basic options.

Note that device specific communication can entail much more than the energy part. For example, for car chargers, OCPP covers user identification, security, session management, and more.

The initiative SmartGridReady CH proposes a unified and machine-readable way to represent communications with different devices as device profiles (XML-format definition files), and to some extent to offer unified sets of datapoints to read/write for different applications (called function profiles) for a given device category, e.g., heat pumps. On one hand, it is not as abstract as other complex ontologies such as SAREF4ENER, but it can be seen as following a simple ontology that concretely takes the form of schema definition files (.xsd) which are templates for device profiles and specify how to describe how to communicate with devices and what data is available for different applications. A data space could integrate SmartGridReady CH device profiles by making them available to connectors, but it is important to mention that Smart Grid Ready CH focuses on real-time communication and control with real-world equipment, an aspect that is not explicitly tackled by data spaces.

EEbus has the main advantage to integrate related to energy with communication specifications, data model based on modern ontology and use case detailed descriptions. It supports different categories of equipment.

SAREF4ENER finally, is the highest-level available ontology for the energy domain.

Overall, lack of widespread adoption is a problem affecting to different degree every ontology and communication standardization effort.

3.3.2 Communication with smart meter

Energy management may in some cases rely on communication with smart meters for data collection. In most cases, smart meters are owned and managed by the distribution grid operator and follow relatively strict national or local regulations. Relevant efforts here include:

a. DLMS/COSEM (Device Language Message Specification/Companion Specification for Energy Metering): is a standard protocol for data exchange between smart meters and data concentrators promoted by the DLMS User Association whose members include manufacturers, utilities, system integrators, etc. DLMS/COSEM has been recognized globally as a standard for data exchange for smart devices since 2002 in the IEC/EN 62056 and the EN 13757 standard suites. It is widely implemented internationally with more than 150 vendors and covers over 1 500 certified device types for various applications. It has three main components:

– COSEM: Companion Specification for Energy Metering which defines a set of standardized data objects, with a given set of attributes and methods, covering all functions of the meter. Each data object is represented by a unique object identification system (OBIS) code, independent of the manufacturer, which enables clients to identify and access the data objects. It also defines a set of standardized services that enable clients to interact with the data objects, based on the specific access rights granted.

– OBIS: Object Identification System is a standardized system that uniquely identifies and describes the data objects used in electricity, gas, water, and thermal metering, as well as abstract data.

– DLMS: Device Language Message Specification defines the structure and format of messages communicated across devices, such as data objects, methods, and services.

b. CSA smart energy/Zigbee: Smart Energy is a standard for interoperable products that monitor, control, inform and automate the delivery and use of energy, water and gas. It is employed in

Smart meter relying on the communication protocols supported by CSA, namely Zigbee. According to the CSA home page, 40 million Zigbee Smart Energy electric meters are being deployed by more than a dozen utility companies, mostly in the United States.

3.3.3 Flexibility applications

Flexibility entails the manipulation of the consumption of distributed resources. It usually involves pooling of resources under the umbrella of one responsible entity (the aggregator) which by combining multiple small resources enter energy services normally not open to small customers. The communication from the aggregator to the end devices can happen through a local HEMS, which in many cases is managed by the aggregator (and hence using proprietary communication, etc.), but could in principle be done through «independent» HEMS provided it relies on standard communication and data models. In this regard, the OpenADR protocol focuses on that link and is designed to facilitate the implementation of Demand-response programs (automated remote control of demand-side resources). OpenADR has seen some level of adoption in the United States with multiple manufacturers supporting this protocol in their devices. Finally, USEF has devised a more general protocol for flexibility, which has only been used in pilot demonstrations, but can be used as an excellent blueprint for other future protocols.

The advantages and drawbacks of the higher-level ontologies/standards listed above are summarized in Table 2.

Table 2 Main advantages and drawbacks of selected ontologies

Ontology / Data model	Advantages	Drawbacks
CIM - IEC 61850	Widely used in industry Existing supporting tools	Not public Too specific for our purpose -> mostly focused on the grid
SAREF4ENER	Academic and institutional support Flexible and rich	Complex Lacks open-source implementations and example
EEBus ontology	Industry support with compatible products Few open-source implementation (although limited) Use cases are described precisely	Not many supported products Complex
OpenADR information model	Specific and practical Some compatible products	Limited in scope and functionalities Not so actively extended anymore

3.4 Link between ontologies and data spaces

IDS uses the terminology of «vocabulary» to represent concrete descriptions of ontologies. IDS itself has introduced an information model which primarily aims at the description, publication and discovery of data products and data processing software within the IDS. It is described fully on its documentation webpage (Mader et al., 2022): «the Information Model has been specified at three levels of formalization. Each level corresponds to a digital representation, ranging from a high-level, conceptual document up to the level of operational code» as depicted in Figure 19.

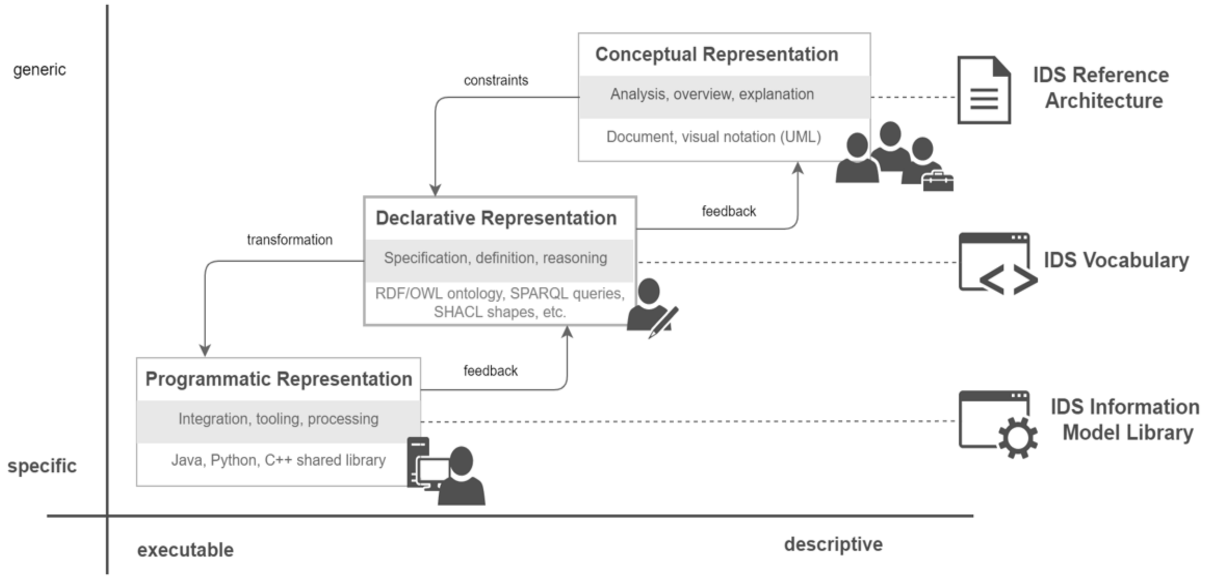

Figure 19 Information model levels. Source: IDSA (2022a)

The description of the three levels of formalization can be summarized as follows (International Data Spaces Association, 2022):

– The Conceptual Representation of the Information Model (CRIM) analyzes and provides an overview of the main concepts of the Information Model, regardless of technology or domain. Descriptions are generic, allowing comparison and shared understanding.

– The Declarative Representation of the Information Model (DRIM) defines the information model of the industrial data space, based on W3C standards and modeling vocabularies. It provides a formal, machine-interpretable specification and allows for sharing and reasoning upon meta-data descriptions. It is a referential model that can derive Programmatic Representations. DRIM is a reusable core model that uses existing domain vocabularies and standards for interoperability.

– The Programmatic Representation of the Information Model (PRIM) helps developers integrate the Information Model with their development infrastructure. It includes a programming language data model and software libraries. It allows developers to create compliant instances of the Information Model without dealing with RDF graph model or ontology processing.

The declarative representation, that forms the IDS vocabulary (see Appendix A1 for more details) is the only normative specification of the Information Model. A set of auxiliary resources, among others, guidance documents, reference examples, validation, and editing tools is intended to support a competent, appropriate, and consistent usage of the IDS Vocabulary. As mentioned, it relies on custom vocabularies for domain-specific elements. That is where specialized ontologies enter the picture.

The interoperability requirements in the IDS directly lead to the usage of commonly known, standardized terms to describe data, services, contracts, and so on. Collection of these standardized identifiers form so-called vocabularies.

In the IDS, however, further requirements occur. The terms of the vocabulary must be machine-readable, also to some degree their descriptions and titles, as well as new terms must be available for lookups. The IDS relies on RDF to encode its attributes and data descriptions.

The IDS Information Model is the central vocabulary that all parties of any IDS share. It is therefore the minimal set of terms all IDS components must understand. In specific domains, however, more and more expressive terms are needed. It is, therefore, a good practice to extend the basic information model with additional vocabularies and provide them in the same ways as the core one.

To do so, a certain service is needed to provide a platform to host, maintain, publish, and document the additional vocabularies. This service is the IDS Vocabulary Hub. It provides IDS-conform endpoints to enable seamless communication with IDS Connectors and infrastructure components. It is important to notice that IDS proposes with the vocabulary hub an infrastructure to host domain-specific vocabularies but does not take care of building the ontologies which would be domain-specific. It is also to be noticed that the hub hosts representations of the information model so it does not perform interpretation, that task would need to be performed in the application.

Connectors might use it to increase the information content of their asset's Self-Descriptions. In the IDS world, this happens by introducing new attributes or values with previously unknown URIs/IRIs. Connectors that read those Self-Descriptions face the challenge of not knowing their semantic meaning at first. They can now lookup (or «dereference») the attribute's identifier at the Vocabulary Hub. The Vocabulary Hub responds with a small RDF document explaining the attribute. This usually includes the type or class of the entity, its label in different languages, and a short description, also possibly in several languages. The Connector can integrate these explanations into its workflows and thereby present the newly discovered meaning to its users.

An implementation of a vocabulary hub that can host and allow collaborative work on vocabularies is represented in Figure 20.

A wider reliance on ontologies or at minimum on standardized data models can greatly contribute to reduce some of the hurdles to the wide adoption of smart energy services, such as advanced energy management, participation of demand-side resources in flexibility mechanisms and markets. Indeed, ontologies are a way to specify commonly agreed terminology and requirements in terms of data / metadata, etc. This should help interoperability of components and services at a low (ideally zero) development cost. In addition, modern representations using modelling languages such as OWL allow machine readable representations of data models, services, etc., which are important to automate executions of various tasks.

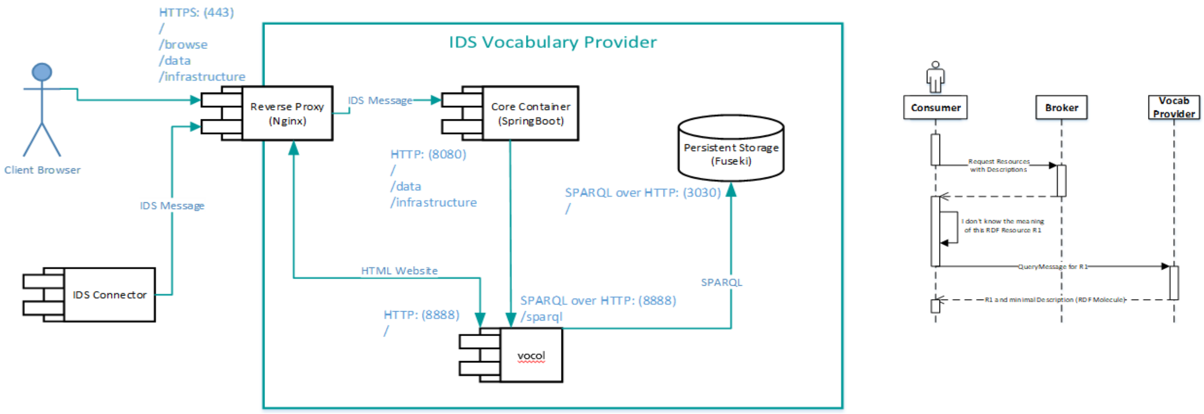

Figure 20 Vocabulary hub architecture as proposed by Fraunhofer. Source: Fraunhofer IAIS (2022)

Nevertheless, the existence of multiple competing ontologies addressing widely overlapping scopes partly undermines this goal. More importantly, to the best of our knowledge, no ontology has gained sufficiently large adoption to justify systematic efforts to further use, support that ontology. We believe that this is largely due to the hurdles involved in using these ontologies:

- Designing services relying on ontologies certainly involve a complexity overhead compared to more minimalistic approaches. Ontologies rely on technologies, languages and tools that are the view of specialists (sometimes called knowledge architects) and are not widely mastered even in fields like computer science.
- Although the ontologies themselves are reaching some maturity and consolidation over time, there is a very problematic lack of publicly available examples of data models or applications built on top of those ontologies (that is for example particularly true for SAREF4ENER based on which virtually no applications were found to have been implemented). This is a major hurdle for practitioners which we believe cannot be expected to start from scratch. How to use and integrate ontologies is not that established. Because these ontologies are often perceived as «academic» or overly «abstract».
- For developers of a single service and especially components, relying on ontologies involves overhead that is not always compensated for by clear immediate benefits. For component manufacturers, allowing for a wide interoperability of their components is only one objective among many others they need to fulfil. This is for example salient for heat pumps where efforts to standardize data exposed have still not given rise to a satisfactory and widely adopted solution. In some cases, business logic can even push against the use of standardized solution to create barriers for competition, etc. In turn, without compatible components, the use of ontologies in applications has diminishing returns as it does not allow to streamline development as advertised (for example specific SW «connectors» or «translators» are often required to support various components).

Still, ontologies are steadily gaining a larger audience and institutional support. Data spaces as per the IDSA initiative have fully recognized the usefulness of ontologies by 1) defining a core «vocabulary» to support the core functionality of any data space and 2) planning for the possibility to include specialized or vocabularies to support domain specific ontologies. It offers a reference implementation for a vocabulary hub that helps to integrate vocabularies in application relying on data spaces.

Nevertheless, practical implementation demonstrating the use of it is still lacking. The prevailing landscape in ontological frameworks reveals a notable absence of widespread adoption that would warrant systematic endeavors to fortify their utilization and support. The limited adoption of ontologies can be attributed to inherent complexities. Designing services based on ontologies involves a higher level of intricacy compared to simpler approaches, using specialized technologies mastered by a select few.

Besides the further work that is needed regarding the ontologies, the legal aspects for such a data space implementation have to be further conceptualized. Therefore, the next chapter will provide an overview of the relevant aspects in this area.

4 Governance: Legal aspects

In this chapter, the primary focus is on the legal aspects of a data space and the associated data handling. The essential legal aspects in every data exchange are analyzed and the issues that necessitate their application in a data space project, such as SINA, are uncovered. As we further explore the complexities of legal considerations, we address the challenges posed by the operation of diverse organizations under varying legal jurisdictions and frameworks.

4.1 Legal documents

The primary focus on legal aspects of SINA, was on the legal documents, including:

- The project's Draft of the Statutes for setting up SINA association («Vereinsstatuten») (Appendix A7).
- The «SINA’s Rulebook for a Fair Data Economy - General Terms and Conditions» serving as a guiding framework and outlining the ethical principles and behavioral expectations for all participants engaged in SINA's data exchange (Appendix A3).

The close connection to the SITRA notebook, a rulebook for a fair data economy, was not only chosen for its efficiency but also because it enhances legal interoperability between other data space projects. As we delve into the legal considerations, the practical application of the Code of Conduct, General Agreement, and Constitutive Agreement is empathized, providing insights into their significance and relevance within SINA.

Participants in SINA's data exchange are not only expected to comprehend these legal documents thoroughly but also to actively implement measures ensuring compliance with the stipulated guidelines.

4.2 Legal foundations

The fundamental legal considerations that refer to data spaces in general are further explained and discussed. By exploring general principles and regulations, we aim to establish a foundational understanding of the legal landscape surrounding these dynamic environments. As data spaces continue to evolve, navigating the legal aspects becomes increasingly crucial, and this exploration sets the stage for a comprehensive examination of specific legal dimensions in subsequent sections.

Data platforms require the implementation and realization of legal requirements. In the case of domain-specific data platforms that deal with known vendors and predetermined business requirements, the policy and regulation are assessed and included in the architecture of the corresponding data platform. However, in the case data spaces, the uncertainty of policy and regulations makes this task very difficult, if not impossible. The legal challenges span many policies and regulations, including privacy and data protection law, national/international data protection legislation (e.g., GDPR or FADP), human rights law (e.g., EU Convention on Human Rights), and business regulations. Data Governance involves managing and maintaining data properly, covering areas like data rights, privacy, and security. It operates at different levels: micro (within an organization), meso (among collaborating organizations), and macro (national or international). While goals may vary, the focus should be connected throughout the data life cycle. For example, the GDPR is a macro-level rule in Europe for personal data, influencing internal practices at the micro level for compliance. In the discussion ahead, the emphasis will be on crucial aspects at the macro level (Curry et al., 2022).

4.2.1 Trade secret

Confidentiality agreements aim to protect sensitive information, but sharing data raises a paradox—focus on empowering data consumers for new insights. Data providers may not control consumer usage fully. In SINA, providers influence shared metadata to guide interpretations. Consider EU Trade Secrets Directive (European Parliament, 2016) and national laws alongside confidentiality measures. Data may qualify for trade secret protection if secret, valuable, and protected. Assess data for protection eligibility, document measures, and define access rights. This, combined with a robust framework, enhances trade secret protection in shared data, offering rights for legal action if misused (Duisberg, 2022, p. 65).

4.2.2 Competition law

Competition law in the digital sphere presents a complex and uncertain landscape that demands resolution. Traditional mechanisms of merger control have evolved in recent years to encompass and regulate situations in which the impact on the market is not solely determined by the size of the merging companies but also addresses scenarios where the central focus of the proposed transaction revolves around a data-abundant target company with limited traditional economic significance (Duisberg, 2022, p. 66).

Furthermore, the investigations carried out by the German Federal Cartel Office against Facebook, which scrutinized the terms of use to identify potential cases of excessive data collection, offer insight into how regulatory authorities are endeavoring to broaden their authority to oversee dominant market positions that hinge on the accumulation of substantial volumes of data and data-driven business models (Bundeskartellamt, 2019).

Ultimately, the EU Commission has tackled the issue of market power imbalances concerning data-rich enterprises, commonly referred to as «data advantage», and has brought to the forefront the imperative to establish novel methods for preventing market distortions and fostering healthy competition. This initiative has notably led to the creation of the new EU Digital Markets Act (Duisberg, 2022, p. 67).

4.3 EU strategy on data: The relevance of data spaces and the applicability to Switzerland

The EU Commission has already taken significant steps in implementing its vision through legislative measures. These measures encompass a multifaceted approach: firstly, establishing a comprehensive cross-sectoral governance framework for data access and utilization; secondly, increasing the accessibility of high-quality public sector data for reutilization; and thirdly, introducing a data act to encourage horizontal data sharing among sectors. These legislative actions are part of the EU Commission's ongoing efforts to build a data-driven society, highlighting their commitment to robust data governance and expanded data-sharing opportunities to drive innovation and economic growth across various sectors (European Commission, 2022). Switzerland's unique status as a non-member of the European Union (EU) gives rise to distinctive considerations when it comes to the EU Data Strategy. Unlike EU member states, Switzerland is not automatically subject to the regulatory framework proposed by the EU Data Strategy. Therefore, in the following discussion, we will also evaluate how different acts affect Switzerland. This is particularly important because as shown in the following the EU Strategy on Data plays a significant role in shaping legal aspects, and since Switzerland is part of SINA, this consideration is of importance.

Table 3 outlines key EU Acts and regulations that play an important role in shaping data spaces, with a focus on data governance, data protection, and the facilitation of data sharing.

Table 3 Overview of EU Act. Adapted from: Sury (2023)

Name of the regulation /directive	Entry into force	Brief description	Addressee	Purpose
Digital Markets Act (DMA)	01 Nov 2022	Law on the regulation of gatekeeper platforms in the digital market	Large online platforms, gatekeepers	Stopping unfair practices by gatekeepers, improving consumer protection
Digital Services Act (DSA)	In force since: Nov 16, 2022 / transition periods until: Feb 17, 2024	Act on the Regulation of Intermediary Services and Search Engines	Online platforms, digital service providers, large-scale intermediary services (VLOPs), and large-scale search engines (VLOSEs)	Protecting consumers and their fundamental rights on the Internet, combating the spread of illegal content online
Data Governance Act (DGA)	June 23, 2023 / transition periods until September 24, 2023, and September 24, 2025	Law on Optimizing Data Sharing in the Single Market	Public agencies, intermediary service providers, data altruistic organizations	Creation of a single European data space, promotion of the dissemination of data from public institutions to the general public
Data-Act (Draft) (DA-E)	Expected end 2024	Act to regulate data generated by the Internet of Things (IoT)	Manufacturers of «smart» products, data processing services (cloud, edge)	Creation of a single European data space, increase of legal certainty, prevention of abuse of market power, use of data for public interest
GDPR	May 2018	Safeguards individuals' personal data by setting	The GDPR primarily addresses organizations and	Ensuring the protection of individuals' personal data and to give them more control

Name of the regulation /directive	Entry into force	Brief description	Addressee	Purpose
		strict rules for its handling and imposing penalties for violations	businesses that handle the personal data of individuals within the European Union	over how their information is collected and used

4.3.1 Data Governance Act

The Data Governance Act, a fundamental component of Europe's data strategy, aims to foster confidence in data sharing, enhance mechanisms for expanding data accessibility, and tackle technical barriers that hinder data reutilization (European Parliament, 2022)

Furthermore, the Data Governance Act is geared towards promoting the establishment and growth of unified European data ecosystems across key sectors. These ecosystems will engage a diverse spectrum of participants, spanning private and public entities, in areas like healthcare, the environment, energy, agriculture, transportation, finance, manufacturing, public administration, and skills development (Duisberg, 2022).

The Data Governance Act officially came into effect on June 23, 2022, and, after a 15-month transition period, it became enforceable as of September 2023.The Act aims to emphasize the importance of data sharing, unlike the GDPR both personal and non-personal, within the European context. It seeks to reduce the dominance of large tech corporations by promoting «neutral data intermediaries» as an alternative model for data sharing.[13]

This new approach places a significant focus on data held in the public sector and facilitates data sharing across European Member States. A central requirement to ensure trust and control in data sharing is the neutrality of these data intermediaries, meaning they only act as middlemen in data transactions and don't use the data for other purposes.[14]

Data intermediaries must be based in the EU or have an appointed representative if they operate from outside the EU. They are also required to follow a notification procedure, which is still under development in the Member States, to inform the relevant registry or authority of their intent to provide intermediary services.[15] This approach places a strong emphasis on the role and responsibility of data space operators.

The Data Governance Act imposes specific notification procedures on data intermediaries for various types of services:

[13] Art. 1 DGA
[14] Art. 33 DGA
[15] Art. 42 DGA

- Facilitating data exchange between data holders and data consumers, either bilaterally or in multilateral data exchanges, or through the creation of platforms and databases for data sharing. This also includes setting up infrastructure for connecting data holders and data consumers.
- Enabling data subjects to share their personal data with potential data consumers, allowing data subjects to exercise their rights under the GDPR (General Data Protection Regulation).
- Supporting data cooperatives, particularly in the context of micro, small, and mid-size enterprises.

It's important to note that the concept of notification does not imply approval by the authorities but serves as a mechanism to establish certain conditions for data intermediary services before they commence their operations. This mechanism also allows for supervisory control to ensure compliance with these conditions, with the potential for imposing significant financial penalties if necessary. Additional measures regarding data access and use are under consideration in the EU Data Act, which has been under discussion since February 2022.

The conditions outlined in Article 11 of the Data Governance Act are particularly relevant in the context of SINA. These conditions include:

- Data intermediaries must use the data they receive solely for making it available to data consumers.
- They should only use metadata collected from data sharing services for purposes related to developing the service, ensuring fair, transparent, and non-discriminatory access, and for facilitating data exchange in compatible formats.
- Data intermediaries must prevent fraudulent or abusive practices and ensure the security and continuity of data services.
- They should also provide advice to data subjects regarding potential data uses and standard terms and conditions associated with such uses.

From a SINA perspective, the Data Governance Act aligns with SINA approach, emphasizing neutral intermediaries, reference architecture, and connector technology to facilitate data sharing between data holders and data consumers in bilateral and multilateral ecosystems. The Act also enforces requirements related to data formats and interoperability.

Operators of data spaces should pay attention to their monetization models. The current draft of the Data Governance Act seems to limit an operator’s ability to generate innovative services through extensive metadata analytics, primarily focusing on «further development» of data sharing services. Further clarification is needed on whether intermediaries can create anonymized metadata aggregations, reports, and use such data for machine learning, provided that the actual metadata is properly protected and not shared for commercial gain or unauthorized purposes. The Data Governance Act is open to debate, especially concerning security requirements and references to «high level of security». It is a part of the broader legislative framework under the European Data Strategy, which is subject to ongoing development and refinement (Duisberg, 2022, p. 73).

4.3.2 Goals of the Data Governance Act

The Act seeks to enhance data accessibility and promote data sharing among various sectors and European Union nations, with the aim of harnessing the potential of data to benefit European citizens and businesses (e.g., through data spaces).[16]

To bring in an example (European Commission, 2022):

- Efficient data management and sharing will empower industries to create novel products and services, enhancing the efficiency and sustainability of numerous sectors in the economy. This is also crucial for training artificial intelligence systems.
- Increased data availability will enable the public sector to formulate improved policies, fostering transparency in governance and enhancing the effectiveness of public services.

Data-driven innovation will yield advantages for both companies and individuals, streamlining our lives and work by (European Commission, 2022):

- Health data: Enhancing personalized treatments, improving healthcare, and aiding in the treatment of rare or chronic diseases, potentially saving around €120 billion annually in the EU's healthcare sector. It also enables a more rapid and effective response to global health crises like COVID-19.
- Mobility data: Saving over 27 million hours of public transport users' time and up to €20 billion annually in labor costs for car drivers through real-time navigation.
- Environmental data: Combating climate change, reducing CO_2 emissions, and addressing emergencies such as floods and wildfires.
- Agricultural data: Advancing precision farming, fostering new products in the agri-food sector, and offering new services in rural areas.
- Public administration data: Providing better and more reliable official statistics, contributing to evidence-based decision-making.

4.3.3 Applicability of the GDPR in Switzerland

The European Data Governance Act does not have direct applicability to Switzerland (Bundesamt für Kommunikation, 2023).

Consequently, the provisions regarding the disclosure of confidential data from public sector entities and the legal obligations for data intermediaries do not extend to Switzerland. The regulation also does not exempt relevant EU bodies from their confidentiality obligations, as outlined in Article 3(3) of the Regulation.

In practical terms, this means that data received by an EU agency from abroad through administrative assistance, typically governed by specific confidentiality clauses, remains unaffected by this Regulation. However, Swiss companies may still be impacted, especially when data intermediaries that offer their

---

[16] Art. 2 DGA

services within the EU but are not based in the EU (e.g., in Switzerland) must adhere to specific regulations, such as appointing a legal representative in an EU Member State, as per Article 11(3).

Furthermore, the Regulation becomes relevant when dealing with the transfer of confidential, non-personal data from public bodies in the EU to Switzerland. In such cases, any further users of this data can only initiate transfers to third countries if adequate safeguards for protecting non-personal data exist in those third countries. These safeguards are considered adequate if the level of protection for trade secrets and intellectual property in the third country is essentially equivalent to that in the EU. The European Commission may declare a third country as meeting these equivalent protection standards through implementing acts (Bundesamt für Kommunikation, 2023).

In practical scenarios, this aspect of the Regulation will likely apply primarily to Swiss companies or institutions (e.g., research organizations or universities) seeking to access confidential data released by EU public authorities and transfer it to Switzerland (e.g., SINA Data Space Ecosystems). This process for ensuring the adequacy of transferring confidential data from public bodies differs from the adequacy decision procedure established in the GDPR, which specifically concerns the transfer of all personal data to third countries (Kommission von der Leyen, 2023).

4.3.4 GDPR/FADP

The General Data Protection Regulation (GDPR) became effective in May 2018, while the revised version of the Federal Act on Data Protection (FADP) was enacted in September 2023. Given the significant similarities between the revised FADP and the GDPR, our primary focus will be on the GDPR. This is particularly relevant as the GDPR is already applicable to certain Swiss businesses due to their ties with the European Union. It's worth noting that each EU member state has its own national data protection law, but we won't delve into these complexities in this discussion.[17]

When considering the perspective of SINA and data sharing, a fundamental question arises regarding the GDPR: which entities assume specific roles (controllers, processors, and joint controllers)[18] to assess their corresponding responsibilities? Additional considerations pertain to the operators of data spaces and their role in providing the necessary technical and organizational measures to facilitate the sharing and processing of (personal) data within SINA.

In a bilateral data sharing scenario, the roles of data providers (acting as the original data controllers) and data consumers (acting as subsequent data controllers) are typically straightforward, resulting in a controller-to-controller data transfer. In this context, the operator of the data space assumes the role of a data processor by providing the necessary infrastructure. The data provider must assess the legal basis for sharing data with the data consumer under Article 6 of the General Data Protection Regulation (GDPR), which could include considerations such as obtaining the data subject's consent, fulfilling a contractual obligation with the data subject, or demonstrating a legitimate interest. The data provider is also responsible for ensuring compliance with additional obligations imposed on data controllers, such as providing privacy notices under Articles 13 and 14 of the GDPR, safeguarding the rights of data

---

17 Art. 3 GDPR

18 Art. 4 para. 7, 8 GDPR

subjects under Articles 15 and subsequent, and fulfilling documentation requirements under Article 30 of the GDPR (Duisberg, 2022, p. 75).

In the context of scientific research, especially concerning special categories of data, such as health data regulated by Article 9 of the GDPR (see also Art. 5 lit. c FADP), data providers may have the option to invoke specific legal justifications provided by Member State legislators through national derogations within SINA.

In multilateral data sharing, the data provider must carefully assess the legal basis for sharing personal data with each data consumer separately. This evaluation may lead to the application of different legal bases depending on the specific data processing purposes envisioned by each consumer. It's important to note that the flexibility to change the intended purpose of data usage is limited under Article 6, paragraph 4 of the GDPR.

If the data provider relies on consent as the legal basis, they must offer robust consent management tools, including the option for individuals to withdraw their consent. In cases where legitimate interest is the chosen legal basis, the data provider must ensure the right to object to data processing is safeguarded.

Moreover, additional complexity arises when multiple data controllers jointly determine how personal data will be processed. This can happen in scenarios involving multilateral or coordinated data sharing, whether it's between the data provider and various data consumers or exclusively among different data consumers. In these situations, the involved parties must establish a joint controllership agreement and collectively assume responsibilities for ensuring compliance with data protection regulations for their jointly controlled processing activities.[19]

For the operator of a data space, it is crucial to provide data controllers with the necessary tools to facilitate the implementation of GDPR compliance for data providers. This can be achieved by offering standardized documentation so that data controllers and processors involved can readily adapt and finalize as needed. However, it's worth noting that achieving a fully automated compilation of relevant documentation is likely a long-term goal that requires further development (Duisberg, 2022, p. 75).

4.3.5 Personal and non-personal data

The GDPR as does the FADP has established a broad reach through its definition of personal data: «personal data means any information relating to an identified or identifiable natural person; an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, on online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person»[20] or according to the FADP «any information relating to an identified or identifiable natural person»[21].

Anonymizing personal data can allow a data controller to process the data without having to adhere to data protection regulations, as long as the data can no longer be linked to an identified or identifiable

[19] Art. 26 GDPR

[20] Art. 4 para. 1 GDPR

[21] Art. 5 lit. a FADP

individual.[22] It's important to note that pseudonymizing personal data, which involves replacing identifiable information with pseudonyms, is reversible and can only serve as an additional security measure during processing. This becomes evident when considering examples like IP addresses, demonstrating that what qualifies as «personal data» extends beyond initial appearances.[23]

Furthermore, the arrival of Big Data and AI has made it clear that what might initially appear as anonymous data or data unrelated to specific individuals can, when combined with other identifiers, eventually be traced back to individuals. However, it's worth noting that the EU Commission acknowledges this distinction and has regulations in place for «non-personal data» in certain areas, such as the Free Flow of Non-Personal Data Regulation (European Parliament, 2018).

Therefore, any concept related to data spaces, like SINA, should be designed to meet compliance requirements for both personal and non-personal data, recognizing the need to protect individuals' privacy while also allowing for the free flow of non-personal data in specified cases.

4.4 Introduction to legal frameworks and standards

In data spaces, different agreements between the business roles are important for effective data sharing. Table 4 provides an illustrative snapshot showing potential roles that could be implemented. It is not a conclusive but flexible example and therefore not a definitive representation as the actual framework may differ. It is important to note that these dynamics may vary depending on the specific characteristics of the data space and the sensitivity of the data used, with some data requiring heightened security measures.

Table 4 Parties to the contract

Roles	Data Provider	Service Provider	Operator	Data User	End User	
Data Provider	X	Data Processing Agreement	SLAs Data Management Agreement		Data Licensing Agreement Data Security Agreement	
Service Provider	Data Processing Agreement	X	SLAs Data Management Agreement		Data Sharing Agreement Acceptable Use Policy	Privacy Policy User Agreement/ Terms of Service

[22] Art. 2 para. 1 GDPR and Art. 2 par. 1 FADP

[23] BGE 136 II 508 E. 3.8

<table>
<tr><th>Roles</th><th>Data Provider</th><th>Service Provider</th><th>Operator</th><th>Data User</th><th colspan="2">End User</th></tr>
<tr><td>Operator</td><td>Data Management Agreement
SLAs</td><td>SLAs
Data Management Agreement</td><td colspan="2">X</td><td>Data Security Agreement</td><td>X</td></tr>
<tr><td>Data User</td><td>Data Licensing Agreement
Data Security Agreement</td><td>Data Sharing Agreement
Acceptable Use Policy</td><td colspan="2"></td><td>X</td><td></td></tr>
<tr><td>End User</td><td>X</td><td>Privacy Policy
Agreement/Terms of Service</td><td>X</td><td>X</td><td colspan="2">X</td></tr>
</table>

4.5 Legal interoperability

The legal landscape surrounding data sharing within the EU and on a national level is characterized by a lack of specific regulations, relying instead on various agreements and soft laws in data spaces. Notably, both EU and national laws, particularly in Switzerland, lack explicit mandatory or default rules for acts that involve granting access to data between parties. This absence may lead to a potential conflict of forms due to diverse contractual frameworks within and across data spaces, rooted in different jurisdictions (IDSA, 2021, p. 12).

A further complication arises from the necessity for the contractual and policy frameworks to align with the regulatory landscape and comply with relevant rules. Striking a balance between fostering innovation and establishing standardized practices in the realm of data poses a formidable challenge. Achieving this equilibrium is crucial for the scalability of innovative practices. Legal interoperability, a key consideration, necessitates the formulation and adoption of standards and harmonized rules. Additionally, fostering a «common language» becomes paramount for legal interoperability. This shared understanding is essential for navigating the intricate legal aspects of data sharing (IDSA, 2023).

5 Economic and ecological aspects

This chapter focuses on exploring various use cases and associated challenges in the building industry by leveraging the SINA architecture built on the principles of the IDSA. The discussion revolves around multiple scenarios that illustrate the application of data sharing and interoperability in energy management and related services.

The concept of data economics refers to an economic framework in which the interactions among market participants are primarily data-driven (Ochs et al., 2019). Within this framework, various structural arrangements can emerge, including platforms, data collaborations, data spaces, and value creation networks. The valuation of data and data assets is described as context dependent (Wessels et al., 2019). Pricing models derived from this context dependency are contingent on the business model of the platform or the specific objectives of the data exchange.

Data spaces, structured intentionally for distribution, cater not solely to individual companies but extend their reach to ecosystems or entire domains. The implementation of their foundational software infrastructure is imperative prior to the realization of business innovation. In ecosystems, participants are organized within a complex network, with these participants potentially operating at different stages of the value chain or even across various industries. Data ecosystems aim to facilitate the exchange and utilization of data, while grappling with the fundamental conflict faced by data generators. This conflict arises from the tension between the necessity of data exchange and the need to protect or maintain the confidentiality of one's own data assets (Steinbuss et al., 2019).

The IDSA's Reference Architecture Model (see section 2.3) serves as a foundation for the creation of a secure, cross-domain data space through technical standardization. This data space allows businesses to tailor their solutions to their specific requirements. SINA, which is based on the IDS-RAM is intended to be part of this international association to form international standards which will play a crucial role in shaping the landscape of data-driven business models enabled by data spaces.

Table 5 provides an overview of the industrial partners involved in developing the SINA use cases and identifying the potential business cases through a number of onsite workshops.

Table 5 Industrial partners participating in development of SINA use cases

Industrial Partner	Industry
Allthisfuture AG / WWZ Energie AG	Energy
arcade solutions ag	Consulting
Bonacasa AG	Building
esolva AG	Consulting
Intelitec AG Stans	Building
Plutinus	Consulting

Industrial Partner	Industry
Privera AG	Building
St. Galler Stadtwerke SGSW	Energy
SIE SA / TVT Services SA	Energy
Zug Estates AG	Building
SmartGridready / IBT	Consulting
St. Gallen Appenzellische Kraftwerke SAK	Energy
Zukunftsregion Argovia	Consulting
SWL Energie AG	Energy

Appendix A4 details the methodology used to engage with industrial partners and outlines the steps taken to co-develop the SINA use cases.

5.1 Ecosystems for potential use cases for SINA

The utilization of data spaces has the potential to emerge as a pivotal tool for holistic management and optimization within ecosystems (see section 1.4 for an overview). The data space could serve as an enabler for third parties to access all the relevant data for the aggregation, analysis, and application of diverse data sources, revolutionizing the «Building Information», «Smart Home and Comfort», and «Energy and Grid» ecosystems. Each ecosystem plays a unique role in reshaping the industry's operations through integrated data utilization. A selective overview of potential use cases within each ecosystem can be seen in Figure 21.

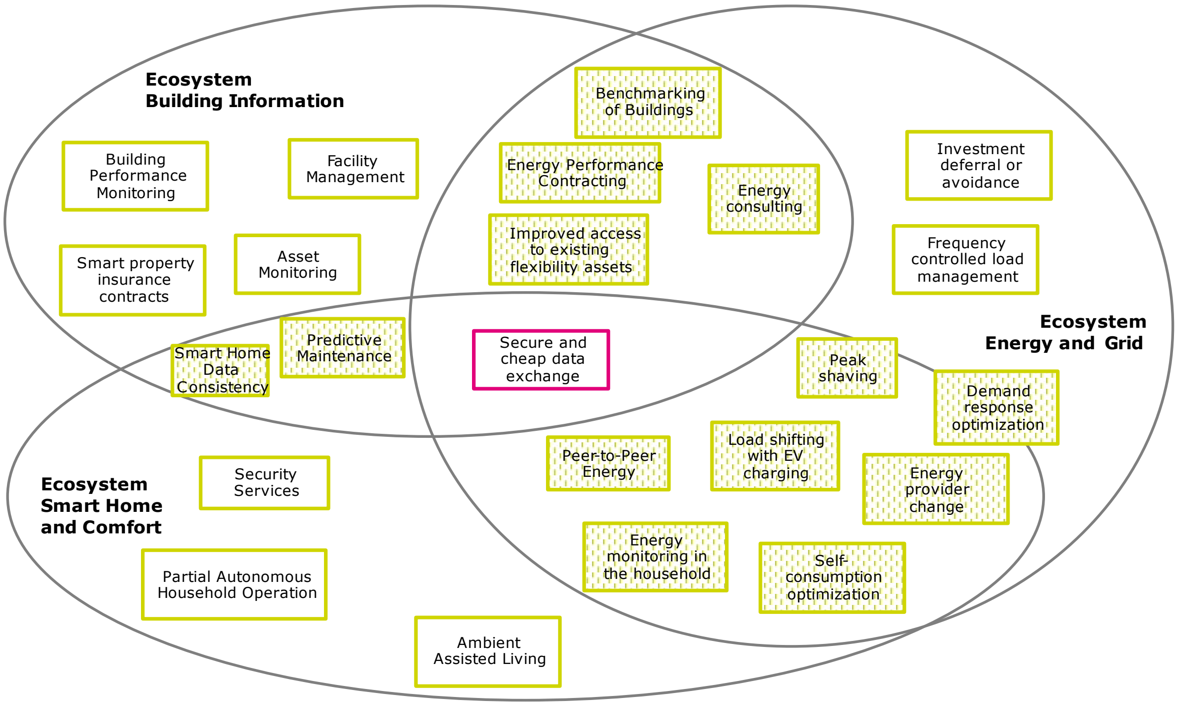

Figure 21 Ecosystems with the identified use cases for SINA

In the following sections the three ecosystems are explained in more detail and potential use cases for each are provided.

5.1.1 Building Information Ecosystem: Enhancing building performance through data solutions

The Building Information Ecosystem is empowered by data solutions, revolutionizing building performance and operational efficiency. An overview of selected use cases is provided in Table 6.

Table 6 Overview of use cases within the Building Information Ecosystem

Name	Brief description
Building Performance Monitoring: Optimizing Operational Insights	Utilizing data spaces enables the collection, analysis, and visualization of real-time data from building sensors. By monitoring temperature, humidity, energy consumption, and air quality, stakeholders gain valuable insights to optimize building performance and identify areas for improvement.
Facility Management: Streamlined Operations for Effectiveness	The data space streamlines facility management by integrating diverse data sources. Predictive maintenance insights and resource optimization tools ensure efficient operations, minimizing unnecessary resource expenditure and enhancing building functionality.

Name	Brief description
Asset Monitoring: Real-time Tracking for Enhanced Efficiency	Implementing asset monitoring systems enables real-time tracking of building assets. This technology enhances maintenance practices, empowering stakeholders to proactively address maintenance needs and ensure optimal operational efficiency as well as having a logbook of historical events.
Benchmarking on Buildings: Performance Evaluation for Improvement	The integration of data spaces enables stakeholders to compare and evaluate building performance against industry benchmarks. This provides a comprehensive assessment that identifies areas for improvement and guides strategies to optimize energy efficiency and operational effectiveness.
Energy Consulting: Tailored Strategies for Efficient Energy Usage	Utilizing data-driven insights, energy consulting services offer tailored recommendations to optimize energy usage and reduce overall operational costs. These recommendations, derived from data insights, streamline energy consumption while maximizing efficiency.
Smart Property Insurance Contracts: Dynamic Risk Management	Data spaces facilitate smart property insurance contracts leveraging IoT and real-time building data. These dynamic contracts are adjusted based on building performance, reducing risks and insurance costs for stakeholders. This innovative approach mitigates risks and promotes a safer, more cost-effective operational environment.
Predictive Maintenance: Efficient System Management	Leveraging historical data within the data space enables the prediction and scheduling of maintenance for critical building systems. This proactive approach minimizes downtime and enhances overall operational efficiency by addressing potential issues before they impact the system's functionality. This use case is further detailed in section 5.3.2.

The integration of data solutions within the Building Information Ecosystem optimizes building performance, fosters efficiency, and facilitates proactive, data-driven decision-making and therefore supports the transformation of building management and operations.

5.1.2 Smart Home and Comfort Ecosystem: Optimizing energy management and comfort

The Smart Home and Comfort ecosystem, empowered by data spaces, introduces innovative solutions for efficient energy management and smart living within households or office buildings. An overview of selected use cases is provided in Table 7.

Table 7 Overview of use cases within the Smart Home and Comfort Ecosystem

Name	Brief description
Ambient Assisted Living: Automated Solutions for Enhanced Comfort and Safety	HEMS uses sensor data and AI to provide ambient assisted living solutions, automating environmental adjustments for the safety and comfort of elderly or disabled individuals. These solutions enhance living conditions and ensure a safer home environment.
Energy Consumption Analysis: Informed Decision-making for Efficiency	HEMS leverages smart meters and IoT devices to collect and analyze energy consumption data within homes. This data, which can be directly collected over manufacturers clouds via SINA connectors, enables informed decision-making, helping homeowners identify patterns and take actions to reduce energy waste, promoting a more efficient household energy usage.
Smart Appliance Integration: Centralized Control for Energy Optimization	Connecting smart home appliances to the data space facilitates centralized control and optimization of energy usage. Users can customize preferences and schedules while promoting energy-efficient practices within the household and optimizing self-consumption of own (PV-based) electricity generation.
Partial Autonomous Household Operation: Convenience through Automated Control	Implementing partial autonomous household operations via the data space allows for automated control of home systems. This use case is going a step beyond smart appliance integration by integrating third-party signal (e.g. dynamic tariffs or forecasting to further enhance convenience for homeowners by automating various household systems) whilst guaranteeing no loss in comfort.
EV Charging Management: Optimizing Electric Vehicle Charging	Leveraging the data space, the electric vehicle (EV) charging is managed and charging schedules optimized based on energy demand and cost while considering grid constraints. This ensures efficient charging while balancing household and grid requirements. This use case can be seen as a subcase of «Smart Appliance Integration: Centralized Control for Energy Optimization».
Demand Response Optimization: Supporting Grid Stability and Efficiency	Utilizing the data space, HEMS participates in demand response programs, adjusting energy usage during peak hours to support grid stability. These adjustments not only benefit the overall energy system but may also incentivize homeowners for their contributions. This use case is shared with the Energy and Grid Ecosystem.

The smart home and comfort ecosystem, supported by data spaces, presents a broad array of solutions for efficient energy management, smart living, and sustainability within households, enhancing overall energy efficiency and the quality of life for homeowners.

5.1.3 Energy and Grid Ecosystem: Stability and optimization of the electric grid

The evolution towards smarter and more resilient grids hinges upon the pivotal role of data sharing and collection. From optimizing energy distribution and integrating renewable sources to empowering consumers and fortifying cybersecurity, the comprehensive utilization of data stands as the cornerstone of innovation and efficiency in the realm of grid management. An overview of selected use cases is provided in Table 8.

Table 8 Overview of use cases within the Energy and Grid Ecosystem

Name	Brief description
Peak Shaving: Managing Energy Consumption for Grid Stability	Utilizing the data space, peak shaving strategies are implemented to manage energy consumption during peak hours. This ensures a more balanced distribution of energy, reducing strain on the grid and promoting stable and efficient energy delivery which avoids costly infrastructure upgrades. Improved demand-side management and efficiency measures lead to more prudent resource utilization and decreased reliance on extensive infrastructure investments.
Grid Congestion Relief: Optimizing Energy Flow	The data space is employed to identify and alleviate grid congestion by managing the flow of energy. Optimization of distribution and avoidance of overloads in specific areas guarantee smoother energy delivery, ensuring a robust grid system. This leads to avoiding costly infrastructure upgrades such as grid reinforcement. Improved demand-side management and efficiency measures lead to more prudent resource utilization and decreased reliance on extensive infrastructure investments.
Improved Access to Flexibility Assets: Balancing Grid Operations	Data space facilitates efficient access and deployment of flexibility assets, such as demand response and storage systems. This balanced deployment ensures optimal utilization, contributing to grid stability (supporting the use cases Peak Shaving and Grid Congestion Relief) and the efficient balancing of energy demand with supply.
Load Forecasting: Strategic Demand Planning	By analyzing historical data and real-time inputs from various sources, load forecasting anticipates energy demand. This proactive planning aids in grid operations, allowing for the anticipation of peak loads, supporting the use case Peak Shaving and the management of the use case Grid Congestion Relief.
Distributed Energy Resource Management: Optimizing Decentralized Sources	Integrating data from decentralized energy sources allows for the optimal use of solar panels, batteries, and electric vehicles, contributing to a balanced grid demand and the efficient management of decentralized energy resources. This can be part of a self-consumption association (ZEV) or a local electricity community (LEGs - lokale Elektrizitätsgemeinschaften).

Name	Brief description
Grid Anomaly Detection and Resilience: Proactive Grid Maintenance	Data from multiple sources is utilized to detect anomalies, such as voltage fluctuations or outages. Proactive responses ensure grid resilience and stability, ensuring an efficient and reliable energy distribution system.
Energy Provider Change: Enhancing Market Competitiveness	Leveraging the data space enables smoother transitions between energy providers. This seamless integration and interoperability foster an open, competitive energy market, encouraging consumer choice and market competition. This use case is further detailed in section 5.3.1.
Peer-to-peer Energy Trading: Community-based Energy Sharing	The data space facilitates peer-to-peer energy trading among homeowners, enabling the direct exchange of surplus energy. This fosters community-based energy sharing, reducing reliance on traditional energy providers and promoting a collective approach to increase the electricity within the low-voltage grid. This use case is further detailed in section 5.3.3.

The energy and grid ecosystem, fortified by data spaces, introduces sophisticated strategies and responses to ensure stability, efficiency, and resilience within the electricity grid, revolutionizing the landscape of electricity management and supply.

5.2 Economic and ecological benefits of data space utilization in the building industry

The integration of SINA within the building industry brings forth a transformative landscape, offering economic and environmental advantages. The streamlined operations and resource allocation, guided by data-driven insights enabled by SINA, could result in cost reductions. The informed decision-making stemming from the utilization of data spaces might minimize wastage and optimize energy usage, effectively reducing overall operational expenses. Moreover, the seamless integration and interoperability enabled by data spaces between energy providers enhance market competitiveness, potentially leading to reduced costs and offering consumers a wider spectrum of choices. This synergy of cost reduction and market competitiveness not only benefits stakeholders but also potentially translates into financial incentives, bolstering economic gains through participation in demand response programs and optimized energy consumption strategies.

On the environmental front, the adoption of SINA could contribute to energy efficiency and sustainability within the building industry. Strategies driven by data insights lead to a tangible reduction in energy consumption, fostering efficiency and conservation. Integration of renewable energy sources and decentralized resource management promotes sustainable practices, significantly reducing the carbon footprint of buildings. Furthermore, stability initiatives, such as peak shaving and grid congestion relief, can avoid grid extension costs and ensure more stable grid operations. The environmental benefits extend further, as the optimization of renewable resources results in a substantial decrease in greenhouse gas emissions, promoting a cleaner and more sustainable environment. Moreover, proactive measures in maintenance and grid anomaly responses amplify environmental resilience, reducing the environmental impact of potential energy disruptions. In conclusion accessing data through a data space like SINA offers multiple opportunities to reduce energy consumption, directly or indirectly, and significantly cut down on CO_2 emissions.

During the discussions of the use cases with the industry partners, it became clear that it is still very difficult to make quantitative assessments of the savings at this stage. This observation has been confirmed in various studies that have investigated the use of data spaces (AKW Group & E-Bridge, 2021; Pöyry Norway AS, 2019). For this reason, we have considered a mixed approach to describe the impacts and benefits of data spaces like SINA: first we assess the impact on a qualitative basis, then the analysis is based on existing calculations. In particular, the extent to which SINA and the data spaces approach can be compared with existing concepts is examined.

Based on a study of the potential impact of the usage of data spaces (section 5.2.1), three topics were studied in more details:

– The net present value (NPV) of SINA as a data space for energy data applications (section 5.2.2).
– CO_2 saving potential when using SINA as a national platform for the exchange of metering data and the additional savings resulting from the expansion of the HEMS installations (section 5.2.3).
– The data transfer and storage costs savings when using data spaces (section 5.2.4).

The integration of SINA within the building industry can manifest a dual benefit - a substantial enhancement in operational efficiency, cost reductions, and consumer choice, intertwined with contributions to sustainability and environmental conservation. The transformative impact of data spaces stretches beyond mere economic gains, presenting a fundamental shift towards a more efficient, sustainable, and environmentally conscious building industry.

It is important to reiterate that data spaces have not only economic and ecological aspects. Creating national data spaces would not just ensure a technological advantage but would also stimulate innovation within our country. The true potential of such infrastructure lies in its ability to inspire and create revolutionary solutions that have not yet been imagined or realized.

5.2.1 Potential impacts of widespread usage of SINA within the building industry

In the building industry, SINA transcends its role as a catalyst for diverse business cases. Beyond its economic impacts, SINA's ramifications permeate the social and ecological realms, with energy conservation emerging as a pivotal dimension.

Impact 1 – Redefining HEMS. SINA's introduction revolutionizes the landscape of HEMS. According to our calculations, outlined in section 5.2.3, the utilization of HEMS presents significant possibilities for both energy and CO_2 reduction.

Impact 2 – Streamlined data transfer and storage in HEMS. SINA transforms the data paradigm for HEMS providers by mitigating the necessity to store data in proprietary cloud systems. By receiving data as a service from OEM/utility clouds, HEMS providers obtain only necessary data, contributing a reduction in data storage and transfer requirements, as this is further detailed in section 5.2.4.

Impact 3 – Enhanced operational efficiency. SINA's implementation facilitates streamlined operations within buildings. The data space optimizes resource allocation, streamlines facility management, and enables predictive maintenance, contributing to increased operational efficiency.

Impact 4 – Market competitiveness and consumer choice. SINA's role in enabling seamless transitions between energy providers or access to offers from energy service companies amplifies market competitiveness and enhances consumer choice. By fostering interoperability, it cultivates a diverse marketplace, benefitting consumers with a wider array of energy service options (potential benefits of a central data exchange platform for metering data are further elaborated by Pöyry (2019)). Data space connectors can act as an enabler for a secure connection to the platform (e.g. for the DSOs) while guaranteeing data sovereignty and data protection compliance.

Impact 5 – Resilient grid operations. The utilization of SINA contributes significantly to grid stability. Strategies like peak shaving, grid congestion relief, and anomaly detection enhance the reliability and resilience of energy grids, reducing energy wastage and enhancing the overall energy distribution system.

Impact 6 – Environmental stewardship. The comprehensive adoption of SINA aligns with broader environmental sustainability goals. Reduced energy consumption, optimized use of renewable resources, and minimized greenhouse gas emissions are integral to achieving a cleaner and more sustainable environment as further detailed in section 5.2.3.

Impact 7 – Technological innovation and job creation. SINA's integration fosters a culture of innovation within the building industry, leading to the creation of new job opportunities and skill development in the technology and energy sectors, thus stimulating economic growth.

Impact 8 – Data security and privacy. The emphasis on data space and interoperability also addresses concerns related to data security and privacy. SINA's approach ensures robust security protocols and privacy measures, bolstering consumer trust and confidence in data handling.

Impact 9 – Urban development and smart cities. Data spaces have the potential to drive the development of the future smart city by facilitating the integration of diverse open and urban data from various municipalities, as well as data from other stakeholders spanning city borders. Therefore, the widespread utilization of an interoperability platform like SINA lays the groundwork for the development of smart cities by improving infrastructure, fostering interoperability and sustainability, and creating more efficient urban spaces.

5.2.2 Cost-benefit analysis of a data space infrastructure in Switzerland

In the following cost-benefit assessment, the use of SINA as a Swiss data space for energy data is examined. As a data infrastructure, it provides the basis for various use cases in the energy sector, such as the ecosystems considered in this study and the corresponding use cases. The net present value (NPV) is used as a financial indicator to assess the profitability of an investment.

The following calculations are based on the approaches identified in the past by the THEMA Consulting Group & devoteam (2018) and the more recent study by Misurio (2020) on improving data exchange in the Swiss energy sector. The study by AKW Group & E-Bridge (2021), which considers the relevant use cases of energy-based business models in the electricity sector, is used as the basis for the comparison of the scenarios.

Data space infrastructure approaches

Data exchange within the energy sector can be organized based on different infrastructures. Today, market players often engage in bilateral communication through a variety of non-standardized interfaces. A similar approach is the decentralized exchange between market partners (P2P approach). However, much optimization potential is seen in a platform-based solution of the data infrastructure in the form of a Datahub Light and Datahub Full (AKW Group & E-Bridge, 2021).

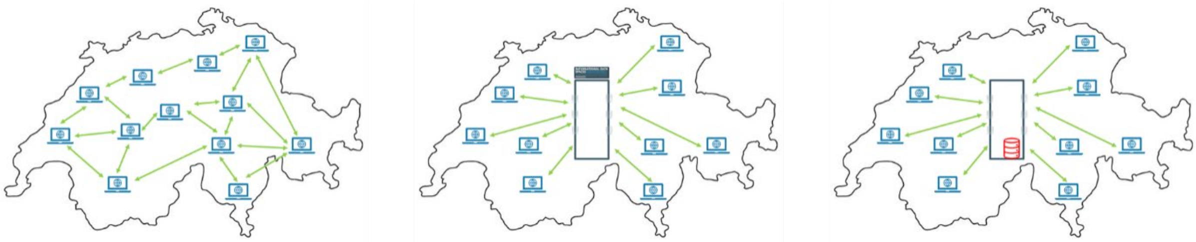

Figure 22 Infrastructure approaches: P2P approach, Datahub Light and Datahub Full

The three different infrastructure approaches have the following characteristics (see Figure 22):

c. Decentralized data infrastructure (P2P approach): A decentralized network based on so-called peer-to-peer communication, in which every market player must establish a connection with every other market player.

d. Datahub Light: Centralized, platform-based approach in which the market players communicate with each other via a central unit. Routing of data. No data is stored, except the necessary user data to establish the connections.

e. Datahub Full: Is based on the Datahub Light. In addition, data is stored centrally, and value-added functions (e.g., automated billing) can be provided, as well as extended statistics, analyses, or plausibility checks of data.

In its current form, SINA or the framework of the IDS corresponds to a Datahub Light: SINA handles data routing, measurement data is not stored. The user data relevant to routing and security is stored in the data space's functional layer. The storage of data and the other functions that characterize the Datahub Full can be provided and made available by various data consumers in the SINA concept. Thus, the SINA data infrastructure is summarized as follows:

– The structure and operation of SINA corresponds to the Datahub Light model.

– The functions and services that can be realized with SINA can match those of the Datahub Full, assuming they are made available by the various data consumers.

Benefits of the use cases and net present value

The AKW Group & E-Bridge (2021) estimated and quantified the benefits, i.e. the cost savings, of a Datahub Light or Datahub Full for various electricity market use cases (see Table 9).

Table 9 Estimates of the benefits (in million CHF per year) in the Swiss electricity market. Source: AKW Group & E-Bridge (2021)

Parameter	Cost savings	
	Datahub Light	Datahub Full
Measurement data exchange	2.1 – 8.4	3.5 – 14.0
Moving in/out, relocation	5.4 – 10.5	6.8 – 13.1
Change of provider	3.5 – 15.3	4.4 – 19.1
Supply management	7.1 – 30.6	8.8 – 38.2
Access to external suppliers	2.2 – 8.7	2.7 – 10.6
Flexibility	0.1 – 0.6	0.2 – 0.7

Taking into account the benefits as well as the costs for setting up and operating the two variants of data hubs, the previous studies (AKW Group & E-Bridge, 2021; Misurio, 2020; Thema Consulting Group & devoteam, 2018) always resulted in a positive net present value for the electricity market as a result of the cost-benefit analysis. AKW Group & E-Bridge (2021) calculated a net present value range of CHF 68.6 million - CHF 406.3 million for the Datahub Light and CHF 69.3 million - CHF 507.5 million for the Datahub Full in the event of full electricity market liberalization and considering the extremes of the options (maximum costs/minimum benefit gain and minimum costs/maximum benefit gain).

Considering the data infrastructure outlined for SINA, the scenario presents the following situation: the costs for structure and operation align with the Datahub Light, while the benefits of functions and services correspond to the Datahub Full. The calculations of the net present value are based on the same values for the lifetime of the data hub and the discount rate as the earlier studies (AKW Group & E-Bridge, 2021; Thema Consulting Group & devoteam, 2018). A lifetime of the data space of seven years and a discount rate of 4.5% were assumed. The scenario results in the following extremes: with minimum costs and maximum benefits, a NPV of CHF 534.2 million is achieved and with maximum costs and minimum benefits, a NPV of CHF 103.7 million (Table 10). Both values are therefore slightly higher than those calculated in the previous study (AKW Group & E-Bridge, 2021). However, it must be pointed out that these values are also subject to uncertainty. Furthermore, certain costs incurred by Datahub Full (e.g., operation of the database, infrastructure for services) are not incurred in the data space but are incurred by the data consumer offering the corresponding services. In this setup, however, the offering of services is open to every data consumer. This can be crucial for the development of innovative services based on data. It reduces the barriers to market entry and leads to more competition in the market. Ultimately, it can be assumed that this will lead to better selection and lower prices for the services offered.

Table 10 Estimates of the net present value (in thousand CHF) for SINA

Parameter	Scenario SINA: Costs Datahub Light, Benefits Datahub Full	
	Minimum costs/ Maximum benefits	Maximum costs/ Minimum benefits
Total one-off costs	9’450	20’650
Recurring costs (incl. deprecia-tion)	3’850	5’300
Benefit gain	96’100	26’400
Net Present Value (NPV)	**534’152**	**103’686**

5.2.3 CO_2 saving potential

It is assumed that a data infrastructure such as SINA helps to save CO_2. On the one hand, savings are made possible by the shared infrastructure and, on the other, indirectly through the functions that the platform enables.

With the inclusion of a study by Pöyry (2019), the savings that a central data exchange platform, such as SINA, could enable in the Swiss energy sector are examined. Furthermore, the study discusses the CO_2 savings that could be achieved if SINA would increase the use of HEMS.

Savings due to an EU-wide data exchange platform

The study by Pöyry (2019) quantifies the benefits of a Europe-wide data exchange platform (DEP) for the energy sector. The framework of data spaces or SINA could be used to build such a platform. This platform is considered as an interoperability platform that will offer retailers, energy service providers, and other eligible market participants a unified and standardized access point to consumers' metering data. Based on the areas «Retail Market», «Energy Efficiency», «Flexibility», «Innovative Use cases» and «general IT costs», the qualitative advantages of a DEP are compared to the counterfactual. The counterfactual is the scenario, without a Europe-wide data exchange, where data is stored and analyzed in various locations or countries (decentralized P2P approach), whereby each EU Member State would develop its own solution.

In the «Energy Efficiency» area, the following use cases, made possible by the DEP, are considered for the evaluation: energy management audits, energy monitoring services and informing consumers of consumption and impact. The resulting consequences for Europe are: energy saved, reduced grid losses, and reduced CO_2 emissions. In the study, the reduction of electricity demand in the residential sector is mainly attributed to targeted energy-based services from third parties. In 2022, the European countries included in the study (total 26) had around 226.3 million households (Eurostat Data Browser, 2023; Statista, 2023). If the calculated values of the study are broken down into households, the following picture emerges (see Table 11).

Table 11 The potential energy and CO_2 savings (per household, per year)

Parameter	Energy saved reduced consumption		Energy saved reduced lower grid losses		Total	
	kWh	kg CO_2	kWh	kg CO_2	kWh	kg CO_2
DEP	108.7	36.2	29.6	9.7	138.4	46.0
Counterfactual	98.6	32.7	27.0	8.8	125.5	41.6
Improvement	10.2	3.5	2.7	0.9	12.8	4.4

If a DEP is used, a household consumes 109 kWh less energy per year. An additional 29.6 kWh can be saved on the grid. In total, 46 kg of CO_2 can be saved. In any case, the DEP scenario performs better than the counterfactual.

These calculations are now used as the basis for the Swiss case. This case assumes the existence of an interoperability platform in Switzerland that would serve as a universal access point for metering data accessible to all market participants. This scenario is akin to a Europe-wide DEP. Calculated for Switzerland, with 3.9 million households (Federal Statistical Office, 2023), the use of a DEP for the «Energy Efficiency» area would results in total savings of 540 GWh a year or 179 kTons of CO_2.

Savings in Switzerland due to the use of HEMS

It is expected that a data infrastructure such as SINA would serve as an enabler for various use cases. One possible effect would be, for example, to increase the market for the use of HEMS solutions. The energy reduction potential when using a HEMS is estimated in various studies to be between 5 % and 70 % (Gualandri & Kuzior, 2023; Karlin et al., 2015; Tiko, n.d.; Tuomela et al., 2021).

Based on these figures the following impact regarding increased energy and CO_2 saving through broad diffusion of HEMS can be calculated. A SINA-based solution is believed to be more cost-effective than a gateway-based solution for HEMS. This makes HEMS more appealing to the 3.9 million private households in Switzerland (Federal Statistical Office, 2023). These households consume roughly 19 TWh per year, approximately 33% (Bundesamt für Energie BFE, 2023) of the overall energy demand in Switzerland.

It is considered realistic that up to 30% (1.17 million) of all households will deploy a HEMS within five years from the deployment of SINA. Based on the previous figures and expert inputs from the Lucerne University of Applied Sciences' «iHome Lab» and The Institute of Building Technology and Energy, we estimated that savings of up to 35% are realistic.

Under this condition, SINA could totally contribute to an energy saving of up to 2.0 TWh and roughly 0.24 Mtons CO_2 emissions (Our World in Data, 2023) annually (see also Table 12). This means annual savings of 1705 kWh or 205 kg CO_2 per household. The potential savings estimate for all 3.9 million households in Switzerland amounts to 6.65 TWh and 0.8 Mton CO_2.

Table 12 Energy and CO_2 savings per year due to the use of HEMS in Switzerland

	3.9 million households (100%)	3.12 million households (80%)	1.95 million households (50%)	1.17 million households (30%)	1 house-hold
Consum-mation households	19 TWh	15.2 TWh	9.5 TWh	5.7 TWh	4872 kWh
Consum-mation equivalent CO_2	2.28 Mton	1.82 Mton	1.14 Mton	0.68 Mton	585 kg
Savings HEMS (30%)	6.65 TWh	5.32 TWh	3.33 TWh	2.0 TWh	1705 kWh
Savings equivalent CO_2	0.8 Mton	0.64 Mton	0.4 Mton	0.24 Mton	205 kg

The comparison of these calculations with those of the EU-wide data exchange platform shows that the figures are very difficult to compare and demonstrate how difficult it is to calculate such values.

5.2.4 Reduction potential of data transfer and storage cost

One added value of data platforms such as SINA is that they reduce general transaction costs for the involved stakeholder. This is partly an effect of simplified interaction via a common technical platform and partly a result of the specific platform services. And further costs are reduced through the standardization of communication and contract components, and the coordinated ecosystem (Roland Berger, 2018).

Given the data infrastructure underpinning the data spaces, the data owners do not have to store the data they need to operate their services in their own proprietary cloud system. Instead, they receive the data from the corresponding data providers, e.g., from the manufacturers' clouds. The service providers must only receive the data they need. It can therefore be assumed that there will be savings in data transfer and storage costs across the entire data transfer value chain.

In the example of a HEMS application, the HEMS function does not have to be carried out in each individual household; instead, the corresponding data is delivered to a central HEMS service (Figure 23).

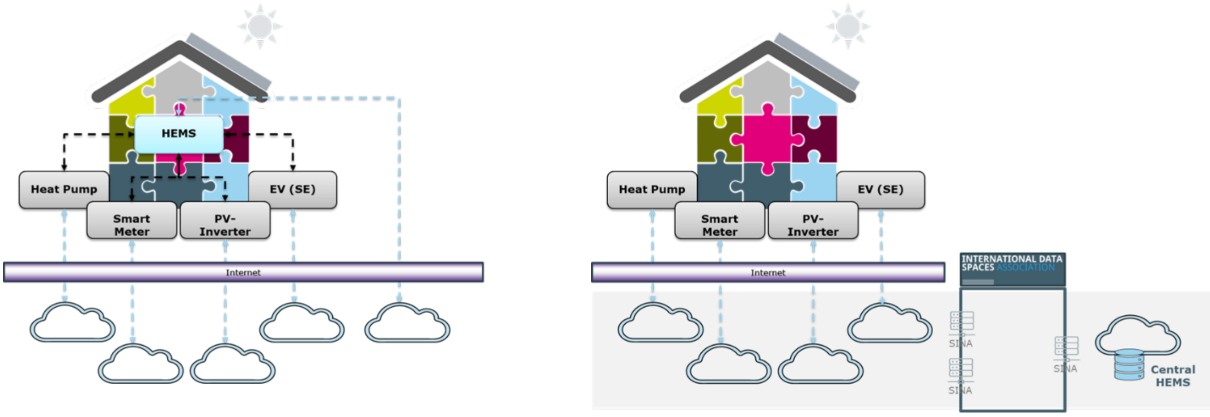

Figure 23 Local and central HEMS

As the SINA data space currently only exists as a demonstrator, it is difficult to estimate in which areas less data flow or data storage will take place, which could ultimately lead to energy and CO_2 savings. No studies could be found that attempted to quantify these values in a similar setting. As rightly noted by Morley et al. (2018), there is great anticipation surrounding «smart» innovations and their potential to reduce energy demand. However, proving this impact is challenging. For instance, the exact amount of electricity consumed by the Internet remains undefined.

The expenses related to transferring, processing, sharing, and storing data between systems or entities are discussed below. These costs can take various forms within a data space environment. The table provided highlights important aspects of data transaction costs and their connection to data spaces such as SINA. An assessment was conducted to determine how data transfer and data storage is minimized (Table 13).

Table 13 Costs associated with the transmission, processing, sharing and storage of data in a data space environment

Costs	Description	SINA
Data Transfer Costs	Costs that arise when data is moved from one location to another, e.g., data transfer between data provider and data consumer. The cost may depend mainly on the volume of data but also on the network infrastructure used.	- A targeted selection of the data to be received by the data consumer can reduce the data volume and thus costs. - Multiple queries of the data could be necessary as the data is not saved, which in turn would increase the data volume and increase the costs.
Data Processing Costs	Costs that are incurred by processing, transforming, or analyzing data. They are related to computing resources and the required operations.	- Reduced data processing can occur if certain data is only processed by one data customer and the outcomes are made available to other stakeholders via the data space. Then this cost would decrease.

Costs	Description	SINA
Storage Costs	Costs for the operation of physical storage devices, cloud storage services and maintenance. The amount of data stored, and the duration of its retention can affect these costs.	- The amount of stored data is reduced because the same data is not stored in multiple locations. This reduces storage costs.
Data Governance and Compliance Costs	Ensuring compliance with regulations and internal policies incurs costs. The necessary measures include monitoring data transactions, ensuring security, and managing the data environment.	- Data governance and compliance are provided by the data space framework and should lower these costs. - Access control to data is provided by the data space framework. Thus, this cost should decrease for the involved parties.
Data Integration Costs	Integration costs may occur when combining data from different sources or systems. These may include interface development, data mapping, and ensuring compatibility between data formats.	- Stakeholders only have to adapt and maintain one interface. The standardization or reduction of interfaces significantly reduces the transaction costs for data access and also the data integration costs.

As can be seen in the overview provided in Table 13, data processing and storage costs can be reduced with a data space infrastructure. The standardized governance, compliance and data interfaces also simplify the operation of the infrastructure and contribute to cost reduction.

5.3 In-depth use case description for applying SINA

This section delves into the practical application of a data space based on use cases. Accessing data within trusted and collaborative cloud infrastructures is essential to deliver comprehensive solutions while upholding stringent security standards. These elements play a pivotal role in building trust and transparency, essential for scaling operations, digitalizing processes, as described in section 5.3.1, and introducing innovative services such as predictive maintenance (section 5.3.2). Further, section 5.3.3 emphasizes the necessity for shared cross-border knowledge representations, semantic data models, and robust capabilities for collecting and sharing data to enable the use case P2P energy trading.

5.3.1 Smoother data exchange between energy providers: Data-driven integration

Brief description of the use case

In this use case, the data space is harnessed to facilitate smoother data exchange between energy providers (e.g., the transition between energy providers for prosumers selling their surplus PV production) within an open, competitive market. The seamless integration and interoperability are provided by standardized and automated procedures. This could for example be applied in changing the energy provider enabled by the connectors accessing the data space.

Data and insights

Data sources encompass smart meter data but could not be limited to it. Historical and real-time data on energy consumption patterns, market pricing, and consumer preferences as well as flexibility potential of IoT devices can contribute to a comprehensive understanding of energy usage and market dynamics.

Value proposition

Leveraging the power of data analytics, this data space streamlines transitions between energy providers, promoting an open and competitive energy market. Seamless integration ensures consumer empowerment, choice, and fosters healthy market competition. An extended use of potential data could also enable real-time insights and improved forecasting within the distribution grid.

Primary objectives

The primary objectives include:

- Facilitate smoother transitions between energy providers.
- Supporting DSOs with digitization and automation of processes.
- Empower consumers with choices in an open market.
- Foster transparency and a healthy competition within the energy market.

SINA roles and actors

This section explores the key roles and actors contributing to the data-driven integration for smoother energy provider transitions in an open market. The integration of data-driven technologies helps optimize processes, enhance transparency, and improve overall efficiency.

Table 14 Actors within the use case smother data exchange between energy providers

Name	Brief description
Distribution grid operator (DSO)	The DSO is the operator of the distribution system, whose main task is to deliver electricity to customers in a cost-effective, sustainable, and efficient way. In the energy provider change use case, the DSO is managing the administrative tasks of managing data of customers who are willing to change providers.

Name	Brief description
Consumers and Prosumers	Entities purchasing and selling energy for their energy requirements and are customers of the DSOs
Platform provider	Technology providers facilitating the platform for enabling a smooth energy provider change in an automated and efficient way.
Regulatory Bodies and Energy Authorities	Government or regulatory entities overseeing energy market compliance and providing the market framework

Value capture and monetization

In Table 15 an overview of the different monetization options for the use case is provided.

Table 15 Description of monetization options for the use case

Name	Brief description
Data Access and API Usage	Charge fees for direct access to the data space and the use of APIs for integrating data into external systems. This can be a revenue stream for third-party applications and services.
Custom Analytics Services	Offer customized analytics services to energy providers, allowing them to extract tailored insights for strategic decision-making. This can include predictive modeling, market trend analysis, and consumer behavior predictions.
Partnership and Collaboration	Collaborate with energy providers, IoT device manufacturers, and other stakeholders to jointly develop innovative solutions. Create partnerships that leverage the data space for mutual benefit. Establish revenue-sharing models with partners based on the value generated through joint initiatives. Monetize collaboration efforts by offering joint services to a broader market.

5.3.2 Data-driven facility management solution with advanced predictive maintenance capabilities

In the following sections the use case of predictive maintenance enabled by using data space technology is described.

Brief description of the use case

Utilizing AI-driven analytics, this innovative data space offers enhanced predictive maintenance solutions for facility management. The system leverages advanced analytics to optimize daily maintenance tasks, such as asset management, operational decision-making, and maintenance scheduling. Facility managers gain access to up-to-date, real-time information regarding predictive maintenance needs, proactive equipment fault diagnosis, and the likelihood of equipment failure, allowing timely replacements and maintenance of end-of-life components.

Data and insights

As owners of data assets, facility managers will source data from various sensors including SCADA, local inverters, meters, portable cameras, UAV-attached cameras, and local weather stations.

Value proposition

The business application of a third party or within the data space analyzes multiple data streams and provides facility managers with crucial intelligence. It ensures optimal facility operation, predictive maintenance, and proactive avoidance of equipment faults. The platform enables real-time identification of maintenance needs, significantly reducing the facility's operational and maintenance costs (O&M). This leads to increased efficiency, reliability, and lower Levelized Cost of Energy (LCOE) while also improving overall facility functions.

Primary objectives

The primary objectives include:

- Provide real-time insights for predictive maintenance needs.
- Enable proactive equipment fault diagnosis.
- Optimize facility operations and maintenance scheduling.
- Reduce operational and maintenance costs.
- Enhance the reliability and efficiency of diverse facility types.

SINA roles and actors

The domain of facility management stands at the brink of transformation due to the potential offered by data-driven solutions and advanced predictive maintenance capabilities. This section aims to explore the diverse roles and essential actors engaged in a comprehensive data space designed for efficient facility management. Through the fusion of technology, data analytics, and operational expertise, the management of facilities has been reshaped. The merger of AI-powered analytics, predictive models, and real-time data streams has empowered facility managers to foresee maintenance requirements, optimize operations, and make informed decisions, potentially enhancing efficiency and decreasing operational expenses. Table 16 provides an overview of the different actors within this use case and their contributions and interactions in leveraging data intelligence for the effective management of diverse facilities.

Table 16 Actors within the use case predictive maintenance

Name	Brief description
Data Asset Providers/Data Asset Owners	Facility management companies/organizations: These entities own, manage, and operate various types of facilities, such as commercial buildings, hospitals, educational institutions, or industrial sites. They are the primary owners of the data generated by the sensors, meters, cameras, and other monitoring devices installed within their facilities.
Data Asset Consumers	Facility management companies/organizations: The same entities that own the data also act as consumers, utilizing the data for internal operations, maintenance planning, and decision-making.
Technical Users	Maintenance technicians and engineers: Individuals responsible for day-to-day maintenance activities, equipment servicing, and facility upkeep. They use the insights derived from data analytics to guide their maintenance routines and troubleshoot issues effectively.
Data Scientists /Data Analysts	Data analysis teams within facility management: This team focuses on analyzing the data, identifying patterns, creating models, and implementing algorithms to predict maintenance needs and optimize facility operations.
Business Users	Facility managers/administrators: These professionals oversee the facility's overall operations and strategy. They use insights from data analysis to make informed decisions about resource allocation, long-term planning, and strategic direction.
Platform and AI Analytics Marketplace	These platforms provide a range of services including data processing, analytics, and a marketplace for AI-driven solutions. The platform charges fees for various services like data quality enhancement, data hosting, and the execution of predictive maintenance analytics through its AI Analytics Marketplace.
Equipment and Sensor Manufacturers	Companies producing monitoring devices: These are the manufacturers creating the sensors, meters, cameras, and other monitoring equipment deployed within the facilities. They might offer support, updates, or integration services for their equipment within the facility management ecosystem.
Maintenance Service Providers	Third-party maintenance contractors: External service providers or contractors offering specialized maintenance services to support facility management. They might integrate with the system to provide additional support based on predictive maintenance data.

To being able to ensure the efficient utilization of data for predictive maintenance and improved facility management, each of these actors plays a crucial role. They are collectively contributing to streamlined operations, reduced costs, and optimized decision-making processes within the facility management landscape.

Value capture and monetization

In the landscape of data-driven facility management, value capture and monetization are pivotal components that underpin the sustainability and effectiveness of the system. This section describes the strategies to capture value and monetize the services, insights, and data-driven solutions offered within the ecosystem (Table 17).

Table 17 Description of monetization options for the use case

Name	Brief description
Service Fees for Data Quality Enhancement	One of the fundamental aspects of data-driven facility management is the enhancement of data quality. This involves ensuring data accuracy, reliability, and accessibility. Facility management entities pay service fees to the platform or service provider for improving the quality and integrity of the data generated within their facilities.
Service Fees for Accessing Data	The data space provides access to a platform to share and/or access the vast amounts of data generated by various sensors and monitoring devices within the facilities. Fees are charged for the organization, and accessibility of this data, making it readily available for analysis and decision-making.
Service Fees for Analytics Execution (as-a-Service)	A third party provides its services in the analytics marketplace within the data space ecosystem and offers predictive maintenance analytics as a service. Facility management entities pay fees for the execution of these analytics, enabling the generation of insights, predictions, and recommendations based on the data collected. This service facilitates informed maintenance decisions and operational optimizations.
Predictive Maintenance Value Quantification	The effective utilization of predictive maintenance tools and insights translates to tangible cost savings and enhanced operational efficiency for facility management entities. By identifying and addressing maintenance needs before they lead to critical failures, these entities reduce downtime, avoid costly repairs, and extend the lifespan of equipment, thereby saving significant operational costs.
Efficiency Gains and Operational Improvements	The implementation of data-driven insights leads to efficiency gains and operational improvements within the managed facilities. Reduced energy consumption, optimized resource allocation, and streamlined maintenance routines contribute to cost savings and improved operational performance.

Name	Brief description
	These improvements translate to real monetary value for facility management entities.
Subscription or Usage-Based Model	The monetization strategy may involve a subscription-based model or usage-based fees, where facility management entities pay for the services based on the volume of data processed, the frequency of analytics execution, or the scale of facilities managed. This model ensures that costs align with actual utilization.
Consultation and Additional Services	The platform or service provider might offer additional consultation or specialized services based on the insights derived from the data. These services provide added value for which facility management entities pay extra fees, further enhancing their ability to make data-informed decisions.

The combination of these strategies ensures that the platform or service provider captures value while providing essential services and insights that contribute to operational efficiency, cost savings, and improved decision-making for the entities managing diverse facilities.

5.3.3 Peer-to-peer energy trading based on SINA

Brief description of the use case

Peer-to-peer (P2P) energy trading is a decentralized energy exchange system that allows individuals or entities to directly trade electricity among themselves. This innovative approach empowers participants to buy and sell excess renewable energy, fostering a more efficient and sustainable energy ecosystem. Prosumers, equipped with renewable energy generation systems, produce excess energy. They utilize smart meters to track surplus production and offer it for sale on the P2P trading platform. Consumers, in need of additional energy, may purchase it directly from these prosumers through the platform, facilitated by smart contracts ensuring secure and transparent transactions.

P2P energy trading is currently in its experimental phase, but it holds the promise of revolutionizing the energy landscape. Advocates argue that it offers a potent solution to meet renewable energy and climate targets while helping to balance supply and demand in an increasingly decentralized electricity grid. Enabled by digital technology, P2P energy trading leverages the increasing generation and capture of data, much like other sectors within the sharing economy. An overview of the key contributions can be seen in Figure 24. If appropriately regulated, it could serve as a cornerstone in the creation of decarbonized energy systems essential for our sustainable future.

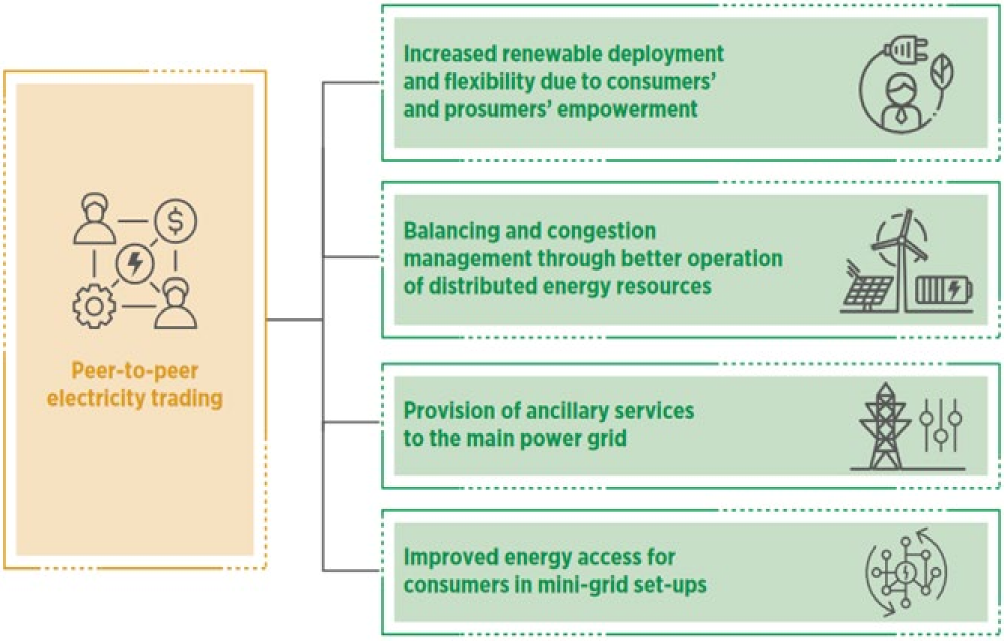

Figure 24 Key contributions of P2P electricity trading to the transformation of the electricity grid. Source: IRENA (2020)

A schematic representation of such a P2P market can be seen in Figure 25.

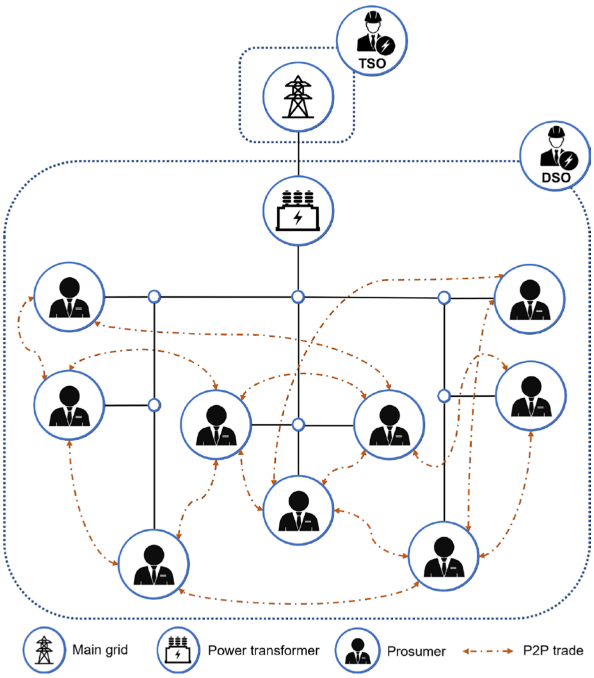

Figure 25 Peer-to-peer flexibility market model. Source: Marques et al. (2023)

Data and insights

The exchange and availability of data plays a pivotal role in enabling P2P energy trading. This data pertains to the consumption and generation behavior of the various actors involved in the P2P trading scheme. Smart metering systems are instrumental in collecting this data, allowing for the efficient operation of P2P energy trading. These systems have advanced significantly, incorporating various technologies, e.g., from sensors, which facilitate innovative approaches to energy management and value production.

The increasing penetration of smart meters, sensors, and the Internet of Things (IoT) means that energy data is now captured at granular levels, even down to individual energy-consuming devices. Machine learning and Artificial Intelligence (AI) have empowered big data analytics, transforming energy data into a valuable asset. Furthermore, distributed ledger technologies like blockchain have brought transparency to energy production, consumption, and demand-side response derived from smart meters. This transparency forms the foundation for value and accounting chains, offering new opportunities for value creation in the energy market.

Data generated within P2P energy trading includes:

– Real-time energy production and consumption data from smart meters.
– Historical energy trading records including prices, quantities, and participants.
– Local weather conditions impacting renewable energy generation.

Insights derived from this data include:

– Forecasting energy surplus and deficit periods.
– Analyzing price trends and demand-supply dynamics.
– Identifying peak trading hours for optimized transactions.

Value proposition

The value proposition of P2P energy trading is comprehensive, offering a direct and transparent avenue for participants to exchange surplus renewable energy. This platform empowers prosumers by enabling the monetization of excess energy while providing consumers access to sustainable energy sources. Furthermore, it fosters a decrease in transmission losses and reduces dependency on centralized energy systems, ultimately fortifying grid resilience. Moreover, the promotion of renewable energy adoption within this framework significantly contributes to environmental sustainability by encouraging cleaner energy practices and reducing carbon footprints.

Primary objectives

The primary objectives include:

– Enable direct buying and selling of surplus energy among participants.
– Facilitate the use of renewable energy sources.
– Empower consumers to have control over their energy usage and distribution.
– Create a more resilient and sustainable energy network.

SINA roles and actors

A P2P energy market involves the decentralized exchange of energy among individual participants, allowing them to generate, consume, and trade energy directly with one another. In such a market, various roles and actors play crucial roles in ensuring the smooth functioning of the system (Table 18).

Table 18 Actors within the use case peer-to-peer energy trading

Name	Brief description
Prosumers (Energy Producers/Consumers)	Individuals or entities generating renewable energy and participating in trading activities.
Consumers (Energy Buyers)	Entities purchasing surplus energy for their energy requirements.
Facility management companies/organiza-tions	These entities own, man-age, and operate various types of facilities, such as commercial buildings, hospitals, educational institutions, or industrial sites. They are the primary owners of the data generated by the sensors, meters, and other monitoring devices installed within their facilities.
Platform provider	Technology providers facilitating secure P2P energy trading. These plat-forms provide a range of services including data processing, analytics, matching and contracting between the different parties.
Regulatory Bodies and Energy Authorities	Government or regulatory entities overseeing energy market compliance and providing the market framework for the P2P energy market.
Distribution grid opera-tor (DSO)	The DSO is the operator of the distribution system, whose main task is to deliver electricity to customers in a cost-effective, sustainable, and effi-cient way. In the P2P use case, the DSO is providing the grid infrastruc-ture. Furthermore, the DSO might also act as a potential buyer of traded energy.
Equipment and Sensor Manufacturers	Companies producing monitoring devices: These are the manufacturers creating the sensors, smart meters, and other monitoring equipment de-ployed within the facilities. They might offer direct access to the devices via the OEM cloud with a data space connector to exploit their flexibility potential.

With the exorbitant rise in energy production costs from fossil fuels such as gas and oil in 2022, the transition to sustainable and renewable energy solutions is being promoted not only in Europe but world-wide. Photovoltaics, especially in large metropolises where tens of thousands of buildings of different sizes and purposes exist side by side, play a central role in producing the necessary amounts of solar energy and ensuring a sustainable energy supply.

Since not all buildings in a district are suitable for the installation of PV systems and the solar energy production of certain (well-suited) buildings is often higher than self-consumption over the course of a day, the other (less suitable) buildings in the same district have so far been unable to benefit from the significant overproduction of solar energy over time. Among other things, there is a lack of secure, standardized, bidirectional interfaces through which the real-time energy needs and every surplus of solar energy productions can be communicated openly between the individual buildings (peers) in the related district.

The acronym P2P comes from the network domain and describes ad-hoc connections that link two systems or applications directly with each other. Such direct connections are very rare in the real world. Almost all network connections are routed via network gateways (e.g., proxies, routers, servers). The primary goal of P2P trading participants is to tackle various challenges related to energy trading, including reducing energy consumption costs, supporting sustainable renewable energy sources (RES), and enhancing social engagement among prosumers.

Value capture and monetization

P2P energy markets hold the potential to create value on multiple fronts. They enable local matching of supply and demand for renewable energy, empower consumers to actively influence energy sourcing, and incentivize investments in renewable generation. This reduction in depletion of natural resources and greenhouse gas emissions aligns with long-term sustainability goals.

Ideally, local P2P energy trading increases revenue for prosumers, driving investments in renewable energy and reducing uncertainty in distributed energy resources. Moreover, P2P energy trading can contribute to the overall management of an electricity network by promoting renewable energy expansion, enhancing flexibility in power generation, balancing supply and demand, and improving access to energy resources. Further, it could also play a role in offering ancillary services (Marques et al., 2023). This section describes the strategies to capture value and monetize the services, insights, and data-driven solutions offered within the ecosystem (Table 19).

Table 19 Description of monetization options for the use case

Name	Brief description
Data Access and API Usage	Charge fees for directly accessing the P2P energy trading via a data space connector and the use of APIs for integrating data into external systems. Fees are charged for the organization, and accessibility of this data, making it readily available for automatically trading and billing within the P2P market.
Subscription Model	Users pay a recurring fee to access the peer-to-peer energy data space, granting them continuous access to real-time energy usage information and transactional capabilities.
Transaction Fees	Platform charges a small fee for each peer-to-peer energy transaction facilitated through the data space, generating revenue based on the volume of transaction.

Name	Brief description
Consultation and Additional Services	The platform or service provider might offer additional consultation or specialized services based on the insights derived from the data. These services provide added value for which utilities or other involved entities pay extra fees, further enhancing their ability to make data-informed decisions.

The next chapter will delve into the details of creating a P2P energy trading proof-of-concept within the SINA project.

6 Implementation aspects based on peer-to-peer energy trading platform

This chapter outlines the creation of a peer-to-peer (P2P) energy trading platform as part of the SINA project, moving towards the implementation of sustainable and renewable energy solutions, based on secure data exchange. The present research work was intended to investigate diverse communication interfaces, build an initial peer-to-peer energy data exchange platform, and assess the whole framework using real energy data gathered from HSLU GEE Lab.

The GEE Lab infrastructure shown in Figure 26 provides us with an entire solar energy production environment, multiple home appliances (e.g., refrigerator, cooker, washing machine) as well as a vehicle charging station. The GEE infrastructure can be exploited to examine real energy data flows to other (virtual) buildings in the nearby area.

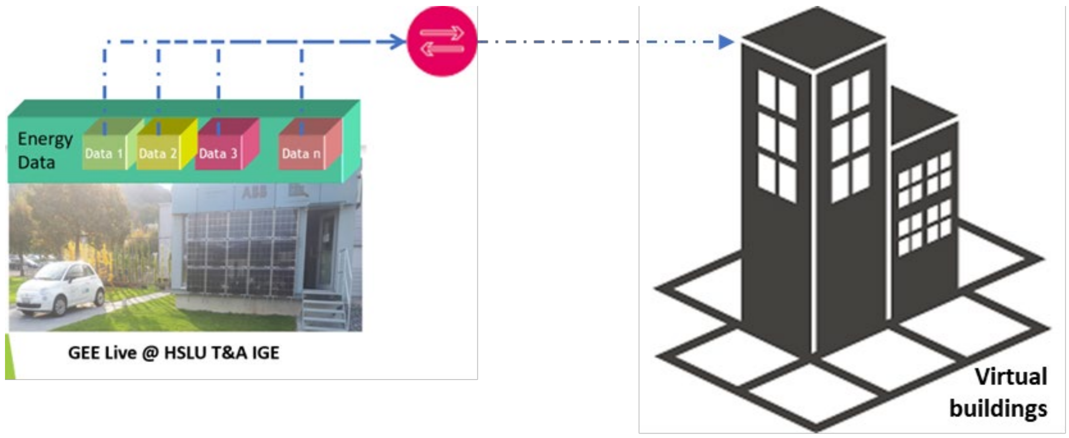

Figure 26 GEE Lab Infrastructure

The outputs of P2P proof of concept are mainly dealing with the data exchange in a specific P2P energy context. Nevertheless, the research is very important to validate and prove that the usage of the standard SINA data spaces connectors will accelerate the interoperability not only between energy applications such as HEMS but also other applications in the building domain e.g., for facility management purposes.

Figure 27 illustrates the role and tasks that a P2P application can take in a local energy network to ensure real-time data connections between «energy consumers» and «energy producers».

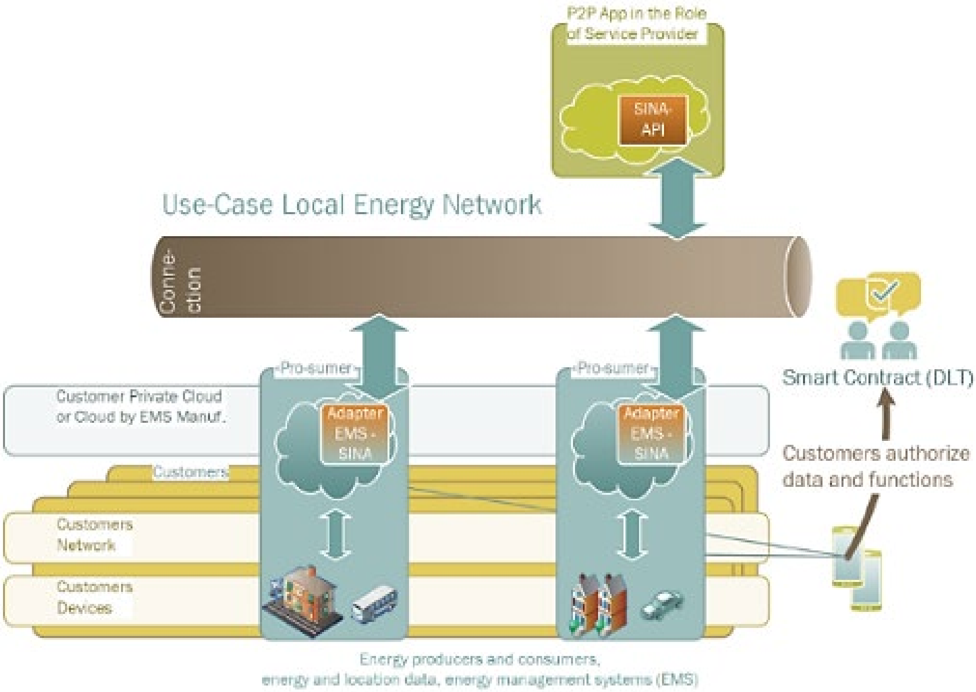

Figure 27 Overview of SINA application for a P2P energy market

The P2P proof of concept addresses multiple technical problems which can be summarized in the following research questions:

– Which buildings' raw energy data are important for building an initial P2P energy network?

– How can related energy data be aggregated, interpreted, and restored for further processing?

– What kind of network protocols and software stacks are required to build core P2P systems?

– How can a P2P network be extended with DLT/BLC to enhance security and redundancy?

Further information to the building energy data that was collected, and how it was used for further processing can be found in Appendix A5.

P2P energy trading demands multi-sided electricity distribution markets that can overcome existing limitations in electricity distribution system management and regulation. It necessitates a trading platform to ensure that individual trades do not create operational constraints. Timestamping generation and outflow metering create the visibility required for subsequent trading and value creation. The creation of differentiation among individual peers can only be achieved through digital trading platforms that establish digital provenance.

However, the use of personal data in the electricity market raises concerns regarding data privacy. Dynamic pricing and aggregation contracts involve the collection of substantial personal consumption data, which must comply with data protection regulations such as the GDPR. This emphasizes the need for sector-specific rules to govern the exchange of data between market actors to ensure compliance with data protection laws.

Several P2P business models rely on blockchain technology for transaction tracking. However, blockchain's immutability creates challenges with data privacy laws. European law, for instance, allows data subjects rights to access, rectify, and erase their personal data. Blockchain's fundamental characteristic of immutability makes it challenging to conform to these rights. To preserve the innovation potential of blockchain technology, experts have proposed a more flexible application of the right of deletion in the context of complex and decentralized IT architectures.

6.1 P2P energy data exchange and analysis framework

Defining which historical and/or forecasted energy data are needed and which associations exist between related data nodes are the first steps enabling an initial inventory of ensemble energy data exchanges and operations that must take place inside the target P2P framework. Figure 28 illustrates data communications between the P2P nodes.

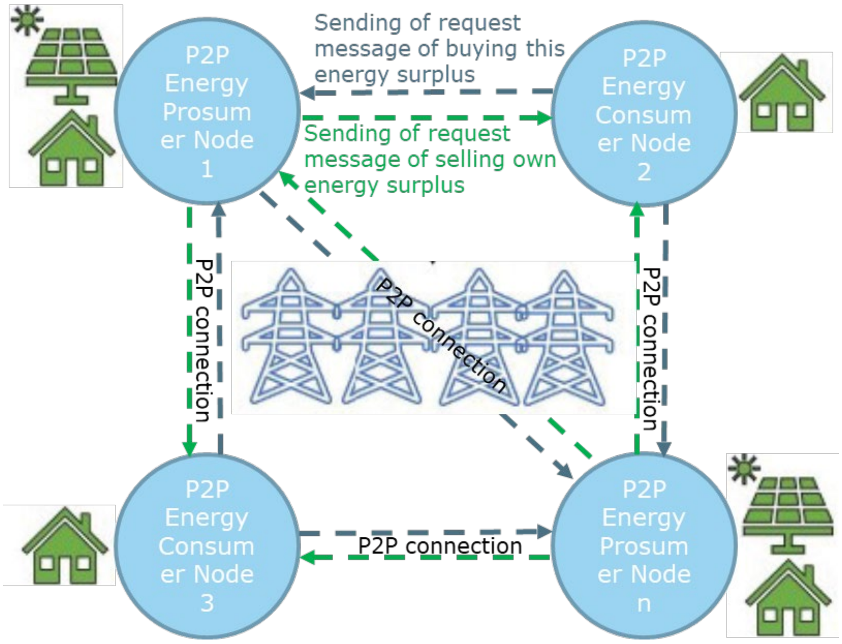

Figure 28 Structure of P2P connections

Each P2P node (consumer or prosumer) participating in the target P2P framework must be designed to enable at least the following services:

- Aggregating and pre-processing its own energy data.
- Estimating its own energy needs for the next day based on historical and/or forecasted consumptions.
- Assessing whether the PV systems installed locally in the building can generate sufficient power covering own electrical energy needs for next day on considering historical and/or forecasts of production and weather data.

- Sending broadcast messages into the local energy network to sell or buy a certain amount of energy.
- Approving and signing a (smart) contract if an agreement is reached and an energy transaction is imminent.

Prosumer nodes may be equipped with energy storage, e.g. a battery. In these cases, related nodes should first store the extra produced energy and then start sending sell broadcast messages to neighbor nodes only when their own batteries are fully charged, and they don’t expect any energy self-demand in the next hours.

6.1.1 Network protocols and software stacks

Every node is responsible for aggregating its own energy data. Therefore, an in-depth analysis of bus and communication protocols (KNX, Modbus, etc.) supported by pre-installed energy consumers (lighting, air conditioning, heating, washing, drying, cooking systems, etc.), producers (PV-production and storage systems) and smart meters on each node is required to get an overview which energy data are logged internally on each system and how to access these data.

Some systems may support IP-based communication and provide web servers enabling access e.g. per Hypertext Transport Protocol Secure (HTTP) to their own logged energy data.

Other systems may not have an interface to the Internet. Those systems may be physically linked to individual smart meters which measure their energy usages, log, and transmit the measured values periodically to a superior application inside or outside the home network. This application may in turn deliver the measured values to any instance soliciting these data.

Each time an energy data bundle (data time series) of an individual building consumer/producer system is requested and received the transmitted data must be pre-processed and saved. For requesting, receiving, preprocessing, and storing related data an integral software programming framework must be integrated. There are no special requirements regarding the software stacks that the programming framework should provide. However, high programing frameworks like Python, .Net and Java are more sophisticated when it comes to developing such kind of backend server applications. These programing frameworks are also providing drivers to interact with the most frequently deployed Data Base Management System (DBMS) (InfluxDB, MongoDB, etc.) used for storing and querying data time series.

For analyzing and forecasting internal energy consumptions and/or productions at a specified node the software programming frameworks, cited already, provide special open sources libraries and tools that can support such processes. Pandas e.g., is an advanced package within Python, which integrates important functions from other libraries like NumPy and scikit-learn. These core functionalities may make data analysis and manipulation of aggregated energy data time series straightforward.

Once one prosumer node estimates their energy productions to be persistently more than their needs in the near future, own software framework must enable opening communication channels to the other nodes participating in the P2P energy network. A unique packet Id must be assigned to each energy data broadcast message sent by the related prosumer node. The message header must include the network address of the source node as well as the address of the target P2P network (broadcast). The payload must also specify the amount of energy to be offered and the period during which the source node can ensure delivering this energy amount.

As soon as another node connected to the P2P network realizes that their energy needs will exceed in the near future self-production capabilities, own software framework must be able to run an internal job to access its own network communication buffer and start reading the energy broadcast messages to check if there is an actual and adequate energy offer between the last messages received from neighbor prosumer nodes. If an offer is available, the software framework should open a direct socket-based connection link between both nodes to exchange and negotiate other in-depth formalities (e.g., pricing).

6.1.2 Integrating smart contract in peer-to-peer energy network

An energy transaction between two nodes participating in a P2P network must be operated automatically and without the intervention of any intermediary third party. However, the finalization of such transactions can be only tackled after both parties accept certain rules, conditions, and approve a digital agreement or smart contract.

Open source DLT/BLC networks like Hyperledger (IBM, n.d.) and Ethereum (Ethereum, n.d.) define a smart contract not as a legal contract but as an executable computer code containing all the required rules for a transaction to be accomplished. A smart contract can thus manage and control energy transactions between two or more parties in the local P2P energy network.

Once both parties in the network sign a P2P smart contract, it will be added to the blockchain, and it is not possible to modify or delete it. If one participating node cannot fulfill the terms mentioned in the P2P smart contract, the other participating node can revoke the signed contract.

The creation, upgrading, and broadcasting (distributing) of needed smart contracts within the target P2P (blockchain) network are operative tasks which are essential and must be therefore done by an additional instance (node) within the network. This third node is responsible for physically controlling and confirming energy injections from the prosumer node and energy transmissions to the consumer node. After mining the smart contract into a block, the smart contract becomes a unique address inside the block, and every participating node can start sending valid transactions to the related address for executing any smart contract functions. Figure 29 shows code snippets from an example energy trading algorithm coded inside a smart contract.

The code snippets include minimal functions which must be realized in a smart contract code to ensure at least network registrations of participating consumer and prosumer nodes as well as instant execution of energy inject transactions issued by prosumer nodes.

Algorithm 1: Energy Trading Algorithm

procedure ENERGYTRADING
txAddr ← hash(public key), msgAddr ← hash(public key of whisper)
Prosumer_i ←. $\text{txAddr}P_i$, $\text{msgAddr}P_i$, O_{pi}, E_i, $S_i(t)$, S_{mi} $(1 \leq i \leq n_p)$
Consumer_j ← $\text{txAddr}C_j$, $\text{msgAddr}C_j$, O_{cj}, $D_j(t)$, D_{mj} $(1 \leq j \leq n_c)$
DSO ← txAddr$D$, msgAddr$D$, $ES(t)$, $ED(t)$
Smart Contract(SmC) ← smcAddr
procedure REGISTER($\text{Prosumer}_1, \ldots, \text{Prosumer}_{n_p}$, $\text{Consumer}_1, \ldots, \text{Consumer}_{n_c}$)
enum State{register, injected, board, match, purchased}
struct EnergyOwnership{
address account; **uint** amount; **State** state; **uint** timestamp;}
if txAddr ∈ txAddrP_i **then**
txAddrP_i.sendtx ($Register$, $timestamp$) ⇒ smcAddr
EnergyOwnership[ ] O_{pi};
SmC.event(O_{pi} ← (txAddrP_i, 0, $register$, $timestamp$))
else if txAddr ∈ txAddrC_j **then**
txAddrC_j.sendtx ($Register$, $timestamp$) ⇒ smcAddr
EnergyOwnership[ ] O_{cj};
SmC.event (O_{cj} ← txAddrC_j, 0, $register$, $timestamp$))
end if
end procedure
procedure INJECTENERGY(E_i)
msgAddrP_i.msg ($Inject$, txAddrP_i, E_i) ⇒ msgAddr$D$
txAddr$D$.sendtx ($Inject$, txAddrP_i, E_i, $timestamp$) ⇒ smcAddr
SmC.require (msg.sender == txAddr$D$)
SmC.event(O_{pi} ← (txAddrP_i, E_i, $injected$, $timestamp$))
end procedure
procedure AGGREGATION ($\text{Prosumer}_1, \ldots, \text{Prosumer}_{n_p}$, $\text{Consumer}_1, \ldots, \text{Consumer}_{n_c}$ during Δt)

Figure 29 Smart contract code snippets

6.1.3 Design of peer-to-peer software framework (demo)

The whole architecture of P2P software framework demo in depicted in Figure 30.

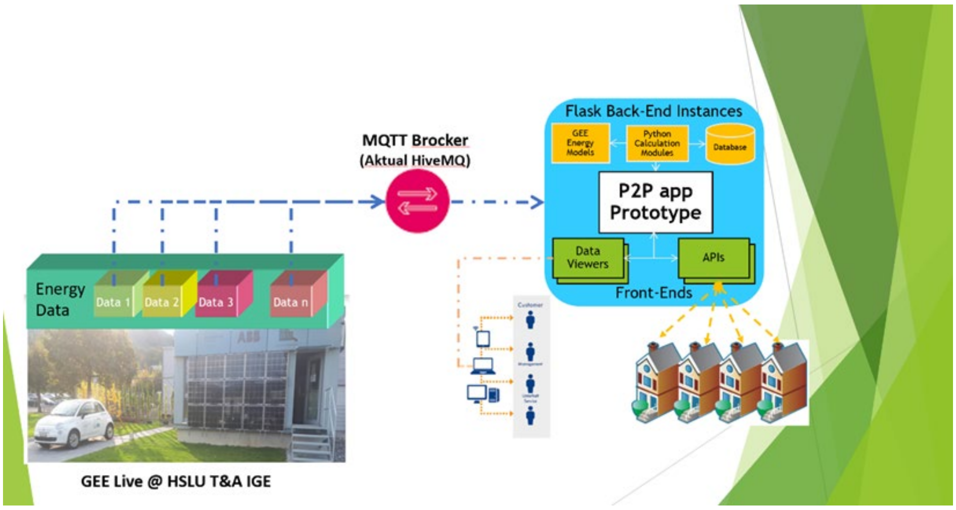

Figure 30 Architecture of P2P software framework (demo)

The P2P software framework consists of a front- and a backend system. On the first layer of the backend system (shown in Figure 30) several instances are running in background and are ensuring following functionalities:

– Subscribing to an external Message Queuing Telemetry Transport (MQTT) broker HiveMQ which is used for periodically publishing actual GEE energy data (productions and consumptions).
– Intercepting and pre-preprocessing GEE energy data acquired via an instance of paho-mqtt (Phypi, n.d.) client.
– Storing the pre-processed GEE energy data internally inside a mongoDB database.

On the second layer of the backend systems concurrent jobs are configured to run in deferent threads and to not disturb each other.

In total 4 threads are defined. The first thread runs a HTTP web server which publishes GEE data stored in MongoDB internally to the frontend system as well as externally towards other nodes in the P2P network. The second thread subscribes continually to the MQTT topics. The third thread intercepts MQTT message, parses, pre-processes these energy data messages, and stores the preprocessed GEE energy data to the internal database. The last one Thread is for running periodically simple EM services to assess GEE energy consumptions & productions.

The app layer consists of the web server, the EM services as well as additional Applications Programming Interfaces (APIs) enabling integration and testing of DLT/BLC-functionality inside the target P2P network.

The frontend of the system enables visualization of actual GEE energy data. Every energy production or consumption data entry stored in MongoDB is accessible via a unique identifier «id». The stored energy value can be retrieved using an API and displayed in related frontend data viewer.

All GEE stored energy (production and consumption) data can also be transmitted to any external instance (node) inside the P2P energy network using HTTP Requests. However, in our demo framework all external energy data exchanges between the nodes participating in the P2P network should be addressed using User Datagram Protocol (UDP) or Transport Control Protocol (TCP)-based sockets.

Figure 31 shows an overview of the software stacks used to build the P2P front -/ backend core packages. The frontend Apps are in majority programed to run in an Angular/Node (Angular, n.d.) and the backend modules in an integrated Python environment. The virtual nodes building the DLT/BLC-Test-Network must be composed to run in a Docker environment. The detailed description of the implementation can be found in Appendix A5.

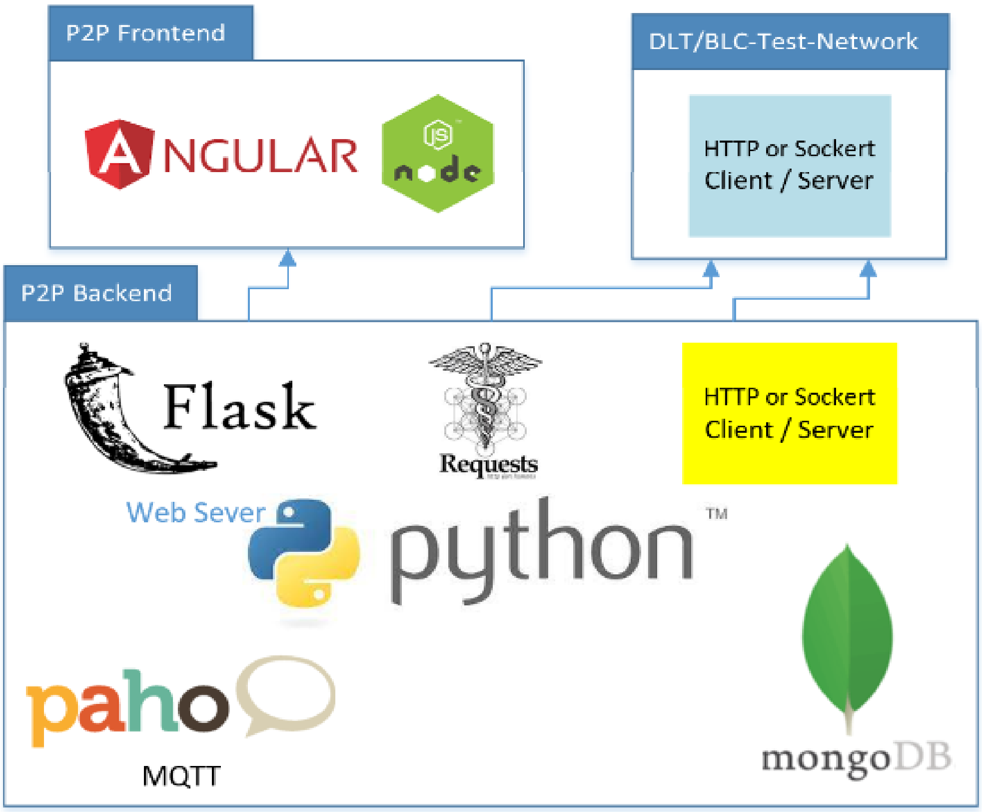

Figure 31 Software stack used for the front- and back-end of the P2P demo

6.2 Buildings energy data exchange in a data space topology for a peer-to-peer implementation

The development of a peer-to-peer (P2P) energy data exchange platform as part of the SINA project represents a significant step forward in the pursuit of sustainable and renewable energy solutions. The research conducted in this report has demonstrated the feasibility of creating a secure, standardized architecture for building data exchange in Switzerland. The three-stage approach, encompassing the establishment of bidirectional IP-based communication channels, the integration of smart contracts, and the simulation of energy data communication scenarios, has provided valuable insights into the practical implementation of such a system. The potential for integrating Distributed Ledger Technology (DLT) and Blockchain Technology (BCT) further enhances the security and reliability of the platform. As we move forward, the lessons learned from this research will be invaluable in guiding future work in this area, particularly in the expansion of the communication model and the exploration of additional use cases.

Enabling secure, trustful, and durable interoperability infrastructures for an ad-hoc exchange of specific buildings data behaviors and particularly buildings energy production and consumption data remains the long-term goal of the P2P work. The instant PoC is analyzing the feasibility and suitability of certain protocols and software stacks as basic framework for an initial P2P energy data exchange layer between buildings in a local energy network without deepening in special topics such as trust, data security and sovereignty, identity and access management (IAM), authentication (AuthN) and authorization (AuthZ), while the future R&D tasks may deepen in these special topics and take advantage of the mature components and constructs provided by data spaces.

In future P2P topology design the data accesses and information flows between the energy consumer, prosumer building nodes and any other organisms that are participating in a local energy network shall be standardized, simplified, and secured. These key requirements can be addressed in the future by adhering to data spaces approaches and deploying certified data spaces Connectors on the top all the nodes that are participating in the local energy network. One or more Connector instances should run in the cloud or on-premises IT environment where the building energy data is maintained. These instances should also be programmed to retrieve requested energy data from the backend energy database and write data back to the database.

Every Connector instance should be enabled to provide descriptions (metadata) describing their own data exchange pre-conditions, so that any other Connector instance running in a neighbor node in the network can get access to these self-descriptions and get information about the technical interface, authentication mechanisms, associated data usage policies, etc. To enable this kind of metadata, exchange a Metadata Broker instance should be integrated within the network.

Always before an exchange of energy data from one node Connector instance to another neighbor node Connector instance in same network take effectively place, both instances must know, trust, negotiate and go a data access agreement (smart contract) with each other. For validating and supervising this kind of (automatic) energy data access agreements a Clearing House instance should be provided within the network and managed by the network service provider. This instance is responsible for maintaining all information relevant for clearing, billing, data access and usage control.

A cost-effective and fast processing (interpretation) of the exchanged energy data can only be reached if the energy consumption and productions data series itself are well described and structured. The IDSA's information layer foresees for this a service called Vocabulary Hub. An instance of the Vocabulary Hub should be deployed and used by an energy data provider connector instance to define and publish the semantics and models associated to an energy data endpoint. This Vocabulary Hub instance can be queried by all nodes participants in the network to get the detailed information model behind a certain energy data record.

Figure 32 shows sample data interactions between two P2P nodes. The IT-environment of every node hosts an IDS's connector instance and a backend instance. Most significant interactions may take place between the consumer and prosumer connector instances. After querying and getting detailed descriptions (metadata) via the IDS's Vocabulary Hub instance about the energy surpluses of an authenticated prosumer node and if the quantity, price, etc. conditions are favorable, the connector instance of the consumer node may ask the related prosumer connector for an individual smart contract. Both connector instances negotiate then related contract and if an accord is reached, the final agreement will be signed by both parties and persisted in every connector instance. The IDS's Cleating House instance will be notified by the prosumer connector instance that an agreement is reached.

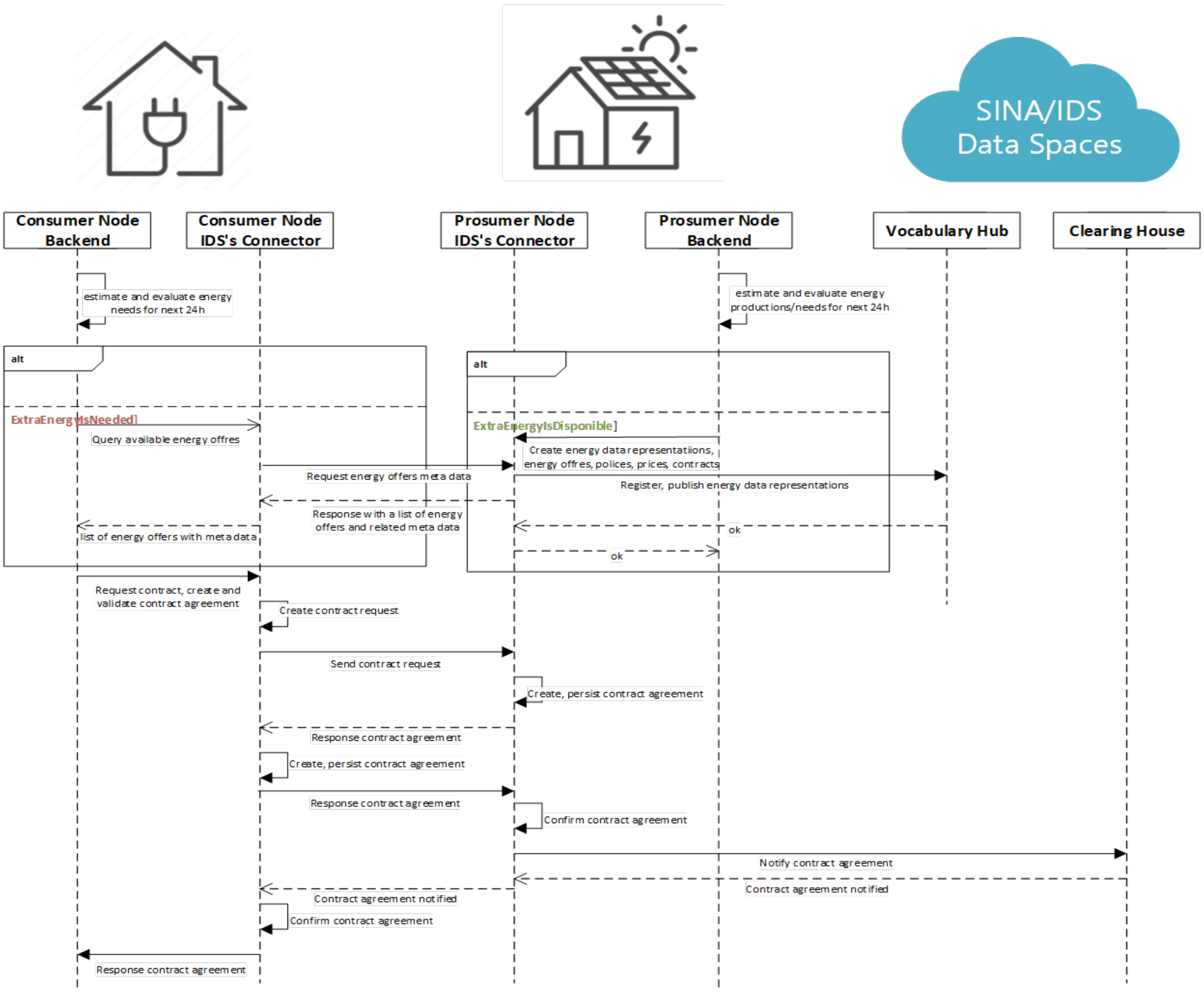

Figure 32 Sequence diagram of data interactions between P2P nodes

Once a smart contract is signed, the consumer and prosumer connector instances can start real data exchange. The Cleating House instance logs related exchanges and can provide in retrospect base information for any subsequent billing process.

Implementing secure communication frameworks that enable the sharing of specific (energy) data trends that may be relevant for other nodes participating in the same local P2P energy network is theoretically feasible. However, setting up and managing parallel DLT/BLC infrastructures for deploying smart contracts and for handling billing information may be very costly, challenging and prone to error. Deploying additional user-friendly Apps on top of an existing local P2P data sharing infrastructure may help to reduce complexity for the residents of the buildings and contribute to a better satisfaction by providing extra functionalities and access of their energy trends data.

7 Governance: organizational structure

The chapter proposes the establishment of an Association as an appropriate organizational structure to coin the SINA data space. It is important to highlight that given the similarities between SINA and the IDSA Project, we will implement the findings and best practices from the IDSA Project within the governance framework of SINA whenever applicable to ensure that SINA benefits from the knowledge and experience gained in the IDSA Project, enhancing its governance structure.

7.1 Legal structure

Central to the governance framework of SINA is a dedicated commitment to cultivating impactful inter-organizational collaboration. Acknowledging the vitality of shared resources and collective endeavors within data spaces, our governance model prioritizes the implementation of well-defined standardized procedures and inclusive decision-making processes. Consequently, the establishment of an ecosystem for data spaces becomes imperative, necessitating the selection of a legal structure aligned with these essential requirements.

An association is being proposed as an appropriate legal structure for SINA. The association's purpose cannot be profit-oriented in accordance with Article 60 of the Swiss Civil Code (Picht & Studen, 2022). SINA's objectives include promoting collaboration, ensuring data privacy, facilitating data analysis, and enabling the development of innovative applications. In essence, SINA aims to enhance data management within organizations, emphasizing that its primary purpose is not profit-driven. Since the activities and the project will last for a while and is not a short-term project, an association is more suitable than a simple partnership according to Art. 530 CO (Schweizerische Eidgenossenschaft, 1911). Another difference between a simple partnership and an association is that the association is a legal person, which means members can join and leave[24]. In SINA we want that new members can join, and the members should not be set on the point of time the legal structure is founded.

Appendix A6 describes in detail the proposed legal structure, prerequisites for founding the SINA association, the roles within the association as well as the membership criteria and the proposed governance model, while Appendix A7 provides the project's Draft of the Statutes for setting up SINA association («Vereinstatuten»).

7.2 Governance instruments

Within the SINA association, governance instruments play an important role in ensuring management and control. These instruments are designed to facilitate seamless operations, compliance, and communication, ultimately fostering a harmonious environment for data sharing. Below, are the key governance instruments outlined in this context (IDSA, 2021, pp. 16–17):

– Standardization: This instrument focuses on formalizing, documenting, and aligning tasks, processes, and guidelines related to data spaces and SINA. Standardization serves as a blueprint, providing a structured foundation for stakeholders to work from.

[24] Art. 70 para. 1 SCC

- Certification: Certification is the mechanism through which stakeholders are validated in their adherence to standardized procedures. It can be further categorized into component certification, which deals with the certification of technical and software components, and organizational certification, which focuses on certifying organizational and legal processes.
- Development: In certain instances, the activity within data spaces may necessitate software development tasks as part of its realization. These development tasks should conform to the established standardization activities and ideally culminate in certification.
- Operations: Operations are at the core of the exploitation and utilization of developed components. As part of the organizational certification, these operations can be certified to ensure compliance with established governance practices.
- Communication: Effective dissemination of information is a vital component, encompassing not only raising awareness about standardization and certification but also incorporating marketing aspects to foster understanding and support.
- Support: Support activities are designed to provide structured assistance to stakeholders engaged in various aspects of the operation, development, certification, and communication activities, ensuring a cohesive and coordinated approach.
- Audits: Serve as an important governance instrument in the SINA association by ensuring data accuracy, compliance with regulations, and the implementation of robust security measures. They help manage and mitigate risks associated with data handling, providing transparency and accountability in operations. Audits also contribute to continuous improvement by identifying areas for enhancement within the governance framework, fostering member confidence and promoting effective collaboration in the shared data environment. Overall, audits play a pivotal role in maintaining high standards of data management and governance within the association (Hess & Ostrom, 2007).

To enhance the effectiveness of the governance concept of data spaces, it is crucial for the executive board of the SINA Association to carefully assess and incorporate the aforementioned instruments into their decision-making processes and strategies.

7.3 Why is an association needed for SINA

In data spaces, the significance of associations is of importance, even in an environment characterized by its lack of a central governing authority. While the concept of decentralization implies a distributed and autonomous system, the presence of a singular association remains crucial, particularly when considering the importance of trust (IDSA, 2021, p. 4).

At the heart of this necessity is the fundamental challenge of establishing and maintaining trust in a decentralized framework. In the absence of a central authority dictating the rules, participants in decentralized data spaces require a common ground to build trust. An association, acting as a unifying force, plays a pivotal role in this context by setting standards, defining protocols, and creating a shared framework that guides the behavior of all involved parties.

Trust is the foundation that facilitates the seamless flow of information within decentralized networks. Associations serve as custodians of this trust by creating and upholding agreed-upon standards. These standards not only ensure the ethical conduct of data transactions but also provide a sense of security and confidence for participants engaged in the exchange.

While the emphasis here is on a singular association, its role is expansive and multifaceted. Beyond governance, the association becomes a symbol of unity within the decentralized data space (IDSA, 2021, p. 15), fostering a collective identity among participants. This shared identity contributes to a culture of transparency and accountability, reinforcing the trust that is essential for the sustainability and success of the decentralized data ecosystem.

The existence of an association remains important in data spaces, primarily due to its role in cultivating and preserving trust. Serving as a unifying force, establishing standards, and mediating disputes, this association becomes the core upon which a reliable, secure, and collaborative decentralized data environment can thrive.

Despite the existence of data spaces, the establishment of a supportive ecosystem is crucial. This is where the role of an association becomes pivotal. The creation of the SINA association not only facilitates the implementation of crucial trust-building instruments but also simplifies the process. The association is tasked with governing the data spaces as far as possible within the decentralized environment. By establishing the association, we lay the foundation for an ecosystem to thrive.

As we embark on founding the SINA association, questions arise about membership criteria and the composition of the executive board, considering various factors, including the perspectives of intermediaries and other important institutions. It's important to note that these criteria are not set in stone and may evolve as needs change. The flexibility of the SINA association is important; it should be dynamic and capable of adapting to the growth and continuous development of data spaces.

8 Conclusion

The SINA project addressed the possibilities of data spaces to include various actors in the building industry and explored use cases for value-adding implementations. Five distinct conceptual data space aspects were investigated: technological, ontological, governance, economic and ecological. The fundamental insight here is that technology is only one part of a data space. From this perspective, data spaces can be described as a means of data integration. Data that is available at a certain location and is required elsewhere does not have to be transferred completely and permanently. Rather, it is sufficient to integrate distributed data to the extent and for the duration that is necessary.

Against this background, the following could be demonstrated with SINA:

– With IDSA or Gaia-X, numerous projects and players within the EU are working intensively on conceiving and realizing diverse data spaces.

– In addition to the technical IDS reference architecture, basic principles in the areas of governance, ontology and use cases are also available. However, these are still at an early stage, meaning that substantial development work is still required for the realization of a use case - such as P2P energy trading - in terms of both technical and interoperability aspects.

– The industry partners involved in this project are showing great interest in further projects to develop a data economy in the Swiss energy sector and building industry based on the use cases discussed. The crucial question will therefore be: «how a framework can be created and designed, that sets the right incentives for the actors involved to make data exchange beyond today's minimal levels possible and attractive?»

From the technical perspective, the following primary objectives have been identified: a) avoidance of redundant data storage; b) enabling data owners to maintain autonomous control over access permissions to their data; c) ensuring the verifiable authenticity of participant identities; and d) the decentralization of data management. While the project has identified challenges in interfacing and certificate management, the developed codebase lays a foundation for future advancements, potentially leading to a library for companies investing in data space technologies. Emphasizing design in ontologies and structured vocabularies within the data space framework enhances data organization and interoperability, supporting adaptability and compatibility.

When looking at the ontologies, it is noticeable that despite advancements, challenges persist, particularly in the underutilization of ontologies and the scarcity of programmatic representations and an absence of widespread adoption that would warrant systematic endeavors to fortify their utilization and support. Addressing these issues necessitates concerted efforts through awareness campaigns, educational initiatives, and collaborative endeavors to establish standardized practices, enhancing the data space's effectiveness.

From a legal standpoint, the requirement for interoperability underscores the multifaceted nature of challenges in data spaces, necessitating a different approach. The inherent difficulties highlight the importance of addressing these problems when crafting contracts and designing the legal framework for data spaces such as SINA. Recognizing that achieving legal interoperability depend on the dynamic and evolving nature of governance and legal aspects, the data space framework needs to be capable to respond to the evolving legal landscape, ensuring that it remains pertinent and resilient.

From an economic and ecological perspective, the positive reception of data spaces among stakeholders is evident. Scalability and cost-effectiveness are tied to the diversity and extent of accessible data sources, presenting potential economic benefits and cost reductions. Furthermore, it could be indicated that a widespread implementation of data spaces could also lead to environmental benefits by a more widespread and faster adoption of use cases leading to increasing energy efficiency and energy savings. During the assessment of various use cases in the ecosystems «Building Information», «Smart Home and Comfort», and «Energy and Grid», our industry partners expressed an interest in further exploring the potential of data spaces in real-world applications and are eager to collaborate on pilot projects to understand how this concept can be implemented. The examination of specific use cases underscores the crucial role of the economic perspective in shaping data spaces.

Data spaces should be perceived as a federated, open infrastructure facilitating sovereign data exchange, grounded in shared agreements, rules, and standards. This approach is essential for fostering a novel data economy marked by entirely innovative and at times disruptive business models. Within the analyzed use cases, two solutions consistently emerged:

- Implementation without data space technology, situated closely to existing processes. However, this approach often results in silo solutions and generally overlooks the potential offered by the data economy.
- Implementation involving data space technology, with active engagement from diverse stakeholders and the exploration of new (disruptive) business models within the framework of the emerging data economy.

To fully harness the advantages of data spaces, effective communication on their benefits compared to alternative technologies is crucial. Strengthening aspects such as streamlined legal agreements, robust data ownership frameworks, and heightened security is vital for informed adoption. Overcoming the tendency towards isolated solutions is key to maximizing the potential impact of data spaces, promoting a holistic and interconnected mindset in implementing data solutions. In the energy sector and building industry, awareness of challenges exists, but a cautious approach prevails. Overcoming this requires building confidence through pilot projects and collaborative endeavors to showcase the benefits of data spaces. This proactive approach can contribute to a more receptive environment for integrating data spaces in the energy sector, fostering sustainable advancements. The development of a P2P energy data exchange platform within SINA demonstrated the feasibility of secure bidirectional IP-based communication channels and sample smart contracts for simulating energy sharing scenarios. However, challenges persist in using distributed ledger and blockchain technologies in such scenarios, highlighting the need for further investigations and advancements. Utilizing a data space as a basis framework for P2P network topology may formalize, standardize, and secure interfaces for energy exchange, but additional evaluations are required for real-time capabilities such as comfort data.

In summary, the SINA project has successfully showcased the adeptness of the IDSA software in establishing a secure data space that respects data sovereignty. The open-source nature of the software, supported by dedicated project groups, ensures ongoing maintenance and development for different components of the data space. While data spaces operate in a decentralized manner, establishing a central entity, such as a SINA Association, is crucial for various reasons, especially for establishing trust but also setting basic guidelines and principles in the data ecosystem. Ultimately, this would also meet the recommendations of the Swiss Data Alliance (2023) which urges Switzerland to get involved in the topic of data spaces.

9 Outlook and next steps

Building a strong data exchange ecosystem goes beyond regulatory mandates and requires the establishment of favorable conditions. The establishment of a digital infrastructure, ensuring data sovereignty, security, and interoperability, is crucial for encouraging market players to engage in data exchange. In order to fully unlock the immense potential of data spaces, it is crucial to explore the dynamic realm of upcoming challenges. These challenges include matters such as legal agreements, data ownership, data sovereignty, trust, and security, all of which are already under consideration in the various data spaces initiatives.

Notably, the European Union strongly supports initiatives and research on data spaces, recognizing their pivotal role in shaping the future of data exchange. According to Gartner (2021), companies that actively participate in data sharing and analytics generate three times more measurable economic benefit than those that do not, underscoring the economic advantages of a collaborative approach.

Nevertheless, the true value of data spaces lies in their ability to enable a trustworthy solution of data exchange between a multitude of different stakeholders. To maximize this impact, concerted efforts are required to incorporate a diverse array of organizations and companies. This inclusive approach is essential for creating a dynamic and thriving data economy.

Furthermore, special attention needs to be directed towards supporting Swiss companies interested in participating in data space pilot projects. By encouraging collaboration and inclusivity, publicly funded initiatives in the data space domain such as SINA and the European IDSA or GAIA-X can foster a cross-sector data exchange that unlocks the potential benefits of data spaces on a broader scale, contributing to economic growth and innovation. In particular, small and medium-sized enterprises require tools to finance innovation projects related to data spaces and to mitigate the risk of failure.

A pivotal aspect of our forward trajectory involves the establishment of the SINA association, a foundational entity dedicated to facilitating the realization of our strategic vision to build up a data space community and further grow the association upon existing international frameworks. Concurrently, the options are analyzed for follow-up projects that extend beyond theoretical exploration and proof-of-concept implementation, aiming to rigorously test and implement the data space framework in collaboration with industrial partners. These initiatives form the nucleus of our commitment to fostering innovation and integration in a future data economy.

References

AKW Group, & E-Bridge. (2021). *Datahub Schweiz. Kern zukünftiger Dateninfrastruktur digitaliserter Strom- und Gasmärkte: Ausgestaltung eines Datahub, der Prozesse, betroffener Daten und Datenschutzfolgeabschätzung*. https://www.bfe.admin.ch/bfe/en/home/supply/digitalization.html

Angular. (n.d.). *Setting up the local environment and workspace*. Retrieved December 18, 2023, from https://angular.io/guide/setup-local

Berkhout, V., Frey, C., Hertweck, P., Nestle, D., & Wickert, M. (2022). Energy Data Space. In B. Otto, M. ten Hompel, & S. Wrobel (Eds.), *Designing Data Spaces : The Ecosystem Approach to Competitive Advantage* (pp. 329–341). Springer International Publishing. https://doi.org/10.1007/978-3-030-93975-5_20

Beverungen, D., Hess, T., Köster, A., & Lehrer, C. (2022). From private digital platforms to public data spaces: implications for the digital transformation. *Electronic Markets*, *32*(2), 493–501. https://doi.org/10.1007/s12525-022-00553-z

Bundesamt für Energie BFE. (2023). *Überblick über den Energieverbrauch der Schweiz im Jahr 2022*. https://www.bfe.admin.ch/bfe/de/home/versorgung/statistik-und-geodaten/energiestatistiken/gesamtenergiestatistik.html

Bundesamt für Kommunikation. (2023). *EU-Digitalpolitik: Analyse der Auswirkungen auf die Schweiz*. https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-94291.html

Bundeskartellamt. (2019). *Case Summary : 15 February 2019 Facebook, Exploitative business terms pursuant to Section 19(1) GWB for inadequate data processing* .

Curry, E., Tuikka, T., Metzger, A., Zillner, S., Bertels, N., Ducuing, C., Dalle Carbonare, D., Gusmeroli, S., Scerri, S., López de Vallejo, I., & García Robles, A. (2022). Data Sharing Spaces: The BDVA Perspective. *Designing Data Spaces*, 365–382. https://doi.org/10.1007/978-3-030-93975-5_22

Daniele, L., den Hartog, F., & Roes, J. (2015). Created in Close Interaction with the Industry: The Smart Appliances REFerence (SAREF) Ontology. *Lecture Notes in Business Information Processing*, *225*, 100–112. https://doi.org/10.1007/978-3-319-21545-7_9/COVER

DETEC, & FDFA. (2022). *Creating trustworthy data spaces based on digital self-determination*. https://www.admin.ch/gov/en/start/documentation/media-releases.msg-id-87780.html

Duisberg, A. (2022). Legal Aspects of IDS: Data Sovereignty—What Does It Imply? *Designing Data Spaces*, 61–90. https://doi.org/10.1007/978-3-030-93975-5_5

Ethereum. (n.d.). *What is Ethereum?* Retrieved December 18, 2023, from https://ethereum.org/en/what-is-ethereum/

European Commission. (2019). *Building a data economy — Shaping Europe's digital future*. https://digital-strategy.ec.europa.eu/en/library/building-data-economy-brochure

European Commission. (2022). *Regulation of the European Parliament and of the Councill on harmonised rules on fair access to and use of data (Data Act)*.

European Parliament. (2016). *Directive (EU) 2016/ 943 Of The European Parliament And Of The Council - of 8 June 2016 - on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure*. Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure

European Parliament. (2018). *Regulation (EU) 2018/ 1807 Of The European Parliament And Of The Council - of 14 November 2018 - on a framework for the free flow of non-personal data in the European Union*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32018R1807

European Parliament. (2022, May 30). *European data governance*. https://eur-lex.europa.eu/EN/legal-content/summary/european-data-governance.html

Eurostat Data Browser. (2023). *Number of households by household composition, number of children and working status within households*. https://ec.europa.eu/eurostat/data-browser/view/lfst_hhnhwhtc/default/table?lang=en

Federal Statistical Office. (2023). *Private households by canton and household size*. https://www.bfs.admin.ch/bfs/en/home/statistics/population/effectif-change.html

Fraunhofer IAIS. (2022). *IDS Vocabulary Provider and It's purpose*. https://github.com/International-Data-Spaces-Association/IDS-VocabularyProvider/tree/main

Fraunhofer ISST. (2021). *Messages | Dataspace Connector*. https://international-data-spaces-association.github.io/DataspaceConnector/Documentation/v6/Messages

Gartner. (2021). *Data Sharing is a Key Digital Transformation Capability*. https://www.gartner.com/smarterwithgartner/data-sharing-is-a-business-necessity-to-accelerate-digital-business

Gottschalk, M., Uslar, M., & Delfs, C. (2017). The Use Case and Smart Grid Architecture Model Approach. *Springer Briefs in Energy*. https://doi.org/10.1007/978-3-319-49229-2

Gualandri, F., & Kuzior, A. (2023). Home Energy Management Systems Adoption Scenarios: The Case of Italy. *Energies 2023, Vol. 16, Page 4946*, *16*(13), 4946. https://doi.org/10.3390/EN16134946

Hess, C., & Ostrom, E. (2007). *Understanding knowledge as a commons : from theory to practice*. 367.

Hyperledger. (2023a). *A Blockchain Platform for the Enterprise — hyperledger-fabricdocs main documentation*. https://hyperledger-fabric.readthedocs.io/en/release-2.5/

Hyperledger. (2023b). *Using the Fabric test network — hyperledger-fabricdocs main documentation*. https://hyperledger-fabric.readthedocs.io/en/release-2.5/test_network.html#creating-a-channel

Hyperledger. (2023c, March 31). *Fabric chaincode lifecycle — hyperledger-fabricdocs main documentation*. https://hyperledger-fabric.readthedocs.io/en/latest/chaincode_lifecycle.html

IBM. (n.d.). *What is hyperledger fabric?* Retrieved December 18, 2023, from https://www.ibm.com/topics/hyperledger

IDSA. (2021). *Governance for Data Space Instances Aspects and Roles for the IDS Stakeholders*. www.internationaldataspaces.org

IDSA. (2022a). *Information Layer - IDS Knowledge Base*. https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/layers-of-the-reference-architecture-model/3-layers-of-the-reference-architecture-model/3_3_informationlayer

IDSA. (2022b). *MVDS-Testbed_1.0*. https://github.com/International-Data-Spaces-Association/IDS-testbed/blob/master/pictures/MVDS-Testbed_1.0.png

IDSA. (2023). *IDSA Tech Talk | Legal Interoperability – legal aspects in data spaces - YouTube*. https://www.youtube.com/watch?v=D7RCfVE8aAo

IEC. (2018). *IEC 62325-301:2018 | IEC Webstore*. https://webstore.iec.ch/publication/31487

IRENA. (2020). *Peer-to-peer electricity trading - Innovation Landscape Brief*. www.irena.org

Jacquette, D. (2002). *Ontology*.

Karlin, B., Ford, R., Sanguinetti, A., Squiers, C., Gannon, J., Rajukumar, M., & Donnelly, K. A. (2015). *Characterization and Potential of Home Energy Management (HEM) Technology*. https://escholarship.org/uc/item/6qd1x5js

Knüsel, L., & Richard, P. (2022). *Die Datenökonomie in der Energiewirtschaft*. www.dena.de

Kommission von der Leyen. (2023). *Die Schweiz und die Digitalstrategie der Europäischen Union*.

Mader, C., Pullmann, J., Petersen, N., Lohmann, S., & Lange-Bever, C. (2022, September 16). *International Data Spaces Information Model*. https://international-data-spaces-association.github.io/InformationModel/docs/index.html

Maghazei, O., Lewis, M. A., & Netland, T. H. (2022). Emerging technologies and the use case: A multi-year study of drone adoption. *Journal of Operations Management*, *68*(6–7), 560–591. https://doi.org/10.1002/JOOM.1196

Marques, J., Soares, T., & Morais, H. (2023). P2P flexibility markets models to support the coordination between the transmission system operators and distribution system operators. *Sustainable Energy, Grids and Networks*, *34*, 101055. https://doi.org/10.1016/J.SEGAN.2023.101055

Misurio. (2020). *SHED Swiss Hub for Energy Data: Teil einer nationalen Energie-Dateninfrastruktur und Wegbereiter für Energiestrategie 2050, Klimaschutz und digitale Innovationen*. https://www.bfe.admin.ch/bfe/en/home/supply/digitalization.html

Morley, J., Widdicks, K., & Hazas, M. (2018). Digitalisation, energy and data demand: The impact of Internet traffic on overall and peak electricity consumption. *Energy Research & Social Science*, *38*, 128–137. https://doi.org/10.1016/J.ERSS.2018.01.018

Motlagh, N. H., Mohammadrezaei, M., Hunt, J., & Zakeri, B. (2020). Internet of things (IoT) and the energy sector. In *Energies* (Vol. 13, Issue 2). MDPI AG. https://doi.org/10.3390/en13020494

Neugebauer, R. (2019). *Digital Transformation*.

Ochs, C., Friedewald, M., Hess, T., & Lamla, J. (2019). *Die Zukunft der Datenökonomie - Zwischen Geschäftsmodell, Kollektivgut und Verbraucherschutz* (C. Ochs, M. Friedewald, T. Hess, & J. Lamla, Eds.). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-27511-2 openHAB. (n.d.). *Semantic Model | openHAB*. Retrieved January 9, 2024, from https://www.openhab.org/docs/tutorial/model.html

Otto, B. (2022). The Evolution of Data Spaces. *Designing Data Spaces*, 3–15. https://doi.org/10.1007/978-3-030-93975-5_1

Our World in Data. (2023). *Switzerland: Carbon intensity: how much carbon does it emit per unit of energy?* https://ourworldindata.org/energy/country/switzerland#carbon-intensity-how-much-carbon-does-it-emit-per-unit-of-energy

Pettenpohl, H., Spiekermann, M., & Both, J. R. (2022). International Data Spaces in a Nutshell. *Designing Data Spaces*, 29–40. https://doi.org/10.1007/978-3-030-93975-5_3

Phypi. (n.d.). *paho-mqtt*. Retrieved December 18, 2023, from https://pypi.org/project/paho-mqtt/

Picht, P. G., & Studen, G. (2022). Civil Law Principles and Family Law. In M. Thommen (Ed.), *Introduction to Swiss Law* (pp. 229–252). sui generis Verlag. https://doi.org/10.38107/026

Pöyry Norway AS. (2019). *Benefit assessment of an EU-wide data exchange platform*. https://elering.ee/sites/default/files/public/tarkvork-konv/tarkv%C3%B5rk/P%C3%B6yry%20-%20Benefit%20Assessment%20of%20an%20EU-wide%20Data%20Exchange%20Platform.pdf

Roland Berger. (2018). *Plattformökonomie im Maschinenbau*. https://www.rolandberger.com/de/Insights/Publications/Plattform%C3%B6konomie-im-Maschinenbau.html

Rusitschka, S., & Curry, E. (2016). Big Data in the Energy and Transport Sectors. In J. M. Cavanillas, E. Curry, & W. Wahlster (Eds.), *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe* (pp. 225–244). Springer International Publishing. https://doi.org/10.1007/978-3-319-21569-3_13

Schneider, G., & Winters, J. P. (2001). *Applying use cases: a practical guide*. Pearson Education.

Schweizerische Eidgenossenschaft. (1911, March 30). *CO | SR 220 - Federal Act of 30 March 1911 on th... | Fedlex*. https://www.fedlex.admin.ch/eli/cc/27/317_321_377/en?print=true

Šestak, M., & Copot, D. (2023). Towards Trusted Data Sharing and Exchange in Agro-Food Supply Chains: Design Principles for Agricultural Data Spaces. *Sustainability (Switzerland)*, *15*(18). https://doi.org/10.3390/su151813746

Statista. (2023). *Number of households in the United Kingdom from 1996 to 2022*. https://www.statista.com/statistics/295545/number-of-households-in-the-uk/

Steinbuss, S., Heitmann, B., Krempel, E., Ochs, M., & Schutzeichel, D. (2019). *IDSA Position Paper GDPR Related Requirements and Recommendations for the IDS Reference Architecture Model*. https://doi.org/10.5281/zenodo.5675903

Sury, U. (2023, September 25). *EU-Digitalstrategie – eine Übersicht*. https://jusletter.weblaw.ch/juslissues/2023/1168/-eu-digitalstrategie_57b4d6c06b.html__ONCE&login=false

Swiss Data Alliance. (2023). *Der europäische Datenraum aus Schweizer Sicht*. https://www.swissdataalliance.ch/publikationen/whitepaper-eu-datenraum

Thema Consulting Group, & devoteam. (2018). *Datahub Schweiz: Kosten-Nutzen-Analyse und regulatorischer Handlungsbedarf*. https://www.bfe.admin.ch/bfe/en/home/supply/digitalization.html

Tiko. (n.d.). *Solutions*. Retrieved December 6, 2023, from https://tiko.energy/solutions/

Trzaska, R., Sulich, A., Organa, M., Niemczyk, J., & Jasiński, B. (2021). Digitalization business strategies in energy sector: Solving problems with uncertainty under industry 4.0 conditions. *Energies*, *14*(23). https://doi.org/10.3390/en14237997

Tuomela, S., de Castro Tomé, M., Iivari, N., & Svento, R. (2021). Impacts of home energy management systems on electricity consumption. *Applied Energy*, *299*, 117310. https://doi.org/10.1016/J.APENERGY.2021.117310

Wessels, N., Gerlach, J., & Wagner, A. (2019). *Willingness-to-Sell Personal Information on Data-Selling Platforms*. *34*. https://aisel.aisnet.org/icis2019/cyber_security_privacy_ethics_IS/cyber_security_privacy/34

Appendices

A1 Reference architecture model

Business layer

In the business layer, the RAM defines the roles of the participants in the data space and categorizes them. It also specifies basic patterns of interaction between the roles and thus contributes to the development of innovative business models by the participants providing digital, data-driven services.

The business layer summarizes the roles in four categories which are explained below.

Core participants

The Core Participants comprise the roles of Data Supplier and Data Customer. The Data Supplier role is divided into Data Creator, Data Owner, and Data Provider. These roles provide data in the data space. The Data Customer role comprises the Data Consumer, Service Consumer and Data User. These roles receive, process and use the data for services.

Intermediary

Intermediaries act as trustworthy institutions and are usually referred to as «platforms». They play a central role and mediate between the large number of data suppliers and data customers. The Intermediary role is divided into Data Intermediary, Services Intermediary, App Store, Vocabulary Intermediary, Clearing House and Identity Authority. These intermediary roles may only be assumed by trustworthy organizations.

The data intermediary is responsible for ensuring that data is exchanged correctly between the data provider and the data consumer. This role primarily provides metadata via a data broker (metadata broker) and thus ensures that data consumers know which data a data provider provides and how.

Service Intermediary is a platform operator that provides metadata on the services offered.

App Store distributes data applications which are running in the context of a connector. In the context of data spaces, apps are not meant as software running on a smartphone. The apps are available for download in the App Store for data space.

Vocabulary Intermediary manages and offers vocabularies such as ontologies, reference data models or metadata elements on a technical level. Vocabulary annotates and describes the data sets.

The Clearing House records all activities within the scope of the data exchange without seeing the data itself. These activities are the confirmations that data has been sent by the data supplier and received by the data customer. Successful transactions can also be settled based on this information. In the event of conflicts, this information can serve as proof of dispatch and/or receipt.

Identity Authority ensures secure operation of the data space by creating, maintaining, managing, monitoring, and validating identity information for the participants in the data space. The Identity Authority consists of a Certification Authority, a Dynamic Attribute Provisioning Service, and a Dynamic Trust Monitoring.

Software Developer

This role includes IT companies that make the software available to the participants of a data space. In the context of data spaces, a distinction can be made between app developers and connector developers.

Governance Body

Governance bodies in the data spaces have the authority and the task of implementing and enforcing guidelines for standardization. They are responsible for building and maintaining trust within the data space. And they ensure the sustainable operation of the data spaces. This role is assumed by Certification Bodies and Evaluation Facilities, Standardization Organizations, and the International Data Spaces Association.

An overview of the interactions between the roles described is shown in Figure 33. Some of the roles (Certification Body, Evaluation Facilities) are not involved in day-to-day activities and are therefore not shown in the schema. Software Providers are linked to all other roles and are therefore not shown, which also applies to the Identity Provider. The diagram also only shows the basic interactions. Additionally, specific interactions are required for data exchange, which are described in the process layer.

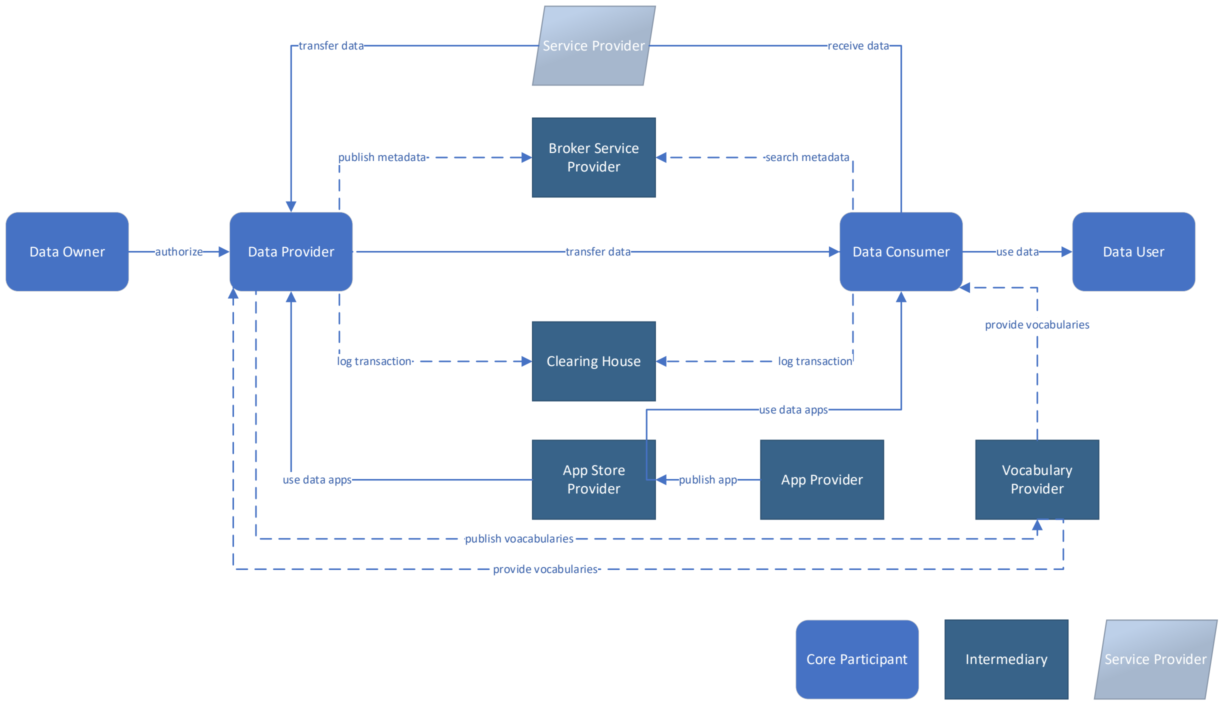

Figure 33 Basic interactions for data exchange and data sharing in the international data spaces

Functional layer

The functional layer specifies the functional requirements for data spaces and the resulting functionalities to be implemented. The functional architecture is divided into six areas, each of which comprises a group of functionalities of the software as shown in Figure 34.

Figure 34 Architecture of the functional layer

Trust

Functional requirements must be met in order to achieve trust. The software must be able to manage roles and identities and carry out user certification. The functionality that manages the roles must consider the respective rights of the roles and implement them. Identity management assigns a unique ID to each connector and issues the certificates. Each connector must also be able to verify the identity of the other connectors. User certification ensures that all participants in the data space undergo certification to create trust in the data space.

Security and Data Sovereignty

The functionalities regarding security and data sovereignty are fundamental in International Data Spaces.

– *Authentication and Authorization*: Connectors must have a valid X.509 certificate (or equivalent) to enable authentication and authorization.

– *Usage Policies and Usage Enforcement*: In International Data Spaces, data owners and data providers can be sure that data consumers' data is always handled in accordance with the applicable usage policies. Such policies can be attached to all data packages and thus restrict the use or forwarding of the data.

– *Trustworthy Communication and Security by Design*: The components in the data space can check whether the other party's connector has a trustworthy software stack. Data is only exchanged if all components involved meet the necessary security requirements and can prove this. If a component does not meet the requirements, no data is sent. Data providers and data consumers can agree on a specific security level and roll out the corresponding connectors.

– *Technical Certification*: Core components of the International Data Spaces, particularly the connectors, must be certified by the Certification Body so that trust can be established between all participants in the data space.

Ecosystem of Data

For data to be described, found, and correctly interpreted in a data space, each data source must be specified at the information level. The Ecosystem of Data consists of three areas: Data Source Description, Meta Data Brokering and Vocabularies.

– *Data Source Description*: Participants in the data space must be able to describe, publish, maintain, and manage different versions of metadata. The metadata describes the syntax and semantics of the data sources. It should also contain information about the domain of the application. In addition, the operator of the connector must be able to add prices, pricing models and usage policies.

– *Meta Data Brokering*: Operators of connectors must be able to provide data and metadata. It must be possible to transfer the metadata-to-metadata brokers, in whose repositories the participants can search the metadata and thus find suitable connectors.

– *Vocabularies*: To create and structure metadata, the data provider can use vocabularies to define the semantics of the data. Vocabularies can be registered in vocabulary hubs, where they can be queried by the participants. Data consumers thus have the opportunity to understand the semantics of a data provider's data. A data space should agree on a set of harmonized vocabularies so that there is a common understanding of the data available in the data space.

Standardized Interoperability

The standardized exchange of data between participants is the fundamental functionality in a data space. Data is exchanged via connectors. The connector is the main technical component in the data space.

– *Operation*: Connectors should run in the IT environment of the connector operator. Alternatively, they can also be installed on mobile devices or in embedded systems. The operators must define the workflow in the connector, and the users of the connectors must be identifiable. All actions in the connector must be recorded.

– *Data Exchange*: The connector retrieves the data from the company system using pull or push mechanisms. It can also write the data back again.

Value Adding Apps

Depending on the application, it may be necessary to process, analyze or transform the exchanged data. Data apps are used for this purpose. Such apps are available in the App Store for all participants.

– *Data Processing and Transformation*: Data apps provide clearly defined functionality. They transform a given input into a defined output format.

– *Data App Implementation*: The developers of data apps provide them with metadata about interfaces, functionality, price models or licenses.

– *Providing Data Apps*: Authorized developers can publish data apps in the App Store after they have gone through an optional evaluation and certification process. This process is controlled by the Certification Body. Authorized users can easily find suitable data apps in the App Store.
– *Installing and Supporting Data Apps*: Connectors allow authorized users to install and uninstall data apps that come from an official app store or alternatively from any other source.

Data Markets

The data exchanged in the data space can have a value, which is why a data space data market must include concepts for clearing, billing and governance.

– *Clearing and Billing*: The data owner determines the pricing model and the price of their data. The clearing and billing process must take this information into account and charge the correct prices.
– *Usage restrictions and governance*: The five aspects of governance in data spaces are data as an economic good, data ownership, data sovereignty, data quality and data provenance.
– *Legal aspects*: Legally valid contracts that can be concluded automatically are required for data exchange. Standardized contracts are required for typical data transactions.

Information layer

The information layer specifies the information model, the domain-independent, common language, i.e. the vocabulary used in the data space. The information model is an essential agreement that is shared by the participants and components of the IDS and facilitates compatibility and interoperability. The main purpose of this formal model is to enable the (semi-)automatic exchange of digital resources within a trusted ecosystem of distributed parties while preserving the data sovereignty of the data owners. The information model supports the description, publication and identification of data products and reusable data processing software (both, hereafter, referred to as digital resources or simply resources). Once the corresponding resources are identified, they can be exchanged and consumed via easily discoverable services. Apart from these core assets, the Information Model describes key elements of the International Data Space, its participants, its infrastructure components, and its processes.

The information model is generic and therefore not linked to any specific domain; domain modeling is a matter of vocabularies and specific data schemas. These provide domain-specific communities of International Data Spaces. The information model also does not provide a metamodel for defining customer-specific data types. Issues that go beyond the modeling of digital resources and their exchange are not considered by the information model. It also does not deal with side effects of data exchange, which can occur in scenarios with time-critical machine operations, for example.

Model representations

The reference architecture model specifies the information model on three formalization levels, which range from the conceptual document at a high level of abstraction down to the operational code (Figure 35).

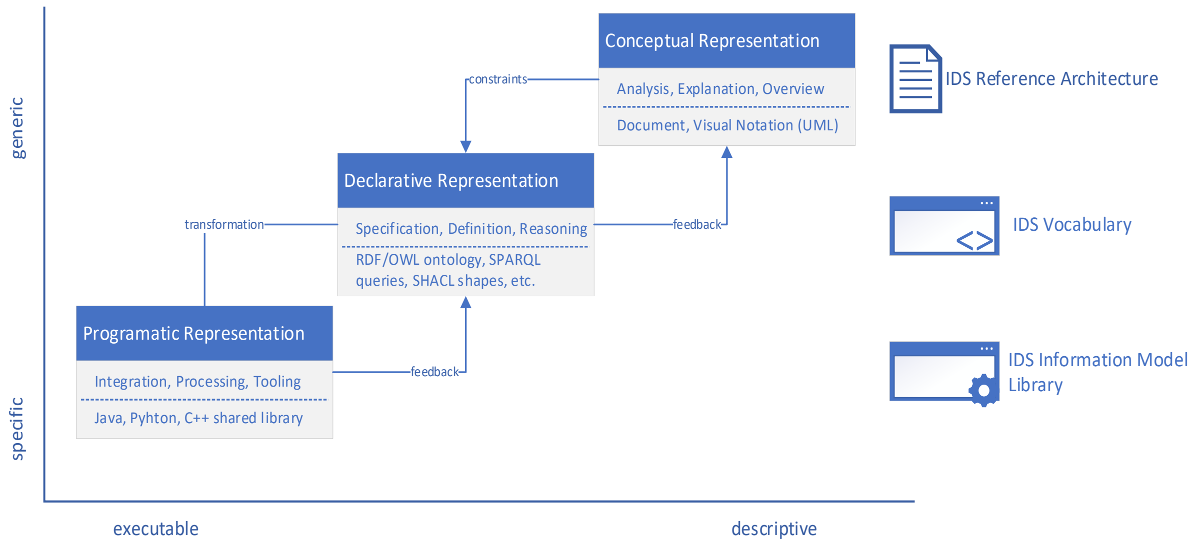

Figure 35 Representations of the Information Model

Each level represents the entire information model in its own way:

– *Conceptual presentation:* The conceptual representation of the information model provides an overview of the most important, largely unchanging concepts without committing to a specific technology or domain. It provides basic information and promotes a common understanding of the concepts. Where available, reference is made to related elements of the declarative representation and a programmatic representation.

– *Declarative representation:* The declarative representation forms the IDS vocabulary and provides a normative view of the International Data Spaces information model. The information model provides a formal, machine-interpretable specification of the concepts and entities of the International Data Spaces. Data resources or software resources are queried using standardized query languages. The declarative representation comprises a complete reference model that enables the derivation of a series of programmatic representations. The IDS vocabulary defines a relatively minimal, domain-independent core model and relies on standard and customized third-party vocabularies to express domain-specific facts. It is common practice to reuse existing specialized vocabularies and standards wherever possible to promote acceptance and interoperability.

– *Programmatic representation:* The programmatic representation of the information model is aimed at software providers by supporting the seamless integration of the information model into a development infrastructure with which software developers are familiar. It consists of a data model in a programming language that is delivered in the form of documented software libraries. The programmatic representation offers the best possible mapping of the IDS vocabulary to the native structures of a target programming language. This approach supports type-safe development, well-established unit tests and quality assurance processes.

Process layer

The process layer specifies the interactions between the various components of the International Data Spaces. The most important processes are:

– *Onboarding:* Onboarding describes what needs to be done to gain access to the International Data Spaces as a data provider or data consumer. Onboarding consists of registering and certifying the company that wants to operate the connector. This is followed by the purchase of a certified connector, which the company then configures and provides. The configuration involves obtaining a unique identity in the data space and the company must also provide a self-description of the connector. The last step required is for the company to connect the connector to its own system so that the data is available via the connector. The connector is now ready to be made known and available in the data space.

– *Data Offering:* The process of data offering describes the data offering or the search for suitable data. If data is only exchanged between two dedicated connectors, the two parties agree on the procedure. In Data Spaces, however, the normal case is that data providers do not know in advance which data consumers require the data. A complete description of the data is therefore an important prerequisite so that others can take advantage of the data. A connector can provide this description as a self-description or make it accessible to other connectors via a metadata broker.

– *Contract Negotiation:* Once connectors have been found, they must negotiate the contract, i.e. accept data offers by negotiating the usage guidelines, which range from simple access restrictions to complex obligations before and after data exchange. The negotiation of contracts is automatic. A data consumer can send a request for a contract to a data provider. The provider can accept or reject this request. A contract agreement can also be concluded between the Connectors. This requires mutual acceptance. A connector can revoke a contract negotiation at any time or save it in the clearing house to achieve a higher level of trustworthiness. Typically, a data consumer initiates the contract negotiation, but it can also work the other way round. It is also possible for Connectors to reject a request for a contract and respond with a counter-contract.

– *Exchanging Data:* The connectors can only exchange data with each other, i.e. transfer data between IDS participants, once all previous process steps have been successfully completed. A connector calls up the data operation that relates to the contract agreement. The architecture does not prescribe any specific protocols or technologies for data exchange. This means that data can be exchanged both synchronously and asynchronously. The data can be requested via pull-request or push-request.

– *Publishing and Using Data Apps:* Another relevant process is the publication and use of data apps. The apps are used to transform or process the data. The apps are published by app providers in the app store and certified by the certification body if required. A connector can search for apps in the app store, download the appropriate app and make the payment. Payment can be made directly between the app user and the app provider or processed via a clearing house. The app is then executed in the context of the connector.

These processes are related to the key values of the International Data Spaces and include most of the roles presented in the business level section.

System layer

Figure 36 shows the components of the data space. At the system level, the data space represents a distributed system in which the participating parties host the different data space components on their infrastructure. The connector initiates the exchange of data to and from the internal data resources of the participating organizations' enterprise systems and the data space. The Connector provides metadata to the Metadata Broker as specified in the Connector's self-description. This description may include information about the technical interface, authentication mechanisms, and associated data usage policies. Usage agreements are transmitted by the Connector to the Clearing House to establish trust. The Clearing House logs information about transactions to establish trust and transparency or to enable billing. The Vocabulary Hub provides vocabularies that specify the semantics of the exchanged data. Connectors can load applications from the App Store that run in the context of the connector and process data according to usage policies.

The System Layer maps the roles from the Business Layer and the processes from the Process Layer to a concrete data and service architecture. This represents the technical core of the data space (Figure 36).

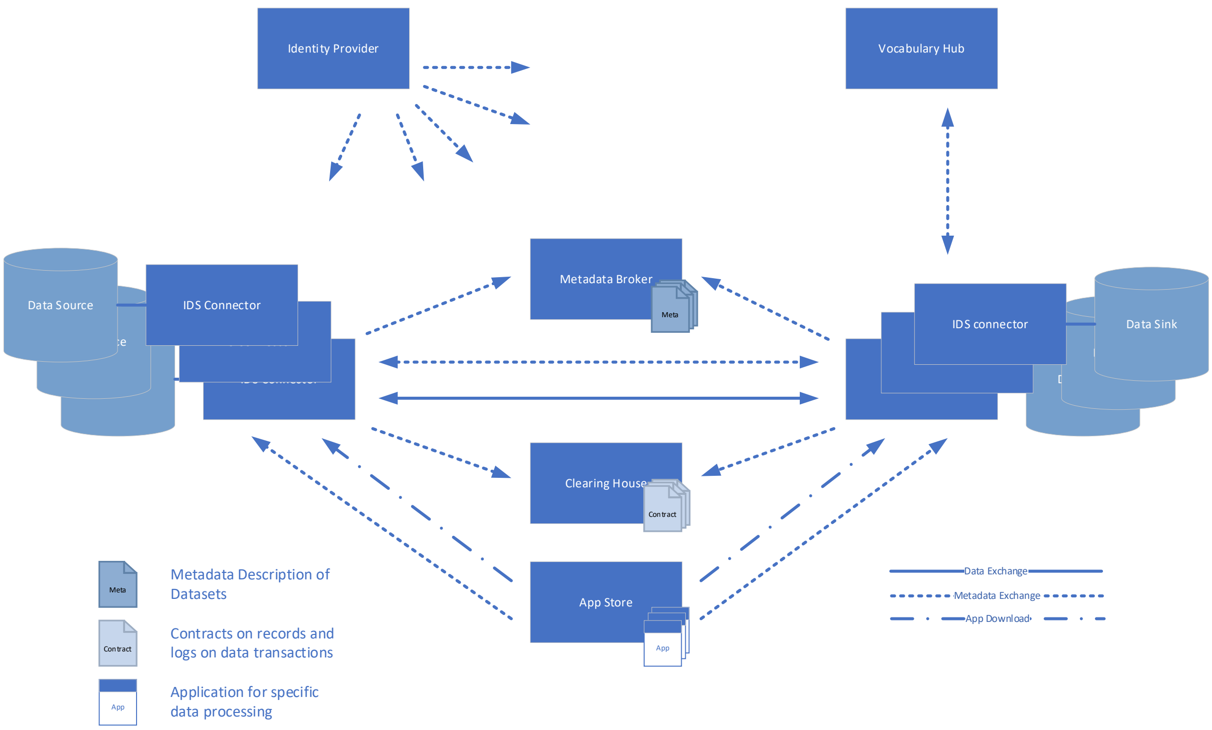

Figure 36 Technical components interaction (interactions to and from identity provider not shown)

The following section contains a more detailed description of the components of a data space, describing the internal structure and the functionality of the components.

Identity Provider

A concept for Identity and Access Management (IAM) is mandatory to base access-based decisions on reliable identities and properties of participants. In this concept, the aspects of identification, authentication and authorization must be defined in order to access the resources in the IDS. The identity provider consists of three components:

a. Certification Authority (CA), which issues and manages the identity claims on a technical level.

b. Dynamic Attribute Provisioning Service (DAPS), which provides tokens with the latest information about the connectors.

c. Participant Information Service (ParIS), which provides business-related information on IDS participants in a form that can be read by machines and humans.

Certification Authority (CA)

Connectors send certificate signing requests (CSR) to the CA, which then signs the CSR if they come from valid connectors. Invalid certificates are rejected. CAs are essential for establishing trust in a data space. It ensures that only registered organizations can operate components in the data space.

Dynamic Attribute Provisioning Service (DAPS)

The DAPS supplements the connector identity with current information and signs it. The DAPS embeds this information in so-called Dynamic Attribute Tokens (DAT) and passes them on to the requesting connectors. DAPS verifies the status and validity of software manifests and company descriptions, which contain metadata on issued IDS certificates. At the same time, DAPS adds time-variable information, such as the location of the devices or the currently supported transport certificates. This procedure reduces the rejection of certificates because dynamic information can be added at any time. DAPS is not a connector, but an external service and is part of the trust establishment in the data space.

Participants Information Service (ParIS)

The ParIS provides business-relevant information of the entities participating in the data space, which has been verified by the data space organization.

The ParIS consists of several functional blocks, which can be implemented with different technology stacks and hosting solutions adapted to the application. The server hosts the IDS endpoints, a database persists the participants' self-descriptions. An IAM checks the identity claims of clients and validates their authorization by using the DATs. The IAM can be located in the surrounding identity provider. An optional index speeds up search queries, and a website allows people to interact with the ParIS.

There are two categories of interaction with the ParIS. In the first type of interaction, the ParIS provides information about the participants when they are onboarded into the data space, as well as when they are updated by the operators of the identity provider. This process is necessary to build the necessary trust. The second type of interaction enables information on participants to be queried and updated at any time.

ParIS enables requesting components to request the desired information from participants using a query language. This information is self-descriptions of the participants, which are created and checked during onboarding. During operation, these self-descriptions can be changed by the relevant participants and thus adapted to the current situation. If participants leave the data space, the self-descriptions are no longer accessible.

The core attributes of participants must be comprehensively maintained between the various components; further synchronization between different ParIS instances is not enforced.

Connector

A data space is a network formed by its Connectors. Each connector enables the exchange of data via the data endpoints it provides, which is why no central instance is required for data storage. Connectors should be accessible via the standard Internet Protocol (IP). A participant can operate several Connectors, for example for load balancing or data partitioning. Connectors can be operated both in local data centers and in the cloud environment.

The connector architecture uses application containers to ensure an isolated and secure environment for individual apps and connector functionalities. An app corresponds to an application that provides an Application Programming Interface (API) for storing, accessing, or processing data. To ensure the privacy of sensitive data, it should be processed as close as possible to the data source. Any data processing (e.g. filtering, anonymization, or analysis) should be performed by the backend services or by apps. Only data that is to be made available to other participants should be offered by Connectors.

Apps are services for implementing business logic within the connector. Apps can be used to process data, connect to external systems, or control the connector. Apps can be downloaded via the App Store and provided by the Connector.

The App Store, the Metadata Broker and the Clearing House are based on the Connector architecture to support secure and trustworthy data exchange with these services.

App Store and Apps

An app is an independent, functional, and reusable software component that can be deployed, executed, and managed on a connector. Connectors can use apps for different purposes, including data apps, adapter apps and control apps, each of which fulfills different tasks in the IDS ecosystem. All app types can be downloaded and fully managed by the connector:

– *Data app*: Data apps are reusable, interchangeable, and independent of the connector. They perform small processing tasks such as data transformation, cleansing or analysis. In order to define data flows, the input and output points of the components involved (data app, connector and backend system) must be linked together. Data apps can be linked in order to carry out several processing steps on the same data.

– *Adapter app*: Adapter apps are also reusable, interchangeable, and independent of the connector. They provide access to company information systems and make them available to the underlying connector. The data flows of adapter apps are also defined by linking suitable input and output points of the components involved.

– *Control app*: Control apps enable the connector to be controlled from external systems and are used to connect backend systems to an IDS ecosystem. In contrast to the previously presented types, the control app works in the administrative control flow and is connector-specific, as it requires programming against the corresponding API of a connector in a specific version for its implementation.

The different app types can be bundled together to create a data processing chain with several apps of all types.

To integrate apps into an IDS ecosystem or link them to other components as described above, an app can be equipped with various endpoints. The endpoints for exchanging data between apps as well as between apps and connectors are mainly divided into those that consume data and those that provide data. A distinction is also made between endpoints that communicate exclusively internally and those that communicate with external components.

In addition to the endpoints mentioned above, there is the configuration endpoint and optionally the status endpoint. The configuration endpoint can be used to actively set or change configuration parameters during the runtime of an app. The so-called status endpoint, which can retrieve information on the status of an app during runtime, is optional.

The App Store is a secure platform for distributing apps. It consists of a directory of available apps and allows users to search for apps based on various criteria. The App Store supports operations for the registration, publication, maintenance, and retrieval of apps as well as the provision of apps to connectors and app users.

An app store also contains a connector to communicate with the connectors of app providers and app users within the data space. Each instance of an app store must comply with the connector certification criteria and provide the functionalities and endpoints of general connectors together with the operations mentioned above.

Metadata Broker

The metadata broker is a connector that makes it possible to register, publish, maintain, and query self-descriptions of the connectors in the data space. The self-description encapsulates information about a connector, its functionalities, and characteristics. It contains information about the interfaces provided, the owners of the component and metadata about the data provided by the component. The self-description is a type of metadata. It should be noted that the metadata broker is not a message broker in the true sense nor would provide similar functions.

A connector that offers a service or data registers its self-description with the metadata broker so that all other participating connectors can query this information. There may be several metadata brokers in a data space. In this case, the data space must use a synchronization mechanism to ensure that the information available in the various metadata brokers is consistent. The metadata broker stores the self-descriptions persistently and has efficient search options so that requesting components can access the required information. The metadata broker also accepts updated self-descriptions and stores them in its database.

The metadata broker itself also provides a self-description for the other components in the data space. The metadata broker must also have a valid identity and use valid DATs in communication. Most use cases require the metadata broker to make its information available in a human-readable format, e.g. via a website.

Metadata brokers are an optional component in data spaces. If they are available, they can be run on the infrastructure of the organizations participating in the data space (including the Data Space Authority).

Clearing House

The clearing house consists of a connector and relies in all its functions on a logging service that records the information relevant for clearing, billing and usage control. The information that is sent to the clearing house is defined in the process layer.

The clearing house uses this information to provide a clearing and settlement service based on user contracts and helps to automate payments between data providers and data consumers. It can also use this information to provide a billing service that enables the data space operator to settle accounts with the participants.

The clearing house can run on the infrastructure of the organizations participating in the data space (including the Data Space Authority).

Vocabulary Hub

The interoperability requirements in data space lead directly to the use of generally known, standardized terms to describe data, services, contracts, etc. The collection of these standardized terms forms so-called vocabularies. In the simplest case, any list of controlled terms can be a vocabulary. To be able to use their content, the corresponding vocabulary documents must be exchanged between the parties involved. The terms in the vocabulary must be machine-readable, including to some extent their descriptions and titles, and new terms must be available for lookup. The IDS information model is the central vocabulary shared by all participants in the data space.

The IDS information model represents the lowest common denominator for all use cases. It is therefore the minimum set of terms that all components in the data space must understand. It must be possible to extend the basic information model with additional vocabularies and they must be available via the Vocabulary Hub. The Vocabulary Hub is a platform for hosting, maintaining, publishing, and documenting the additional vocabularies. It provides endpoints to enable seamless communication with connectors and infrastructure components. Vocabulary Hubs provide access to the defined terms and their descriptions, display changes and provide an overview of the different versions.

Vocabularies are described in detail in Chapter 3 Ontological conceptual aspects.

Perspective certification

In order to guarantee data security and data sovereignty, the data space relies on consistent certification of all components. In this way, the Data Space ensures that all players and components adhere to the rules of the data space. Certification applies to all levels of the architecture model.

On the business layer, the certification perspective identifies the roles for certification. The roles of certification body and evaluation body belong to the governance body from the business layer. Certification applicants are organizations that have a role in one of the three categories «Main participant», «Intermediary» and «Software/service provider». The certification scheme describes for each role which certification level is required and what the focus of the certification is. At the functional level, the technical core components must fulfil the core security requirements, which is the basis for the compliance part of the certification of a core component. The assessment of conformity within the scope of certification also includes the information level. Compliance regarding functionality and protocols is considered.

Compatibility with the information model is also assessed. If the processes of the process layer are relevant for the conformity of a component or organization, they are also evaluated during certification regarding compliance with the processes. About the system layer, the certification of the components focuses on the security requirements. In this way, certification ensures that the components are suitable for forming a trustworthy data space.

In addition to the aforementioned Trust Levels, three Assurance Levels are applied as part of the certification process. At the level with the lowest requirements, an organization can carry out a self-assessment. At the next level, external bodies carry out an assessment of corporate strategies and processes. At the level with the highest requirements, external bodies also check measures to monitor compliance with company policy. At the component level, the assurance level at the lowest level means that organizations carry out a self-assessment using checklists and perform automated interoperability tests. The next level comprises external concept testing, including functional and security tests. At the highest level, external parties also evaluate the concepts and carry out tests and source code audits.

To ensure the high quality and transparency of the certification process, all evaluation facilities must first be approved by the impartial certification body. This process is divided into a preparatory phase in which all relevant information and documents are compiled. The audit phase then involves checking whether the evaluation facility complies with all the requirements of the data space and can be approved as an evaluation facility.

Perspective security

A strategic requirement for data spaces is to provide secure data exchange. This is imperative to build and maintain trust between the components and organizations participating in the data space. IDSA's data space security architecture includes measures to identify components in the data space, protect communication and data exchange, and control the use of the exchanged data. Connectors ensure that the specifications and requirements of the security architecture are applied to the interactions and operations in the data space.

In relation to the levels of the architecture, the security perspective is expressed in different ways. In the business layer, the security architecture forms the basis for trust in order to exchange and process data confidently. The various roles are responsible for establishing a trustworthy ecosystem. On the functional layer, this leads to functional requirements for data exchange and data processing. For example, the system must decide which transactions are permitted and which are restricted. Only then can many business models be implemented, especially if sensitive data is to be transferred to trusted business partners. Data usage policies are a key element in this. A common vocabulary is defined on the information layer, which enables basic information to be exchanged between the components of the data space. Such information includes the data usage policies, but also the functionalities of the components. The vocabulary ensures that everyone understands this information. The security requirements must be consistently adhered to in the processes in the process layer. In addition to creation, this also includes permanent monitoring and error correction. The technical implementation of the security requirements in the components takes place on the system layer. The security perspective provides a complete overview of the concepts for ensuring trust and security for all components in the data space.

Since players in a data space, who do not always already know each other, exchange data with each other, the data space infrastructure provides reliable information about the identity of the organizations and components involved. All players can query this information and thus build trust with the business partner. The basis for this is identity and trust management, which is anchored in the data space components. Each component and all actors have an identity that is secured by certificates. The certification authority (CA) of the data space is responsible for managing the certificates.

The hardware and software components of a data space, which can be hosted centrally or operated in a distributed manner, form the Trusted Computing Base (TCB). To ensure secure operation, hardware components such as Hardware Security Modules (HSM) or Trusted Platform Modules (TPM) are also required, which primarily enable the secure generation and management of keys. The use of Trusted Execution Environments should be considered to enable secure execution of the software. The platforms must implement several security requirements. These requirements are the integrity and authenticity of the software stacks, integrity verification and remote attestation, protection of the integrity and confidentiality of data at rest, isolation of processes, tracking of events, protection of keys in use and protection against malicious administrators.

Which of the requirements are to be implemented depends on the required trust level. In the data space, a distinction is made between three trust levels: 1) interoperability in the data space, 2) complete function for controlling data usage, and 3) additional protection against internal attacks. To implement security in accordance with the trust levels, a secure platform is required that can guarantee the secure and isolated execution of the applications. The platform provides security mechanisms and enables the applications to fulfil the security-relevant requirements. Secure data transmission requires secure communication with identification, authentication and authorization of the components. In addition, protection of the confidentiality and integrity of the transported data must be guaranteed.

The platform provides the appropriate access controls to ensure that data is used in accordance with the data usage policies. Of the various possible access control models, Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) are most commonly used.

Perspective governance

In the reference architecture model, the governance perspective defines the roles, functions and processes of the data spaces from the perspective of governance and compliance. A key point here is to describe the requirements to be fulfilled by the business ecosystem, which are necessary for secure and reliable interoperability between the companies. The architecture of International Data Spaces does not impose any restrictions on cooperation between organizations and does not require compliance with predefined rules. The participating organizations determine the rules to be followed. The architecture as a functional framework enables the rules to be implemented. The International Data Spaces support governance issues through various measures:

- Providing an infrastructure for data exchange, interoperability between organizations, and the use of new digital business models.
- Trustworthy relationships between data owner, data provider and data consumer.
- As trustee for the mediation between participating parties.

- Facilitating the negotiation of agreements and contracts.
- Strive for transparency and traceability of data exchange and data use.
- Enabling private and public data sharing.
- Consideration of participants' individual requirements.
- Offering a decentralized architecture that does not require central authority.

The management of data-related resources through decision-making rights, responsibilities, roles, and ownership makes data governance a fundamental element in the data space ecosystem. Data governance makes it possible to successfully engage in a collaborative ecosystem. It is therefore important to create organizational structures and processes that clearly describe who can make what kind of decisions and what responsibilities are associated with these decisions.

The relationship between governance and the five levels of the reference architecture model is described below. Governance considers the business aspect of data ownership, data provision and data consumption. Governance also describes concepts for core services, such as metadata brokerage. At the functional level, interoperability and connectivity must also be ensured from a governance perspective. The governance perspective plays an important role at the information level, especially in the basic vocabulary, also in that the metadata for the data space itself is defined in the vocabulary. Onboarding, data exchange and data use are the processes at process level that are directly linked to the governance perspective. At system level, the governance perspective plays an important role because the technical implementation of various security levels for data exchange must comply with the governance requirements.

The governance model defines a framework of decision-making rights and processes in relation to the definition, creation, processing, and use of data. While the governance activities provide the general direction of the decision-making system, data management comprises three groups of activities relating to the creation, processing, and use of data. These three groups relate to management, metadata, and the lifecycle of data.

The following matrix in Table 20 for the assignment of responsibilities (RACI matrix) supports the assignment of these activities to enable a governance mechanism in the data space ecosystem.

Table 20 RACI-Matrix; R: responsible, A: accountable, C: consulted, I: informed, S: supported. With focus on R, A and S.

	Activity	Data Owner / Data Provider	Data User / Data Consumer	Metadata Broker	Clearing House
Management	Determine data usage restrictions (execute data ownership rights)	R, A	-	S	-
	Enforce data usage restrictions	-	R, A	-	-

	Activity	Data Owner / Data Provider	Data User / Data Consumer	Metadata Broker	Clearing House
	Ensure data quality	R, A	-	S	-
	Monitor and log data transactions	S	S	-	R, A
	Enable data provenance	S	S		-
	Provide clearing services	S	S	-	R, A
Metadata	Describe and publish metadata	R, A	-	S	-
	Look up and retrieve metadata	-	R, A	S	-
Data Lifecycle	Capture and create data	R, A	-	-	-
	Store data	R, A	S	-	-
	Enrich and aggregate data	S	R, A	S	-
	Distribute and provide data	R, A	-	S	-
	Link data	S	S	R, A	-

From a governance perspective, data is an economic asset. Governance therefore focuses its activities on enabling new digital business models through the exchange of data. However, this considers the demands of data owners. This data ownership is another issue that must be respected by governance. In the legal sense, there is no ownership of data, as it is an intangible asset. This is why IDS is using the term data sovereignty. This expresses the fact that the data owner always retains control over the use of their data. To derive the greatest possible benefit from data processing and use, data quality is of great importance. In the context of International Data Spaces, all participants can view and assess the quality of the data sources. This transparency should lead to data providers taking the maintenance of their data seriously. In International Data Spaces, transparency also applies to the data transactions themselves. The clearinghouse stores all transactions so that it is possible to trace the origin of the data.

A2 Overview table of existing ontologies

The following resource provides a more comprehensive list of ontologies pertaining to the building and energy domains: http://smartcity.linkeddata.es/ . Table 21 offers an overview of the ontologies, supplemented with associate online resources.

Table 21 Overview of relevant ontologies

Ontology / Data model	Authors	Field / Specific scope	Link
CIM	IEC	Energy / Power grids, energy markets	https://www.entsoe.eu/digital/common-information-model/
IEC61850 data models	IEC	Energy / Distribution substations, DERs	https://iec61850.dvl.iec.ch/
SAREF4ENER	ETSI	Energy / All use cases	https://saref.etsi.org/saref4ener/v1.1.2/
EEBus ontology	EEBus association	Energy / Home and industrial energy management	https://www.eebus.org/
OpenADR information model	OpenADR alliance	Energy / Demand Response events and tariffs	https://www.openadr.org/
USEF flex protocol	USEF	Energy / Communication between flex aggregator and flex providers	https://github.com/USEF-Foundation/UFTP https://www.usef.energy/app/uploads/2020/01/USEF-Flex-Trading-Protocol-Specifications-1.01.pdf
CSA Smart energy	CSA	Energy / Energy management and metering	https://csa-iot.org/all-solutions/smart-energy/
OpenHab semantic model	OpenHaB Community and foundation	Home automation	https://www.openhab.org/
IoT Lite	W3C	IoT / all use cases	https://www.w3.org/Submission/iot-lite/
Things description	W3C	IoT / all use cases	https://www.w3.org/TR/wot-thing-description11/

Ontology / Data model	Authors	Field / Specific scope	Link
SAREF	ETSI	All / NA	https://saref.etsi.org/saref4ener/v1.1.2/
OneM2M	oneM2M partnership	IoT / all use cases	https://www.onem2m.org/technical/onem2m-ontologies
Bricks	Open source – community driven	Building / building description	https://brickschema.org/ontology
EEPSA	Academia	Building / Energy efficiency and comfort	https://iesnaola.github.io/eepsa/EEPSA/index-en.html
BIM ontology	ISO	Building / Life cycle management	https://www.bsigroup.com/en-GB/iso-19650-BIM/
SAREF4BLDG	ETSI	Building / all use case	https://saref.etsi.org/saref4bldg/v1.1.2/

A3 SINA's Rulebook for a Fair Data Economy - General Terms and Conditions

1 APPLICABILITY, SCOPE, AND GOVERNANCE

The SINA association is established by the statutes, which is signed by the Founding Members of the ecosystem.

The provisions of these General Terms and Conditions will become applicable to and legally binding on the data sharing agreements of the Parties to the ecosystem upon the execution of the Accession Agreements, as applicable.

If a discrepancy arises between any of the terms and conditions established in the Statutes, any Accession Agreements and these General Terms and Conditions, including any of their appendices or schedules, any such discrepancy will be resolved in accordance with the following order of priority:

the clauses of the Statutes;

Dataset Terms of Use and related Schedules;

these General Terms and Conditions

Any amendments to or derogations from these General Terms and Conditions must be agreed upon in the general meeting in order to be valid.

2 DEFINITIONS

2.1 In these General Terms and Conditions, the following capitalised terms and expressions have the following meanings, and the singular (where appropriate) includes the plural and vice versa:

2.2 **«Accession Agreement»** means the agreement that governs the admission of associates of the SINA association and participants of the ecosystem.

2.3 **«Affiliate»** means any individual, company, corporation, partnership or other entity that, directly or indirectly, controls, is controlled by, or is under shared control with Party.

2.4 **«Appendix»** means any appendix to the Statutes.

2.5 **«Confidential Information»** refers to trade secrets as defined in the EU Directive 2016/943 of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, point (1) of Article 2 provided it is: (a) if disclosed in writing or in other tangible form, clearly marked as confidential or proprietary by the disclosing Party at the time of disclosure, or (b) if disclosed in other than tangible form, identified as confidential at the time of disclosure and confirmed and designated in writing to the receiving Party within fourteen (14) calendar days from the disclosure as confidential information by the disclosing Party.

2.6 **«Statutes»** means the legal document that outlines, the rules, structure and governance of the SINA association, specifying its purpose, membership requirements and operational procedures.

2.7 **«Data»** means any information that Data Providers have distributed, transmitted, shared or otherwise made available to the ecosystem based on the Statutes and during its period of validity as further defined in the respective Dataset Terms of Use.

2.8 **«Ecosystem»** means the group consisting of the Parties who share Data in accordance with the Statutes on SINA.

2.9 **«Data Provider»** means any natural person or an organization that provides Data for the Parties to use via the ecosystem.

2.10 **«Dataset»** means a collection of Data whose use the Data Provider authorizes via the Data ecosystem. Datasets and their related terms and conditions are defined more in more detail in the respective Dataset Terms of Use.

2.11 **«Dataset Terms of Use»** means the terms under which the Data Provider grants a right to use the Data included in the Dataset to the Service Providers and/or End Users.

2.12 **«Derived Material»** means information derived from Data or information that is created as a result of the combination, refining and/or processing of Data with other data. In case there is a need to clarify the borderline between Data and Derived material, additional requirements for what is not considered Derived Material shall be identified in the respective Dataset Terms of Use,

2.13 **«End User»** means any of the Parties to which Service Providers provide Data and/or services or to which the Data Provider provides Data, and which do not redistribute the Data further.

2.14 **«Founding Members»** are the founding members of the association.

2.15 **«Intellectual Property Rights»** means patents, trademarks, trade and business names, design rights, utility models, copyrights (including copyrights in computer software), and database rights, in each case registered or unregistered and including any similar rights to any of these rights in any jurisdiction and any pending applications or rights to apply for the registration of any of these rights.

2.16 **«Operator»** means any Party that provides data system or any other infrastructure services for the ecosystem that are related e.g., to identity or consent management, logging or service management.

2.17 **«Operator Service Agreement»** means any service level agreements governing the services provided by any of the Operators to the ecosystem or to its Members.

2.18 **«Party»** or «**Member»** means a member or an associate of the SINA association.

2.19 **«Associate»** means not a member of the SINA association but is obliged to adhere to its rules and regulations and actively participates in the ecosystem.

2.20 **«Personal Data»** has the meaning set forth in Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (the General Data Protection Regulation) («GDPR»).

2.21 **«Schedule»** means any schedule to the Dataset Terms of Use.

2.22 **«Service Provider»** means any of the Parties that combines, refines, and processes data and provides the processed Data and/or a service, which is based on the Data, to the use of End Users, other Service Providers or Third-Party End Users.

2.23 **«Third Party**» means a party other than a Party.

2.24 **«Third Party End User**» means any Third Party that receives any Data directly or indirectly from any of the Service Providers.

3 ROLE-SPECIFIC RESPONSIBILITIES

3.1 The potential roles defined under these General Terms and Conditions are (1) the Data Provider, (2) the Service Provider, (3) the End User and (4) the Operator. A Party may simultaneously occupy multiple roles. In such a case, the relevant Party must comply with all applicable obligations related to each role and relevant Data. In addition, Third Party End User is a role recognized under these General Terms and Conditions as applying to any stakeholders who are not a Party to the Statutes but who receive Data.

Data Provider

3.2 The Data Provider will be responsible for defining the Dataset Terms of Use for any Data that the Data Provider makes available within the ecosystem. This includes the right to define the purposes for which relevant Data can be processed, the right to allow the redistribution of Data to End Users and, where applicable, to Third Party End Users, and the right to prohibit the unauthorized use of Data and the right to cease sharing Data within the ecosystem. The Data Provider must notify the Parties to whom the Data Provider makes the Dataset available of any new Dataset Terms of Use, after which the Dataset Terms of Use will bind the other Parties. Unless otherwise defined in the applicable Dataset Terms of Use, any changes introduced by the Data Provider to the applicable Dataset Terms of Use will become effective within thirty (30) days from the relevant Parties to the ecosystem being sent a notification of such change. Changes to the Dataset Terms of Use must not have retroactive effect.

3.3 The Data Provider shall provide Data for the use of the ecosystem in a machine-readable form and by a method as defined by the Data Provider in the applicable Dataset Terms of Use (e.g., application programming interface, downloadable package, or other method).

3.4 As an exception to the above clause 3.3, the Data Provider may undertake to grant the right to use certain specific Datasets or types of data to the ecosystem for a fixed period, in order to protect investments made in the ecosystem by other Parties in good faith.

Service Provider

3.5 The Service Provider will be responsible for processing Data in accordance with the Statutes, Accession Agreement, and the applicable Dataset Terms of Use.

3.6 The Service Provider must keep records of its processing activities and deliver, on request, reasonably detailed reports on usage, processing and redistribution of Data to the relevant Data Provider(s).

End User

3.7 The End User must use Data in accordance with the Statutes and the applicable Dataset Terms of Use.

Operator

3.8 The ecosystem may involve one or several Operators. The Operator(s) are responsible for providing the ecosystem with services that facilitate the operations of the relevant ecosystem, such as authentication, identification, and identity/consent management services or for ensuring data security or providing technical data protection solutions for the ecosystem and as further defined in the applicable Operator Service Agreement.

3.9 Any Operator Service Agreement(s) concluded with the Party/Parties and the Operator(s) may be included as an Appendix to the Statutes.

3.10 Operator shall adhere to any regulatory requirements such as notifications required by applicable legislation.

4 REDISTRIBUTION OF DATA

4.1 The Parties shall have the right to redistribute the Data to the other Parties, unless such redistribution has been specifically prohibited under applicable Dataset Terms of Use. Parties can redistribute Data to Third Party End Users only if permitted under the applicable Dataset Terms of Use.

4.2 If the Data Provider chooses to allow redistribution of the Data to Third Party End users, the Data Provider shall be responsible for determining those Dataset Terms of Use which apply to the redistribution. A Service Provider must include such terms and conditions concerning Data redistribution into any agreements or terms and conditions with Third Party End Users.

4.3 Notwithstanding the above, the Parties shall have the right to redistribute Data to their Affiliates, unless applicable Dataset Terms of use explicitly prohibit such redistribution. Each Party shall be responsible for ensuring that their Affiliates comply with the Statutes.

4.4 Derived Material and its Redistribution.

4.5 Rights to Derived Material shall belong to the Party generating such Derived Material and the restrictions of use set out for the Data in the Dataset Terms of Use shall not cover Derived Material. Any restrictions for the use or redistribution of Derived Material shall be explicitly set out in the Dataset Terms of Use, if any.

4.6 The Parties are entitled to redistribute Derived Materials to the other Parties and any Third Party, unless specifically prohibited in the applicable Dataset Terms of Use.

4.7 Processing and Redistribution of Personal Data.

4.8 The redistribution of any Personal Data or Derived Materials created on the basis of any Personal Data may be subject to more detailed requirements and restrictions. Each Data Controller shall on its own behalf ensure that any redistribution and other use of Derived Material shall take place in accordance with applicable data protection legislation. Additionally, any conduct between Data Controllers and Data Processors shall be subject to the applicable Data Processing Agreements. The Parties may also choose to separately agree on more detailed stipulations about the processing of Personal Data as part of the Dataset Terms of Use.

5 GENERAL RESPONSIBILITIES

Data security, protection and management

5.1 Each Party must designate a contact person for data security matters, who is responsible for the relevant Party's data systems that are connected to the ecosystem and for the implementation of the Party's security policy.

5.2 Each Party to the ecosystem must have sufficient capabilities to process Data securely and in accordance with the relevant data security standards and data protection legislation. The Parties must implement and maintain suitable technical, organisational, and physical measures that are in line with good market practice, by considering the nature of the Data processed by the Party. Each Party must have the capability to properly perform its obligations under the Statutes and applicable Dataset Terms of Use and, where necessary, to cease processing activities without undue delay for any relevant reason.

5.3 The aforementioned capabilities include e.g. the capability to control Data and its processing by being aware of

(i) the origins of the Data (specifically whether the origin is the Party itself, another Party or Third Party);

(ii) the basis for processing Data;

(iii) the restrictions and limitations that apply to processing Data; and

(iv) the rights and restrictions that apply to redistributing or refining Data.

5.4 Parties must also be capable of recognizing Data and removing or returning it if the basis for the processing of Data expires. The obligation to remove or return Data is not applicable to Derived Materials.

5.5 Any identified data security breaches must be duly documented, rectified and reported to the affected Parties without undue delay. All involved Parties have a mutual responsibility to contribute reasonably to the investigation of any data security breaches within the ecosystem.

Subcontractors

5.6 The Parties will have the right to employ subcontractors to perform their obligations under the Statutes. Where and to the extent that the outsourced functions require it, the Parties may allow their subcontractors to access Data. The Parties will be responsible for the subcontracted performance as for their own.

6 FEES AND COSTS

6.1 Data is shared within the ecosystem free of charge, unless otherwise defined in the applicable Dataset Terms of Use.

6.2 Each Party will bear their own costs related to accessing the ecosystem and operating as a Member or associate of the SINA association.

6.3 Unless otherwise agreed by Parties, the joint costs incurred for the maintenance and administration of the ecosystem will be allocated in equal shares between the Parties. For the avoidance of doubt, the maintenance and administration of the ecosystem does not include the costs of Data where applicable and as defined in the Dataset Terms of Use in question.

7 CONFIDENTIALITY

7.1 The Parties must use any Confidential Information they receive in connection with the operation of the ecosystem and/or regarding the ecosystem only for the purposes for which such Confidential Information has been provided. The Parties must not unlawfully use or disclose to Third Parties any such Confidential Information they have become aware of in the course of the operation of the ecosystem.

7.2 If a Party is, under the applicable law or an order issued by a competent authority, obliged to disclose another Party's Confidential Information to the authorities or Third Parties, the obliged Party must promptly notify the affected Party whose Confidential Information will be disclosed of such disclosure if so permitted under the applicable law or the competent authority's order.

7.3 The confidentiality obligations established in these General Terms and Conditions will survive the termination of the SINA Association.

8 INTELLECTUAL PROPERTY RIGHTS

8.1 The Intellectual Property Rights of the Parties must be respected and protected in connection with the operation of the ecosystem.

8.2 Being a member or an associate of the SINA association and sharing any Data within the ecosystem does not result in the transfer of any Intellectual Property Rights. More specific provisions, if any, concerning the Intellectual Property Rights that relate to specific Datasets are included in the applicable Dataset Terms of Use. For the avoidance of doubt, any new Intellectual Property Rights created by a Party will vest in the creating Party as further defined in the applicable legislation governing Intellectual Property Rights.

8.3 Data Provider is responsible for ensuring that it has sufficient rights for the provision of Data in accordance with the Dataset Terms of Use.

8.4 The Parties are entitled to utilize software robots or other forms and applications of robotic process automation or machine learning or artificial intelligence when processing Data. In accordance with the aforementioned and the applicable Dataset Terms of Use, the Parties have the right to learn from Data and to use any professional skills and experience acquired when processing Data.

9 DATA PROTECTION

9.1 Any Personal Data processed within the ecosystem must be processed in accordance with the applicable data protection laws and regulations.

9.2 Terms that are not defined here, have the meaning stated in the GDPR or other applicable data protection laws (e.g. FADP).

9.3 For the purposes of processing Personal Data within the ecosystem, any Parties disclosing or receiving Data are, individually and separately, assumed to be controllers under the provisions of the GDPR. The said Parties are also assumed to be processing Data on their own behalf unless the Parties have concluded a written Data Processing Agreement that sets out the subject matter and duration of the processing, the nature and purpose of the processing, the type of Personal Data and categories of data subjects and the obligations and rights of the controller and the processer. Where any such Data Processing Agreement is applicable in general to certain Dataset(s) or services provided under the Statutes, it must be included as an Appendix to the Statutes.

9.4 The Parties must prevent the unauthorized and unlawful processing of Personal Data by employing appropriate technical and organisational measures. The Parties must ensure that persons allowed to process Personal Data have committed to keeping such data confidential or are bound by an appropriate statutory obligation of confidentiality.

9.5 Personal Data that is shared within the ecosystem can be transferred within the European Union and the European Economic Area (EEA). This kind of Personal Data can also be transferred outside the EU and the EEA in compliance with the applicable data protection legislation and case law, unless otherwise prescribed by the applicable Dataset Terms of Use.

9.6 The Parties commit to provide reasonable assistance to the other Parties where such assistance is needed in order for the other Party to comply with its obligations under the applicable data protection legislation.

10 TERMINATION AND VALIDITY

10.1 The Parties are entitled to continue to use any Data received through the ecosystem prior to the termination of the Statutes, unless otherwise determined in the applicable Dataset Terms of Use or agreed by the Parties in the Statutes. In such case, the clauses governing use of Data in these General Terms and Conditions, Dataset Terms of Use and/or in the Statutes, remain in force according to the Clause 17.1.

10.2 Any Party may choose to leave the association as defined in the Statutes. Notice of termination must be provided in writing to the Members of the SINA association.

10.3 In the event that there are more than two members to the SINA association and one member commits a material breach of the provisions of the Statutes, the general meeting will have the right to terminate the membership with the breaching member or associate with immediate effect. Notice of any such termination must be provided in writing to all members and associates.

10.4 If the breach can be rectified, the non-breaching Party/Parties may resolve to suspend the performance of their obligations under the Statutes until the breaching Party has rectified the breach.

10.5 Where a Member's membership in the ecosystem is terminated as a consequence of the Member's material breach of the Statutes, the breaching Member's right to use the Data will end at the date of the termination. The breaching Member must cease to use the Data and, upon request by any Party, verifiably return or destroy Data and any copies of Confidential Information including copies thereof. However, the breaching Member is entitled to retain the Data as required by the applicable law or competent authorities provided that the breaching Member notifies the Data Provider of such a data retention obligation by the date of termination.

11 LIABILITY

11.1 The Parties will only be liable for direct damages that result from a breach of the provisions of the Statutes as defined hereinafter and where applicable, in the Statutes. Any other liabilities are hereby excluded, unless otherwise specifically defined in the Statutes.

Parties are not liable for loss of profits or damage that is due to a decrease or interruption in production or turnover, or other indirect or consequential damages.[25]

11.2 The Parties will not be liable for any losses, damages, costs, claims or expenses howsoever arising from a mechanical or electrical breakdown or a power failure or any other cause beyond the reasonable control of the Party; and the Parties must fully compensate any damages resulting from an intentional or grossly negligent breach of the provisions set out in the Statutes.

11.3 Each Party, severally and not jointly, will be liable for any infringements of personal data obligations set out in the GDPR in accordance with Article 82 of the GDPR.

12 FORCE MAJEURE

12.1 No Party will be liable for injuries or damage that arise from events or circumstances that could not be reasonably expected beforehand and are beyond its control (force majeure).

12.2 A Party that is unable to perform its obligations due to an event of force majeure must inform other Parties of any such impediment without undue delay. These grounds for non-performance will expire at the moment that the force majeure event passes. This clause is subject to a long-stop date: where performance is prevented for a continuous period of one hundred and eighty (180) days or more, the Parties are entitled to terminate the Statutes as set forth in clause 10.5 or 10.6, as applicable.

13 AUDIT

13.1 A Data Provider will be entitled to audit the Parties processing the Data made available by the Data Provider at its own expense, including also material and reasonable direct costs of the audited Party. The purpose and the scope of the audit is limited to verifying compliance with the material requirements of the Statutes, the applicable Dataset Terms of Use, and applicable legislation.

13.2 The Parties are responsible for imposing the same auditing obligations as set out herein on their Affiliates and the Parties will act in good faith to ensure that the objectives of the Data Provider's audit rights materialize with regard to the subcontractors of a Party.

13.3 The auditing Party must notify the audited Party of the audit in writing at least thirty (30) days prior to the audit. The written notice must disclose the scope and duration of the audit and include a list of requested materials and access rights.

13.4 The audited Party is entitled to require that the audit is conducted by a mutually acceptable and/or certified independent Third Party.

13.5 The Parties are required to retain and provide to the auditing Party and/or the Third Party auditor, for the purposes of the audit, all records and documents as well as access to all

[25] Parties may wish to note that the concept of indirect or consequential damage varies between different jurisdictions.

necessary data systems and premises and to interview personnel that are of significant importance for the audit. Records and documents thus retained must span to the previous audit or to the accession of the audited Party to the ecosystem, whichever is later.

13.6 The auditing Party and/or Third Party auditor may only request such records and documents and such access to data systems and premises and to interview personnel that are of significant importance to the audit.

13.7 All records, documents and information collected and disclosed in the course of the audit constitute Confidential Information. The auditing Party and/or Third Party auditor may not unlawfully utilize or disclose Confidential Information that it has become aware of in the course of the audit. The auditing Party represents and warrants that any Third Party auditor, where applicable, complies with the applicable confidentiality obligations. The audited Party is entitled to require that the auditing Party and/or Third Party auditor or any other persons participating in the audit sign a personal non-disclosure agreement provided that the terms and conditions of such a non-disclosure agreement are reasonable.

13.8 The results, findings and recommendations of the audit must be presented in an audit report. The audited Party is entitled to review any Third Party auditor's audit report in advance (and prior to it being provided to the relevant Data Provider(s) by the Third Party auditor). The audited Party is entitled to require the Third Party auditor to make any such changes to the audit report that are considered reasonable while taking into account the audited Party's Confidential Information and the applicable Data Provider's business interests in the Data. The audited Party must provide its response to the audit report within thirty (30) days. If no response is provided, the audited Party is considered to have accepted the contents of the report.

13.9 If the auditing Party justifiably believes the audited Party to be in material breach of the obligations imposed thereupon in the Statutes, an additional audit may be conducted.

13.10 In the event that the audit reveals a material breach of the obligations imposed in the Statutes or the applicable Dataset Terms of Use, the audited Party will be liable for reasonable and verifiable direct expenses incurred as a result of the audit.

14 APPLICABLE LAWS AND DISPUTE RESOLUTION

14.1 The agreement incorporating these General Terms and Conditions is governed by and construed in accordance with the laws of Switzerland without regard to its principles of private international law and conflict of laws rules.

14.2 Any dispute, controversy or claim arising out of or in relation to the agreements based on the General Terms and Conditions, or the breach, termination or validity thereof, shall be finally settled by arbitration in accordance with the Swiss Chambers' Arbitration Centre. The number of arbitrators shall be one, the seat of arbitration shall be Zurich, Switzerland and the language of the arbitration shall be German.

15 OTHER PROVISIONS

15.1 Unless otherwise agreed by the Parties, any amendments to the Statutes and its Appendices must be made in writing and signed by all members.

15.2 If any provision of the General Terms and Agreements or any applicable Dataset Terms of Use is found to be invalid by a court of law or other competent authority, the invalidity of that provision will not affect the validity of the other provisions established in the Statutes.

15.3 Each party represents and warrants that it is validly existing and in good standing under the applicable laws of the state of its incorporation or registration. Each Party also represents and warrants that it has all required power and authority to execute, deliver, and perform its obligations under the Statutes and, where applicable, to bind its Affiliates.

15.4 The Parties intend to create an ecosystem that is subject to a single set of contractual terms, and nothing contained in the Statutes may be construed to imply that they are partners or parties to a joint venture or the other Parties' principals, agents or employees. No Party will have any right, power, or authority, express or implied, to bind any other Party.

15.5 No delay or omission by any Party hereto to exercise any right or power hereunder will impair such right or power, nor may it be construed to be a waiver thereof. A waiver by any of the Parties of any of the covenants to be performed by the other Parties or any breach thereof may not be construed to be a waiver of any succeeding breach thereof or of any other covenant.

16 NOTICES

16.1 All notices relating to these General Terms and Conditions must be sent in a written or electronic form (including post or email) or delivered in person to the contact person and/or address specified by the respective Party in the Statutes or in the applicable Accession Agreement. Each Party will be responsible for ensuring that their contact details are up to date. Notices will be deemed to have been received three (3) days after being sent or on proof of delivery.

17 SURVIVAL

17.1 Clauses 1, 2, 3, 4, 5, 8, 9, 11, 14, 16, and 17 of these General Terms and Conditions will survive the termination of the Statutes in its entirety together with any clauses of the Statutes that logically ought to survive the termination of the SINA association.

17.2 Clause 13 of these General Terms and Conditions will survive for a period of three (3) years following the termination of the SINA association in its entirety.

17.3 Clause 7 of these General Terms and Conditions will survive for a period of five (5) years following the termination of the SINA association in its entirety.

A4 Methodology to derive the use cases and potential business models for SINA

Use cases, in general, are used for describing the requirements of the system, often helping in creating and validating the proposed designs (Schneider & Winters, 2001) describe a use case as a «collection of possible sequences of interactions between the system under discussion and its external actors, related to a particular goal». A use case typically captures the static view of actors being related to the proposed system as well as peer-to-peers the dynamic view describing the relation between actors and the system (Gottschalk et al., 2017).

The exploration of events whereby various actors interact with the proposed system (e.g., new technology adoption) in different ways allows to identify how value can be created and the goals of the actors fulfilled (Gottschalk et al., 2017; Maghazei et al., 2022). Figure 37 demonstrates how use cases logic could be applied to trial new technologies and search for the right fit (Maghazei et al., 2022).

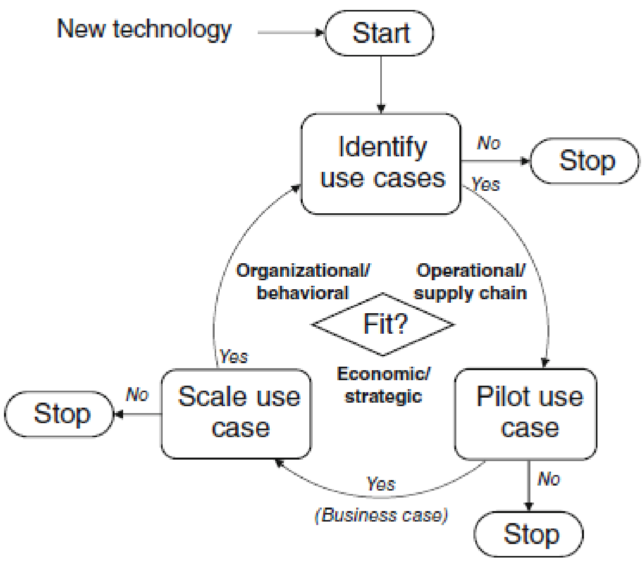

Figure 37 Technology adoption model following the use case logic. Source: Maghazei et al. (2022)

Having developed this process Maghazei et al. (2022), and later Netland et al. (2022), advise to start by focusing on developing the use cases and only then proceed to developing a business case for how cutting-edge technology could be applied to solve a specific business problem.

In the next section the steps taken to develop and validate SINA use cases are presented and their applications explained.

Use case development and validation

Establishing the use cases for SINA was approached systematically, iterating between the stages of use case development and use case validation with industry partners (Table 22).

Table 22 Stages in development and validation of use cases

Step	Activity	Outcome
Step 1 Use case scoping	Academic and grey literature review conducted to draw potential use cases. Use case scoping workshop conducted with industrial partners.	One workshop was conducted. Initial twenty use cases identified.
Step 2 Use case ecosystem	Further academic and grey literature review was performed to detail the use cases. Three overarching use case ecosystems identified.	Twenty use cases detailed. Three ecosystems identified.
Step 3 Use case validation	Conducting use case validation workshops with industrial partners.	Ten workshops conducted.
Step 4 In-depth use case description	Developing three in-depth use cases to demonstrate potential SINA applications.	Three in-depth SINA use cases developed.

Step 1 Use case scoping

Academic and grey literature review was conducted to get an understanding of what potential use cases could be enabled by SINA. Following the initial review, industrial partners were invited to engage with the development of SINA use cases.

On 19 December 2022 a use case scoping workshop was conducted with 10 industrial partners where use cases were identified, discussed, and prioritized. In this workshop the participants spent time detailing the pre-defined use cases, as well as defining new use cases as presented in Figure 38.

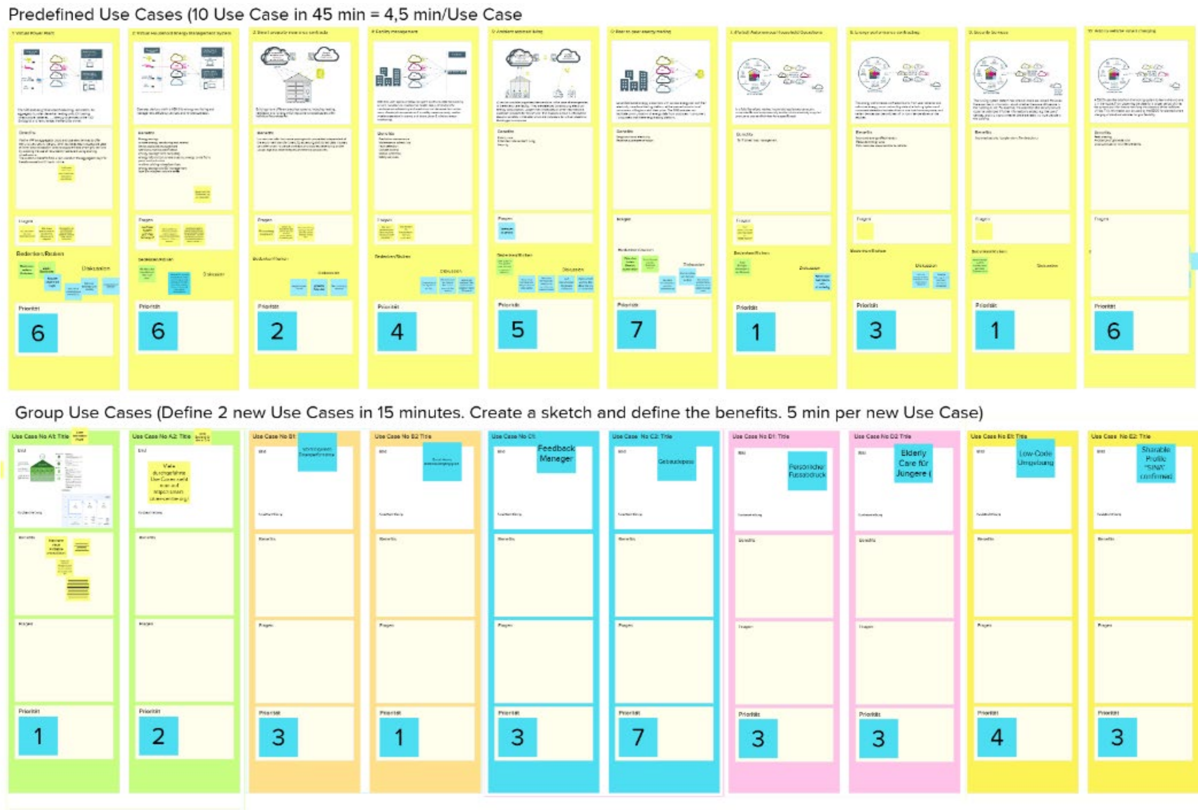

Figure 38 Use Case scoping workshop template

In the use case scoping workshop, the industrial partners were further asked to prioritize the use cases (see Table 23). The prioritization exercise has led to the identification of use cases holding the most potential for further development.

Table 23 Use case scoping workshop

	Use case	Priority
Predefined	Virtual Power Plant (VPP)	6
	Virtual Household Energy Management System (vHEMS)	6
	Smart property insurance contracts	2
	Facility management	4
	Ambient assisted living	5
	Peer-to-peer energy trading	7

	Use case	Priority
	(Partial) Autonomous Household Operations	1
	Energy performance contracting	3
	Security Services	1
	Grid-to-vehicle: smart charging	6
New	Asset Information Model	1
	Utility Backend for the el-Grid	2
	Value-added data performance	3
	Smart home data consistency	1
	Feedback Manager	3
	Building pass (predictive maintenance)	7
	Personal footprint	3
	Elderly Care for the younger	3
	Low code environment	4
	Sharable profile SINA confirmed	3

Further reflection on the prioritized use cases has led to the identification of three use case ecosystems (Figure 39).

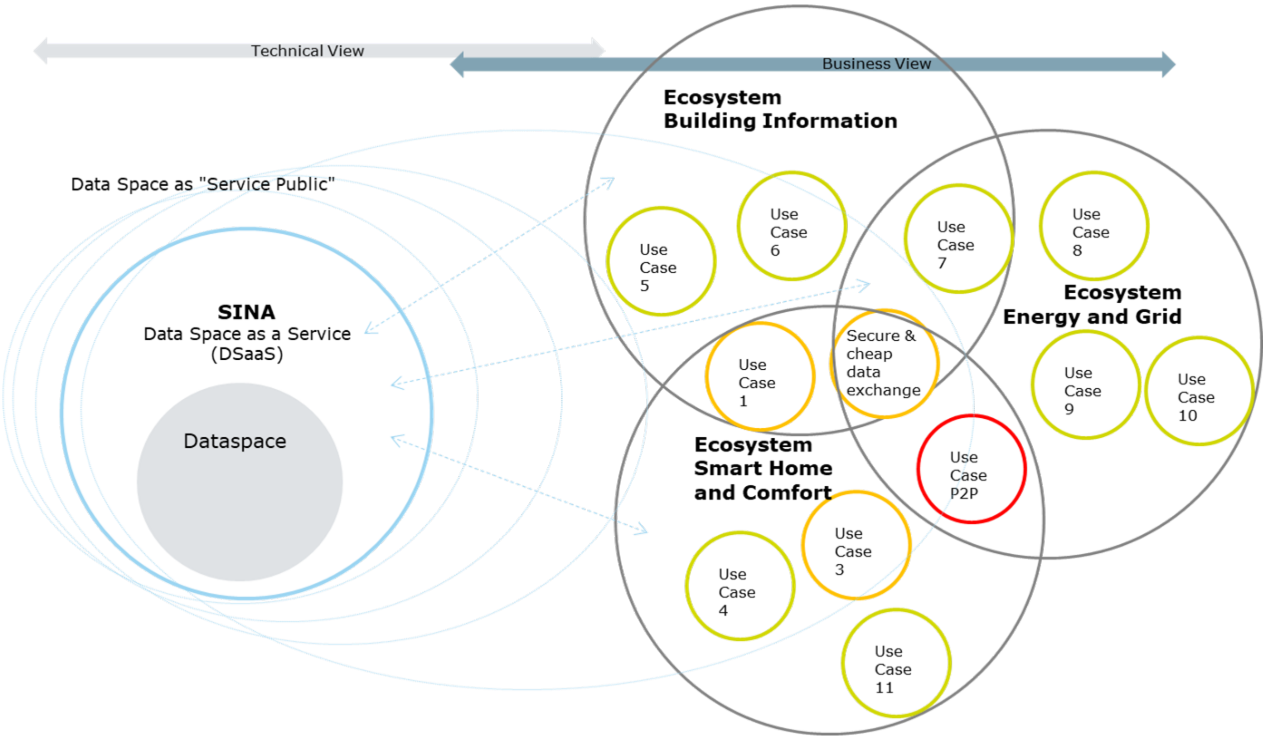

Figure 39 Use Case Ecosystems

Step 2 Use case ecosystems

Further literature review was performed to detail the use cases populating each of the identified use case ecosystems. The use cases populating each of the identified ecosystems were detailed with the narrative description and shared with industrial partners in subsequent use case validation workshops (Step 3 Use case validation with industrial partners)

Use case validation workshops were conducted with industrial partners from August 2023 to October 2023. The research team consisting of a minimum of two and maximum of four researchers conducted nine 2,5-hour workshops, one of which was followed up with an additional workshop. Each use case validation workshop was initiated with reviewing the aims of the project and the progress made to date, followed by a presentation of the identified SINA use case ecosystems, and use cases. The workshop participants were then invited to engage in a facilitated brainstorming exercise where the promising use cases were developed further, including: detailing the tasks to be solved by the use case, designing the potential value proposition, sketching an actor map, reflecting on the use cases' feasibility, as well as their potential economic and ecological benefits (Figure 40).

The input provided by the workshop participants was captured and shared with them after the workshop to obtain any further insights and reflections.

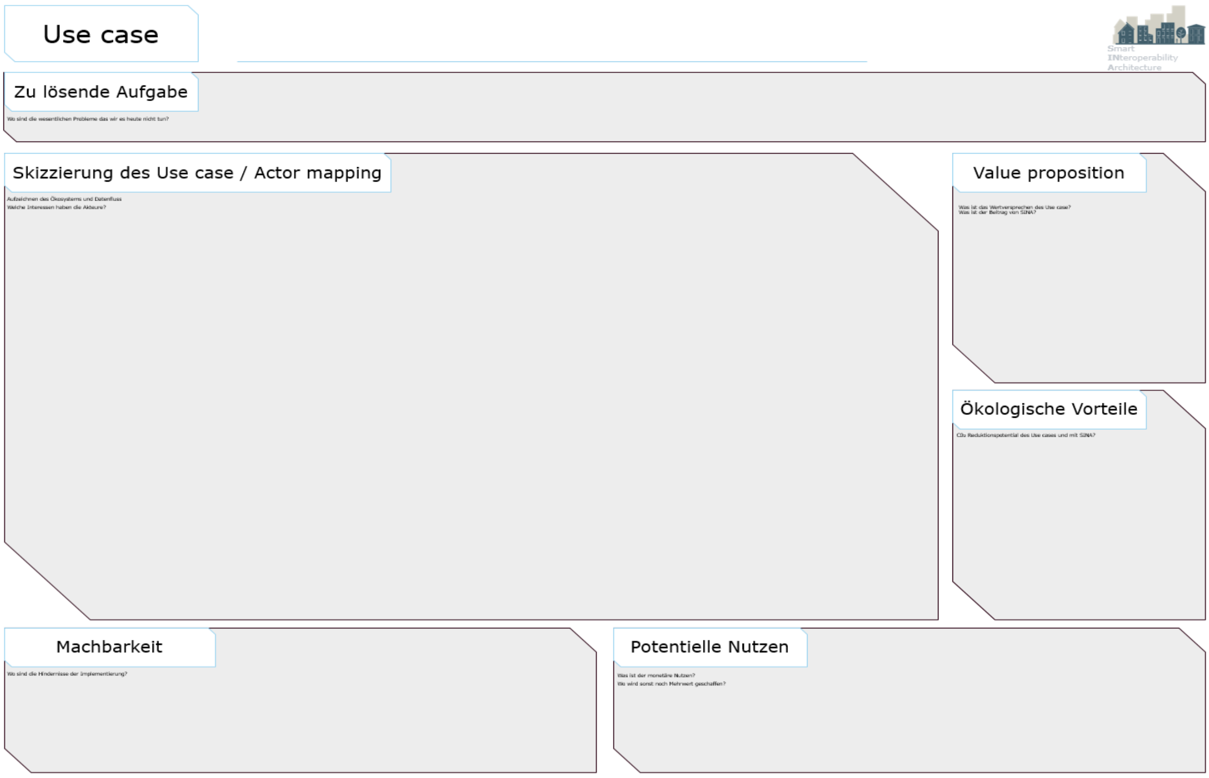

Figure 40 Use case validation workshop template

Step 3 In-depth use cases

Processing the findings from the use case validation workshops has enabled developing three in-depth use cases to demonstrate potential SINA applications as follows:

– Smoother data exchange between energy providers: Data-driven integration
– Data-Driven Facility Management Solution with Advanced Predictive Maintenance
– Peer-to-peer Energy Trading

The following section presents the findings emerging from the use case development and validation workshops conducted together with industrial partners, starting with an overview of the different stakeholders, which can be involved within a data space for the building industry.

Roles and stakeholders in the built environment

This section emphasizes the significance of delineating a comprehensive list of stakeholders for a data space (see Table 24). Identifying key players, including end-users, data owners, IT professionals, and regulatory bodies, is crucial for successful data space planning and management. A well-defined stakeholder list serves as a guide for effective communication, collaboration, and decision-making, promoting resilience and adaptability in the built environment.

Table 24 Stakeholders which could potentially interacting with a data space in the building industry

Role	Role / Responsibility
Power Generation	Responsible for power generation, utilizing various energy sources.
Power Transmission (TSOs)	Operates high-voltage transmission grids, connecting generation units and distribution grids.
Power Distribution (DSOs)	Manages low or medium voltage networks, delivering power to end-users.
Wholesale Market Operations	Computes wholesale prices based on production costs and demand forecasts.
Retailers	Sell power to end-users, forecasting and purchasing from the wholesale market.
Balance Services (BRPs)	Intermediary ensuring the consumed quantity matches reserved amounts
Aggregators	Orchestrate multiple loads for grid balancing and ancillary services, benefiting prosumers and grid stability.
Power Consumers	Includes all electrical appliances consuming power, with special consideration for prosumers.
Energy Efficiency and Management providers	Offers energy optimization services, reducing costs and grid dependency for customers.
Facility/Building Management	Optimizes energy use and balances supply and demand in various buildings, ensuring regulatory compliance.
City Planning Authorities	Responsible for various urban planning, development, infrastructure, and environmental initiatives.
Technology Providers and Integrators	Offer hardware, software, and integrated solutions for building automation and data management.

Role	Role / Responsibility
Regulatory and Standards Bodies	Set standards and regulations for energy efficiency, building codes, and safety measures.
Research and Development Institutions	Conduct studies and research to advance sustainable building practices and renewable energy solutions.
Insurance Companies and Risk Management	Analyze and provide insurance services for buildings, assessing risks related to energy usage and infrastructure.
Environmental and Sustainability Groups	Advocate for environmental conservation and green initiatives in building and energy sectors.
Maintenance and Service Contractors	Offer maintenance and servicing for building infrastructure and energy systems, ensuring operational efficiency.
Real Estate Developers and Property Owners	Influence the implementation of energy-efficient solutions in new constructions and renovations.
Financial Institutions and Investors	Provide financial support and investment for energy-efficient building projects and infrastructure development.
Smart Cities Initiatives and Municipalities	Develop and implement smart city initiatives integrating technology for energy management and sustainable urban development.
Consultants and Energy Advisors	Provide expertise in energy management, sustainability, and regulatory compliance for stakeholders.
Supply Chain and Material Suppliers	Offer building materials, equipment, and technology for constructing energy-efficient and sustainable buildings.
End-Users and Tenants	Individuals or organizations occupying buildings, influencing demand patterns and adherence to energy-efficient practices.

A5 SINA Proof of Concept

Buildings energy data collections

In this proof of concept, a variety of network connection approaches will be investigated. Therefore, the focus in the first stage was enabling bidirectional IP-based communication channels to supply energy data in real time from a certain building such as GEE Lab qualified in this scenario to be an energy producer and consumer (prosumer) at the same time toward one (virtual) house qualified to be only an energy consumer.

In the proof-of-concept, a One-to-One communication model will be conceptualized, implemented, and tested. This initial approach can be extended in future works to enable a One-to-Many communication model which remains the appropriate and most realistic model reflecting the energy associations between an energy producer building and all other surrounding buildings and/or PV charging stations participating in the local energy network.

In the second stage, the WP4 will research and integrate an adequate software environment enabling an instant processing of smart contracts between the participants in the local energy network as well as an immutable persistence of the relevant energy transaction and billing data. At this stage, the integration of DLT and BCT technologies will be evaluated.

In the third and last stage, the whole P2P demo software framework will be integrated to simulate and test the well-known energy data communication scenarios between the different participants connected to the local energy network.

Electrical energy consumptions

Nowadays, the usage of electrical energy in buildings is indispensable. Almost all building apparatuses are interconnected with a certain in-house electrical infrastructure and consume considerable amounts of energy.

The SFOE analyzed in a study published at the end of 2022 the progression of energy consumptions in Swiss households during the last decade (2000 – 2021) and came to the result that despite the remarkable positive reduction (-5.7 %) of thermal energy consumptions in Swiss households, the usage of electrical energy is evermore continuing to augment.

The author of the study found out that the total electrical consumptions between 2000 and 2021 have risen by +21.5%. Figure 41 shows the partition of electrical energy consumptions in Swiss households in 2021. More than a quarter (25.1%) of the electrical energy consumptions in the related year are deployed for room heating and 13.4% are going in water heating. A further 14% are consumed in home electrical cooking and dishwashing devices. 8.2% of the total electrical energy consumptions fall to the costs of cooling and freezing devices and another 8.1% are used for entertainment, information, and communication purposes. About 19 % of the total electrical energy sum is consumed in building lighting (6.5%) systems, washing and drying (6.3%) engines as well as air conditioning, ventilation, and automation infrastructures (6.3%)

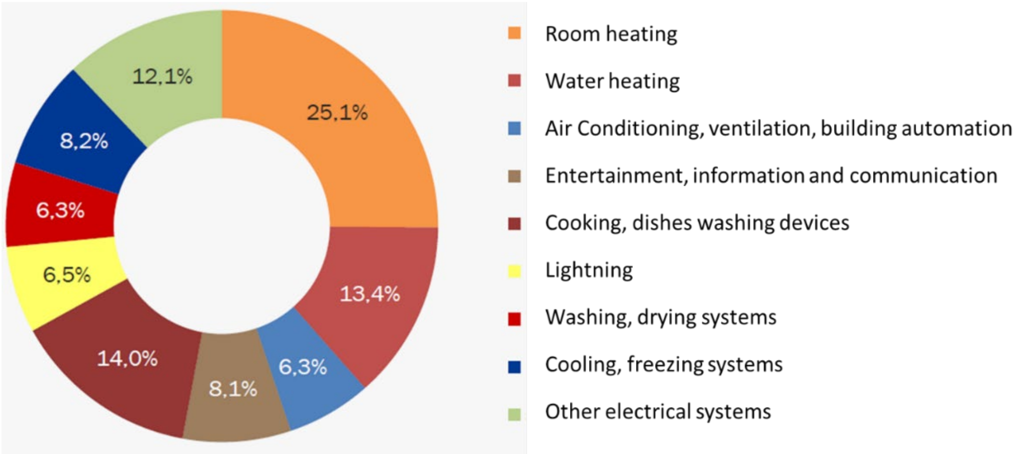

Figure 41 Partition of electrical energy consumptions in Swiss households in 2021

To anticipate expected energy peak demands in the future, new methods and IT solutions are needed to previously estimate energy loads and consumption behaviors of the households e.g., a month, a week, or just a day before. Performing Week-, Day-ahead energy estimations or forecasting requires integrating an advanced P2P energy data exchange and analytics framework.

In the first layer of the mentioned framework, adequate data communication drivers and data storage systems are required to continuously collect, and log energy consumptions of the different home energy consumer systems listed i.e., in Figure 41.

In the second layer of the target framework, configurable task scheduling tools must be designed to create and automate many jobs. These jobs must be linked to the target energy management instances that must be run on the third layer, namely the application layer.

The application layer provides the required intelligent services to analyze, learn from the historical energy consumption log files served by the first layer, and perform e.g. a Day-ahead (d-1) energy loads forecasting.

Figure 42 illustrates a d-1 load forecast performed using EMHAS - Energy Management for Home Assistant, an open-source software designed for residential households to optimize energy usage while considering factors such as the household's actual electrical energy needs and power generation from solar panels.

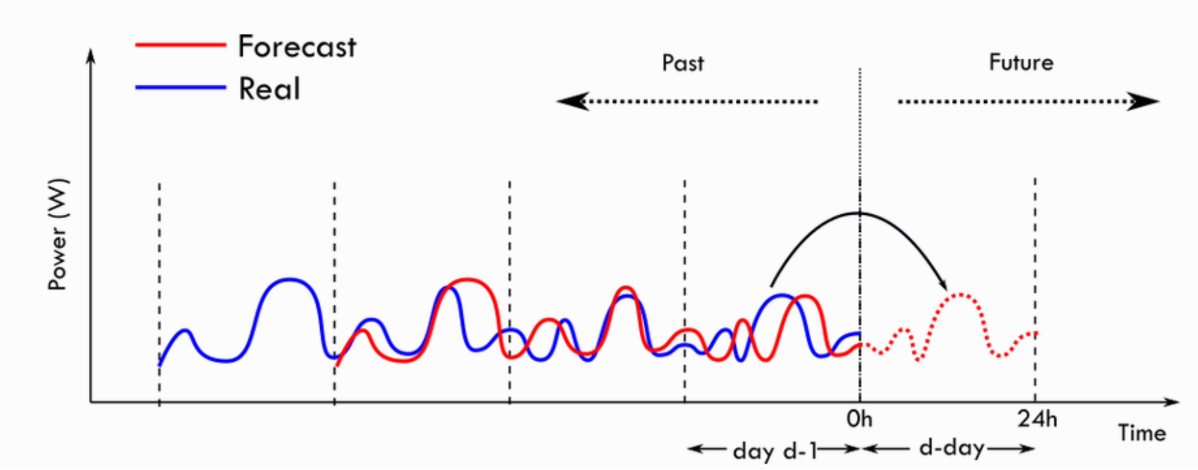

Figure 42 Example Day-ahead load forecast

An accurate forecasting of electrical energy demands of a certain house also relies on other data inputs including historical and short-term forecasts for weather data.

Electrical energy productions

The availability of real historical data related to PV energy productions in a prosumer household is decisive for estimating whether the amount of energy generated in the short term is sufficient for own energy needs or an electrical power extra must be purchased.

In the cases where own PV production data are regularly recorded, the target framework can be equipped on the first layer with required interfaces to the installed PV systems or their smart meters to enable an inline interception and logging of related instant PV energy production data.

Alternatively, the second layer of the target framework should be programmed to run periodically specific jobs sending requests to an online paid cloud service like Solcast data API and receive weather as well as power production data forecasts based on the geographic position of the related building. The next graph shows live data and forecast of expected half-hour PV productions at GEE Live position. Figure 43 was generated using the online Solcast toolkit. The related data objects are also delivered as files in JSON and CSV formats.

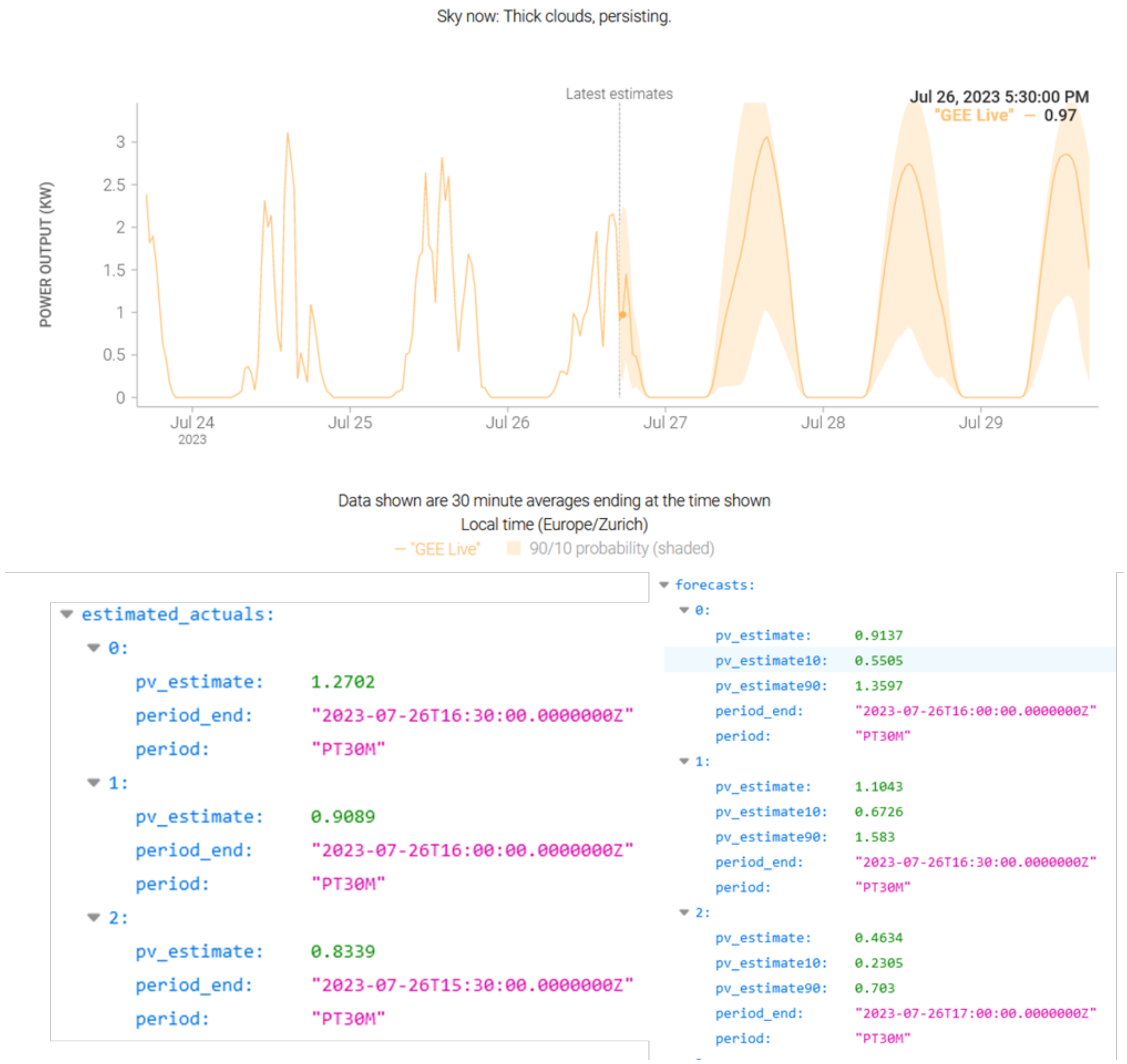

Figure 43 Example PV production live data and forecasts served by Solcast API

The Solcast toolkit can provide PV production forecasts for a certain location with a minimum rate of 5 min. However, in our current P2P scenario, triggering a forecast every 15 min may be sufficient.

Results and findings of the PoC

GEE energy data aggregation, pre-processing, and persisting

All GEE energy data relevant for proving the proposed P2P concept are sent regularly to an external MQTT broker running in HiveMQ cloud. Our target P2P software framework, which is subscribing to related MQTT topics, intercepts all transmitted messages, parses, pre-processes, and persists the relates energy values in the mongoDB instance running locally. Figure 44 illustrates energy dataflows from GEE live to the target P2P software framework.

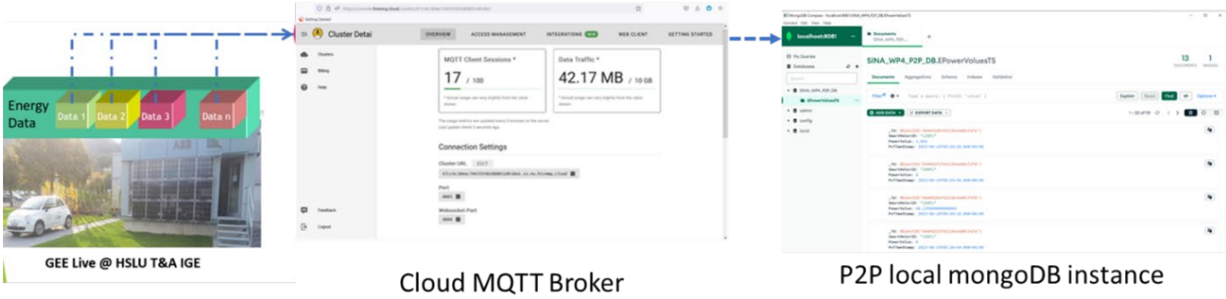

Figure 44 Energy dataflows from GEE to target P2P software framework

Each energy data entry in the database has a unique timestamp, a unique identifier (_id), a power value as well as a unique smart meter ID identifying the smart meter which is linked physically to the related GEE consumer/producer system.

GEE energy data visualization and interpretation

The P2P backend system provides different APIs to access GEE power values stored within the mongoDB instance. These APIs can be requested per HTTP to get an overview about the actual energy production and consumption behaviors at GEE.

The P2P frontend provides several services to request the related backend APIs and monitors GEE energy behaviors. Figure 45 presents an extract of the P2P APIs served by the backend system and the P2P APPs deployed at the frontend to give in-depths about GEE Live energy production and consumption behaviors.

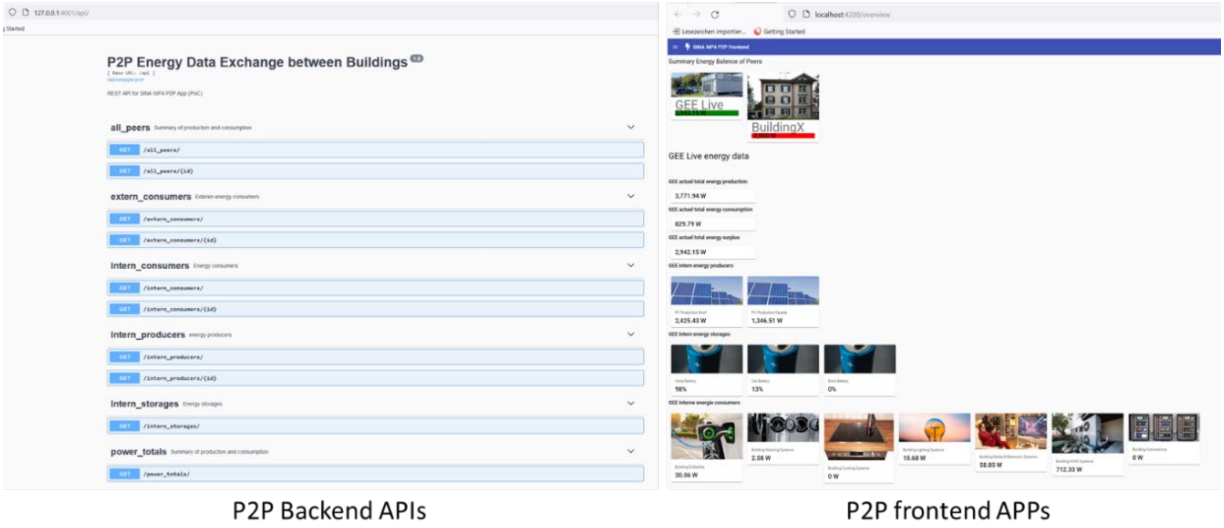

Figure 45 P2P APIs and APPs

The P2P backend system assesses in an accurate rate (at time every 300s) in background whether the total energy produced at GEE exceeds the total energy demands and calculates the actual energy plus or minus.

Once a sufficient energy surplus is noted, the backend system triggers an event to transmit a broadcast message to the neighbor nodes in local P2P energy network to inform them that an energy extra is there. The message includes the amount of energy surplus and a unique address to the prosumer node issuing the message. Before sending the broadcast selling message, a connection to the already initiated P2P energy network must be built. Next screenshots demonstrate an immediate data exchange between the GEE Live and a virtual neighbor node running on the local P2P energy network 127.0.0.1 (Figure 46).

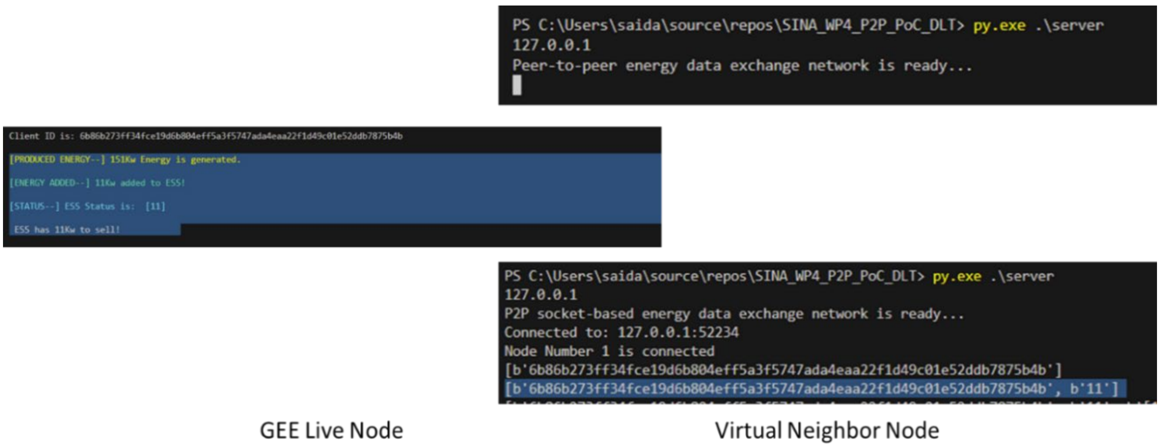

Figure 46 Broadcast energy selling message from GEE to a virtual neighbor node

Deploying smart contract in a P2P-DLT/BLC-Test-Network

Creating Fabric-Test-Network

The deployment of smart contracts requires setup up and operating an extra DLT/BLC-infrastructure. The important and complex DLT/BLC network administration jobs must be automatically processed every time an energy (asset) transfer must be accomplished between an individual prosumer node e.g., GEE Live and another node inside the local P2P energy network.

Developing and managing an extensive DLT/BLC-infrastructure is not in the scope of this research project, therefore only a Hyperleder-Fabric-Test-Network have been integrated and programmed for deploying an example smart contract as describe in (Hyperledger, 2023a).

The Fabric-Test-Network is deployed using a bash script which processes all tasks needed to configure and start the peer organizations (prosumer and consumer nodes) participating in the local P2P network.

The bash script composes and runs 3 docker containers, representing 2 peer organizations and an order organization in the Test-network. The script creates for each composed organization or node participating in the Test-network its own identities and authority certificates before adding related organizations to the network.

Once the two peers and order organization are running successfully on Test-Network, the script creates a Fabric channel for transactions between the first organization and the second organization which is further detailed in (Hyperledger, 2023b). Channels are a private layer of communication between specific network members. Channels can be used only by organizations that are invited to the channel and are invisible to other members of the Test-Network. Each channel has a separate blockchain ledger. Organizations that have been invited «join» their peers to the channel to store the channel ledger and validate the transactions on the channel.

As shown in next Figure 47, the Test-network creation is only succussing, when the new created channel is joined by both organizations, the GEE live prosumer node as well as the neighbor virtual consumer node.

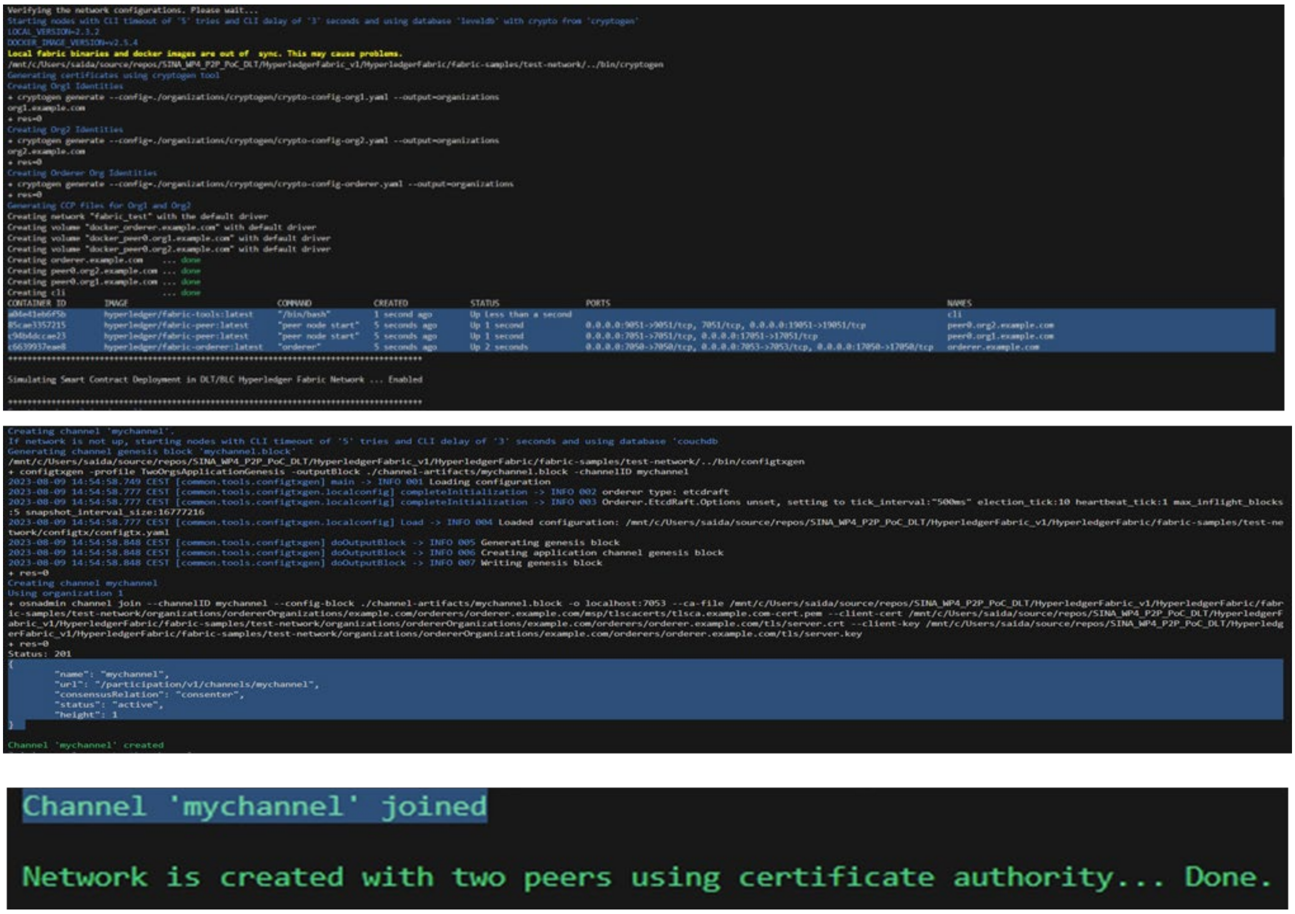

Figure 47 Successful creation of a Fabric-Test-Network

Joining an extra peer organization to the Fabric-Test-Network is at any time possible. For this, the bash script generates for the Org3 its own identities, authority certificates and updates the already created channel to enable the Org3 to store the channel ledger and validate the transactions on related channel Figure 48 and Figure 49 are showing all successfully created software instances inside the Fabric-Test-Network.

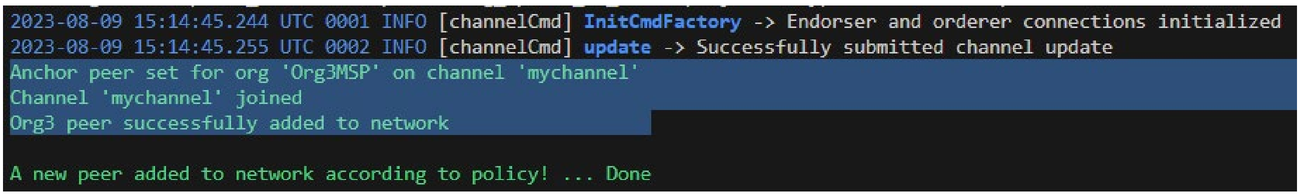

Figure 48 Joining an extra organization to the Fabric-Test-Network

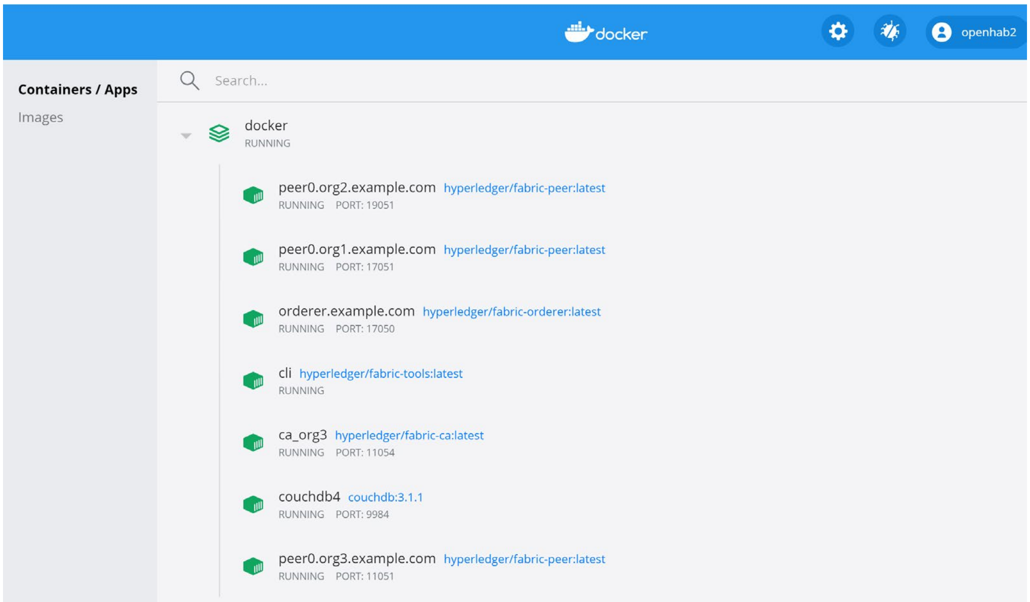

Figure 49 Fabric-Test-Network software instances

Deploying smart contract

While the integration of a Fabric-Test-network was successful, the deployment of a sample smart contract (chaincode) programed in the language go failed. The bash terminal is reporting as shown in Figure 50 an error massage that the compiler «go.exe» was not found in the shell environ path, even though last go version «1.20.5» as well all other dependencies are correct installed and added to the environment path.

```
Deploying main smart contract for registration...

deploying chaincode on channel 'mychannel'
executing with the following
- CHANNEL_NAME: mychannel
- CC_NAME: basic
- CC_SRC_PATH: ../../mychaincode/
- CC_SRC_LANGUAGE: go
- CC_VERSION: 1.0
- CC_SEQUENCE: 1
- CC_END_POLICY: NA
- CC_COLL_CONFIG: NA
- CC_INIT_FCN: NA
- DELAY: 3
- MAX_RETRY: 5
- VERBOSE: false
Vendoring Go dependencies at ../../mychaincode/
/mnt/c/Users/saida/source/repos/SINA_WP4_P2P_POC_DLT/HyperledgerFabric_v1/HyperledgerFabric/mychaincode /mnt/c/Users/saida/s
rk
scripts/deployCC.sh: line 59: go: command not found
/mnt/c/Users/saida/source/repos/SINA_WP4_P2P_POC_DLT/HyperledgerFabric_v1/HyperledgerFabric/fabric-samples/test-network
Finished vendoring Go dependencies
+ peer lifecycle chaincode package basic.tar.gz --path ../../mychaincode/ --lang golang --label basic_1.0
+ res=1
Error: failed to normalize chaincode path: failed to determine module root: exec: "go": executable file not found in $PATH
Chaincode packaging has failed
Deploying chaincode failed
```

Figure 50 Deploying a sample smart contract

The deployment of related chaincode failed during the execution of the operation «package-ChainCode()» (Figure 50).This first operation and all other successive operations (install chaincode on the peer organizations 1 and 2, querying whether the chaincode is installed, etc.) are parts of an advanced bash script «deployCC.sh».

Extensive analysis and debugging of the errors occurred during running the chaincode deployment script have been not pursued, because this requires not only an expertise in shell and go programing but also in-depth know-how about all fabric chain code lifecycle transitions from the chaincode packaging to the commitment of the chaincode to the owner channel.

Next diagram published at Hyperbaric link (Hyperledger, 2023c) illustrates the lifecycle of a new defined chaincode (named MYCC) which is committed to the Channel (Figure 51).

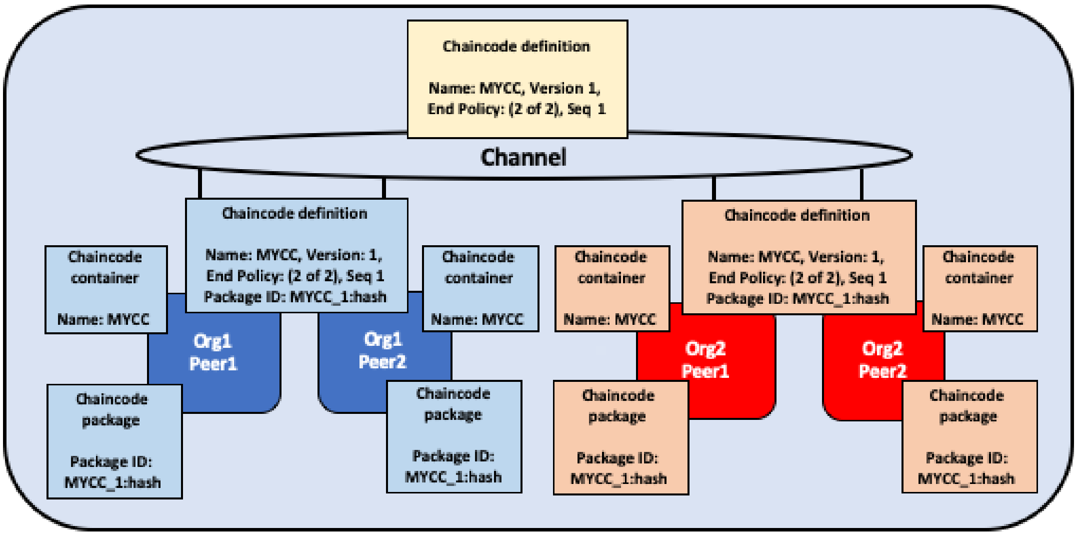

Figure 51 Hyperlager Fabric Chaincode Lifecycle

At the end of such a complex lifecycle each peer organization may then create and run their own new chaincode execution environment inside a unique docker container and start working with the new chaincode program.

Following these approaches our already created P2P-DLT/BLC-Test-Network may be extended in case of a successful chaincode deployment with 3 additional chaincode deployment containers. Each of these virtual instances is associated to a single peer organization within the network. Final P2P-DLT/BLC-Test-Network may consist of 10 docker containers running on the same P2P network 127.0.0.1.

A6 Organizational structure and governance

For a governance concept in the Data Space Project, it is important to address inter-organizational and intra-organizational aspects (IDSA, 2021, p. 6). In the interest of focus, we decided to only discuss certain aspects of intra-organizational governance and put the focus on inter-organizational governance.

Central to the governance framework of SINA is a dedicated commitment to cultivating impactful inter-organizational collaboration. Acknowledging the vitality of shared resources and collective endeavors within data spaces, our governance model prioritizes the implementation of well-defined standardized procedures, and inclusive decision-making processes. Consequently, the establishment of an ecosystem for data spaces becomes imperative, necessitating the selection of a legal structure aligned with these essential requirements.

Legal structure

Selecting the appropriate legal structure for data space project like SINA involves a thorough examination of various options, the following criteria should be considered in the decision:

- Capital
- Risk/liability
- Independence
- Taxes
- Social security
- Duration

The association's purpose cannot be profit-oriented in accordance with Article 60 of the Swiss Civil Code (Picht & Studen, 2022). SINA's objectives include promoting collaboration, ensuring data privacy, facilitating data analysis, and enabling the development of innovative applications. In essence, SINA aims to enhance data management within organizations, emphasizing that its primary purpose is not profit-driven. Another reason the association is a fitting legal structure for SINA are that the activities and the project will last for a while and is not a short-term project, therefore an association is more suitable than a simple partnership according to Art. 530 CO (Schweizerische Eidgenossenschaft, 1911). Another difference between a simple partnership and an association is that the association is a legal person, which means members can join and leave.[26] In SINA we want that new members can join and the members should not be set on the point of time the legal structure is founded.

The supreme organ in an association is the general meeting of members (following «the general meeting»), this concludes that the members are in charge. For the founders, this means that they may lose their influence on further development when new members join - even if they are on the board. This is because the board is the executive body, which is subject to the mandatory supervision of the general meeting.[27] As a legal entity, an association can have assets and debts, it can even inherit and can

[26] Art. 70 para. 1 SCC

[27] Art. 64 SCC

independently sue and be sued in court. It is one of the most important advantages of this legal form, that only the association's assets are liable for its debts.[28]

The downside is that potential contractual partners are aware of the limited liability provided by association law. It is quite possible that a newly established association may need to provide upfront payments or other assurances for substantial contractual considerations.

Therefore, the association is a suitable legal structure for data spaces, considering its alignment with ideal purposes, long-term orientation, openness to members, democratic decision-making, and crucial liability limitations. While acknowledging the diverse landscape of legal entities, the association stands out as a robust and community-oriented foundation for the development and maintenance of SINA.

Founding of the association

The successful founding of the SINA Association necessitates an organized constituent founders' meeting.[29] It is imperative that this meeting is convened well in advance, assuming that all preceding steps for association formation have been undertaken. Primarily among these requirements is the preparation of written statutes, a critical element that requires approval during the founders meeting.

Key Steps in the Constituent Founders' Meeting

1. Statutes Approval:

The statutes of the association must be available in written form for scrutiny during the meeting.[30]

Approval of these statutes is a central agenda item, as it lays the foundation for the legal framework of the SINA association.[31]

2. Organ Election:

Following the approval of statutes, the election of association organs specified within the statutes is conducted.[32]

Subsequent elections may include offices such as president, treasurer, etc., depending on the association's organizational structure.

3. Documentation Requirements:

Minutes of the founders' meeting are obligatory for legal validity.[33]

The minutes must document:

---

28 Art. 75 SCC
29 Art. 60 ff SCC
30 Art. 60 para. 1 SCC
31 Art. 60 para. 1 SCC
32 Art. 61 SCC
33 Art. 60 para. 2 SCC

- Appointment of the chairman of the founders' meeting and the keeper of the minutes.
- Naming of founding members (which can be detailed in an attached attendance list).
- Resolution on association foundation in accordance with 60ff ZGB, outlining the association's purpose and name.
- Approval of the association's statutes.
- Election of the board of directors, auditors, and their acceptance of the election.

4. Signature and Compliance:

The minutes, integral to the legal foundation of the association, must be signed by hand by both the chairman and the recording secretary.

Strict adherence to these documentation and signature procedures is paramount for legal compliance.

This foundational process lays the groundwork for the association's governance structure and sets the stage for its future endeavors in the SINA data space domain.

Members of the SINA association

The important inquiry of membership criteria and influential roles within the SINA association must be addressed. As emphasized earlier, the general meeting, comprised of SINA association members, wields significant power.[34] Therefore, the selection of association members demands careful consideration. It is essential to examine stakeholders in SINA and assess various business roles to determine eligibility and potential influence within the association.

Core participants in data spaces, such as data providers and users, are not deemed suitable for membership in the SINA association one reason being their numbers, which would pose a significant barrier. Therefore, general participants are not considered appropriate candidates for membership.

Intermediaries and, in particular, operators of platforms are important in ensuring the success of facilitating the expansion of the data economy and facilitating the data exchange. Given their pivotal role, it is of advantage that these entities actively engage in implementing governance instruments and fulfilling other essential responsibilities. Therefore, it should be considered if they should be members of the association. Central are other entities for membership, including industrial companies, technology providers, research institutions, and government entities. These participants contribute valuable knowledge and influence, enhancing the overall diversity and expertise within the association.

While not all participants in the SINA data space ecosystem are members of the association, they are still required to adhere to the general terms and conditions, as well as the code of conduct, playing by the established rules of the association. The significance of these ground rules lies in fostering trust among participants, a critical element within the realm of data spaces. Having a set framework ensures

[34] Art. 64 para. 1 SCC

a common understanding and adherence to ethical and operational standards, thereby strengthening the foundation of trust among diverse stakeholders in the ecosystem.

Governance model

The general meeting serves as the ultimate decision-making authority for the Data Network in SINA.[35] Its primary role is to foster cooperation among the participating parties, oversee the strategic administration of the Network and to enhance trust among all participants.

The general meeting is intended for the following tasks:[36]

- Decisions on all matters that are not delegated to the Executive or other bodies (so-called association resolutions).
- Admission and exclusion of members.
- Election of the Executive Board and any other bodies.
- Supervision of the activities of all bodies.
- Dismissal of the Executive Board and other bodies at any time.

Key functions of the SINA general meeting are (Pitkänen & Luoma-Kyyny, 2022):

- *Coordinating and Decision-Making*: The general meeting ensures coordination and makes decisions concerning the Data Network's business, legal, technical, and ethical aspects. It is tasked with proposing changes to maintain the Network's purpose and compliance with relevant requirements. The general meeting has the authority to propose changes to the Statutes and its Appendices. It can also approve new members, data sets, and data set terms of use as necessary.
- *Composition, Meetings, and Organization*: The members of the SINA association appoint the Executive Board. The General Meeting also selects a chairperson and a secretary. The Chair leads meetings and can delegate this responsibility, while the Secretary handles administrative tasks. Representatives are expected to actively participate in meetings and may appoint substitutes or proxies when needed. Regular meetings are convened at least once every three months, with extraordinary meetings called upon written request. Virtual meetings are permissible, but an annual face-to-face meeting is mandatory. The Secretary's responsibilities include meeting preparation, document management, and day-to-day coordination, but the Secretary cannot make binding declarations without explicit authorization.
- *Meeting Agenda*: Meetings follow a standard agenda that includes introductions, review of previous meeting minutes, ongoing matters (such as agreement changes and new memberships, accession agreements), operational updates, and consideration of relevant items.
- *Quorum and Decisions*: A quorum is established when the Chair or their representative and at least two-thirds of the General Meeting or their representatives are present. The General Meeting aims to reach a consensus but may vote on decisions, when necessary, with the Chair holding the casting vote. Decisions are made by a majority of two-thirds or half of the Representatives

[35] Art. 64 par. 1 SCC

[36] Art. 65 SCC

present, depending on the issue. Certain decisions, including amendments to the Statutes, require a two-thirds majority of all Members.

– *Admission of New Parties*: New parties can join the ecosystem with approval from the executive board, which may require a qualified majority, including specific Data Providers. If a decision materially affects a Party's rights or obligations and is objected to, that Party can terminate the Constitutive Agreement within thirty days.

– *Invited Attendees*: The General Meeting can invite individuals to attend meetings, subject to the Chair's approval. Non-Network Member invitees may need to sign a non-disclosure agreement, unless waived by the Chair, to ensure confidentiality. The Chair is responsible for verifying the invitee's commitment to confidentiality.

A7 Statutes of SINA Association

Statuten
des Vereins «SINA Association»

1. Name und Sitz

1.1 Unter dem Namen «SINA Association» besteht ein Verein, der den vorliegenden Statuten und den Vorschriften in Art. 60 ff. ZGB untersteht. Der Verein soll in das Handels-register eingetragen werden.

1.2 Sitz des Vereins ist in Horw, Luzern.

2. Zweck

2.1 Der Verein bezweckt ausschliesslich und unmittelbar gemeinnützige Zwecke. Sein Zweck liegt in der Förderung von Wissenschaft und Forschung auf dem Gebiet der Erarbeitung und Schaffung sicherer Architekturen für den künftigen Austausch und die Bearbeitung industrieller Daten, insbesondere in der Schweiz und Europa

2.2 Durch folgende Massnahmen soll der Vereinszweck unter anderem realisiert werden:

a. Zusammenführung und Festlegung von Nutzeranforderungen an die Struktur eines Data Space.
b. Durchführung von Informationsveranstaltungen, insbesondere durch zeitnahe Veröffentlichung der erlangten Erkenntnisse.
c. Teilnahme des Vereins an der Entwicklung von Leitlinien und Gesetzgebungsprozessen, die für den Vereinszweck sowohl auf nationaler als auch auf internationaler Ebene von Bedeutung sind.
d. Anstoss zur Umsetzung von Förderungsprojekten und Forschungsprojekten.

2.3 Einer Umwandlung des Vereinszwecks müssen alle Vereinsmitglieder zustimmen.

3. Mitglieder

3.1 Mitglieder des Vereins können juristische Personen, Stiftungen, Behörden sowie Personengesellschaften sein, welche den Zweck des Vereins anerkennen und fördern.

3.2 Der Vorstand entscheidet über die Aufnahme von Mitgliedern nach schriftlich eingereichtem Aufnahmegesuch. Der Entscheid des Vorstandes ist endgültig. Ein ablehnender Entscheid muss nicht begründet werden.

3.3 Die Mitgliedschaft an dem Verein beschränkt die Parteien nicht daran, sich an anderen Datennetzen, Plattformen, Ökosystemen oder sonstigen Kooperationen zu beteiligen oder von Dritten bereitgestellte Dienste zu nutzen. Darüber hinaus hindert oder beschränkt die gemeinsame Nutzung von Daten innerhalb des Netzes den jeweiligen Datenanbieter nicht daran, diese Daten nach eigenem Ermessen mit Dritten zu teilen.

3.4 Andere Teilnehmende (sogenannte Associates), am SINA ecosystem, namentlich Data Provider, Service Provider sind nicht zwingend Mitglieder des Vereins müssen jedoch die Dokumente im Anhang akzeptieren.

4. Mitgliedschaft, Rechte und Pflichten

4.1 Alle Mitglieder verpflichten sich, den Verein – im Rahmen ihrer rechtlichen und tat-sächlichen Möglichkeiten – bei der Verwirklichung seines Zwecks zu unterstützen.

4.2 Sie verpflichten sich die Mitgliedbeiträge fristgerecht zu bezahlen.

4.3 Jedes Mitglied ist berechtigt, an Abstimmungen teilzunehmen und Anträge an die Vereinsorgane zu richten. Sowohl das aktive als auch das passive Wahlrecht stehen den Mitgliedern zu, wobei die Ausübung durch die in Abschnitt 5 genannte Person er-folgt. Das Stimmrecht ruht, solange das Mitglied mit der Beitragszahlung im Rück-stand ist.

5. Mitgliederbeitrag

5.1 Der Mitgliederbeitrag wird von der Vereinsversammlung jährlich festgelegt. Er beträgt maximal CHF xx pro Jahr.

5.2 Mitglieder haben für das Kalenderjahr, in welchem ihre Aufnahme erfolgt bzw. ihre Mitgliedschaft erlischt, den anteilsmässigen Mitgliederbeitrag zu entrichten.

6. Erlöschen der Mitgliedschaft

6.1 Erlöschensgründe

Die Mitgliedschaft erlischt durch

a) Austritt;

b) Ausschluss;

c) Verlust der Rechtsfähigkeit bei juristischen Personen.

6.2 Austritt

Der Austritt kann unter Einhaltung einer Frist von 3 Monaten auf das Ende des Kalenderjahrs schriftlich gegenüber dem Vorstand erklärt werden.

6.3 Ausschluss

6.3.1 Der Vorstand kann ein Mitglied ohne Angabe von Gründen vom Verein ausschliessen. Der Ausschluss erfolgt nur nach Anhörung des Mitgliedes und wird diesem schriftlich erklärt. Der Ausschluss gilt per sofort.

6.3.2 Der Ausschluss ist endgültig. Die Möglichkeit eines Rekurses an die Vereinsversammlung besteht nicht.

6.4 Verlust der Rechtsfähigkeit bei juristischen Personen

Die Mitgliedschaft ist nicht rechtsgeschäftlich übertragbar.

7. Organisation des Vereins

7.1 Organe

Die Organe des Vereins sind:

a. die Vereinsversammlung;
b. der Vorstand;
c. die Revisionsstelle
d. Geschäftsführung, wenn bestellt.

7.2 Vereinsversammlung

7.2.1 Oberstes Organ des Vereins ist die Vereinsversammlung. Ihr stehen folgende Befugnisse zu:

a. Genehmigung des Protokolls der letzten Vereinsversammlung;
b. Abnahme des Jahresberichts, der Jahresrechnung, des Jahresbudgets und des Berichts der Revisionsstelle;
c. Entlastung des Vorstandes und der Revisionsstelle
d. Festsetzung der Mitgliederbeiträge und des Jahresbudgets;
e. Wahl und Abberufung des Vorstandes und der Revisionsstelle;
f. Behandlung von Anträgen des Vorstandes und der Mitglieder;
g. Änderung der Statuten;
h. Auflösung des Vereins;
i. Beschlussfassung über die Gegenstände, die der Vereinsversammlung durch das Gesetz oder die Statuten vorbehalten ist.

7.2.2 Die ordentliche Vereinsversammlung findet innerhalb der ersten 6 Monate eines Kalenderjahres statt. Die Einladung erfolgt mindestens 30 Tage im Voraus schriftlich oder per E-Mail durch den Vorstand und enthält die Traktanden, die Anträge des Vor-standes sowie den Jahresbericht, die Jahresrechnung und den Bericht der Revisionsstelle.

7.2.3 Anträge von Mitgliedern zuhanden der Vereinsversammlung sind schriftlich und spätestens bis 31. Januar eines Kalenderjahres an den Vorstand zu richten. Der Vor-stand ergänzt die Traktandenliste um die fristgerecht eingegangenen Anträge.

7.2.4 Eine ausserordentliche Vereinsversammlung wird auf Beschluss des Vorstandes, auf Antrag mit schriftlicher Begründung von mindestens 1/5 der stimmberechtigten Mitglieder oder auf Antrag der Revisionsstelle einberufen. Die Einladung erfolgt mindestens 20 Tage vor der Versammlung.

7.2.5 Den Vorsitz der Vereinsversammlung führt der Präsident, bei dessen Verhinderung der Vize-präsident des Vorstandes oder ein anderer von der Vereinsversammlung gewählter Tagespräsident. Der Vorsitzende bezeichnet einen Protokollführer und 2 stimmberechtigte Mitglieder für die Ermittlung von Abstimmungs- und Wahlergebnis-sen.

7.2.6 Über die Beschlüsse der Vereinsversammlung ist ein Protokoll zu führen, das vom Vorsitzenden und vom Protokollführer unterzeichnet wird. Die Mitglieder sind berechtigt, das Protokoll einzusehen.

7.2.7 Abstimmungen und Wahlen finden offen oder auf Beschluss der Vereinsversammlung schriftlich statt.

7.2.8 Jedes Vereinsmitglied hat eine Stimme und kann sich mittels schriftlicher Vollmacht durch eine Drittperson vertreten lassen.

7.2.9 Die Vereinsversammlung fasst ihre Beschlüsse und vollzieht ihre Wahlen mit der absoluten Mehrheit der anwesenden Mitglieder, sofern nicht eine zwingende Vorschrift des Gesetzes oder die Statuten etwas anderes bestimmen. Bei Stimmengleichheit hat der Vorsitzende den Stichentscheid. Kommt bei Wahlen im ersten Wahlgang die Wahl nicht zustande, findet ein zweiter Wahlgang statt, in welchem das relative Mehr entscheidet.

7.3 Vorstand

7.3.1 Der Vorstand besteht aus 7 - 16 Mitgliedern. Sie werden von der Vereinsversammlung für die Amtsdauer von 2 Jahren gewählt. Wiederwahl ist zulässig. Eine Abberufung ist jederzeit und fristlos möglich.

7.3.2 Die Vereinsversammlung wählt den Präsidenten. Im Übrigen konstituiert sich der Vorstand selbst und bestimmt die Zeichnungsberechtigung. Grundsätzlich gilt Kollektivunterschrift. Der Vorstand besteht aus dem Präsidenten, Vizepräsidenten, Aktuar und Kassier. Ämterkumulation ist nicht zulässig.

7.3.3 Dem Vorstand obliegt die Leitung und Vertretung des Vereins. Er kann in allen Angelegenheiten Beschluss fassen, die nicht nach dem Gesetz oder den Statuten der Vereinsversammlung zugeteilt sind. Es sind dies insbesondere:

a. Führung der laufenden Geschäfte und Organisation des Vereins;
b. Vorbereitung und Durchführung der Vereinsversammlungen;
c. Aufnahme und Ausschluss von Mitgliedern;
d. Buchführung.

7.3.4 Der Vorstand wird auf Antrag des Präsidenten oder auf Verlangen eines Vorstandsmitgliedes einberufen. Er ist beschlussfähig, wenn die Mehrheit der Mitglieder anwesend ist. Die Sitzungen sind zu protokollieren.

7.3.5 Jedes Vorstandsmitglied hat eine Stimme. Beschlüsse erfolgen mit dem einfachen Mehr der Anwesenden. Bei Stimmengleichheit hat der Präsident den Stichentscheid.

7.4 Revisionsstelle

7.4.1 Die Vereinsversammlung kann eine oder mehrere natürliche oder juristische Personen, welche nicht Mitglied des Vereins sein müssen, als Revisionsstelle für die Dauer von einem Amtsjahr wählen. Das Amt endet mit der Abnahme der letzten Jahres-rechnung. Eine Wiederwahl ist zulässig. Eine Abberufung ist jederzeit und fristlos möglich.

7.4.2 Das Geschäftsjahr fällt mit dem Kalenderjahr zusammen. Das erste Geschäftsjahr dauert vom Gründungsdatum bis zum Ende des laufenden Kalenderjahres. Auf den 31. Dezember wird die Jahresrechnung abgeschlossen und ein Inventar erstellt. Die Jahresrechnung wird von der Revisionsstelle geprüft.

7.4.3 Die Revisionsstelle erstattet der ordentlichen Vereinsversammlung schriftlichen Be-richt über die Prüfung der Jahresrechnung und stellt Antrag auf Erteilung oder Ver-weigerung der Décharge gegenüber Kassier und Vorstand.

7.5 Geschäftsführung

7.5.1 Der Vorstand kann einen oder mehrere Geschäftsführende zur Führung der laufen-den Geschäfte bestellen und die Tätigkeit der Geschäftsführung durch eine Geschäftsordnung näher regeln.

7.5.2 Die Geschäftsführende nehmen an den Sitzungen des Vorstands teil und können sich beratend einbringen.

8. Vereinsvermögen, Haftung und Nachschusspflicht

8.1 Das Vermögen des Vereins setzt sich aus den Mitgliederbeiträgen, Überschüssen der Betriebsrechnung, allfälligen Schenkungen, Veranstaltungsbeiträgen und Vermächtnissen sowie Sponsoring zusammen.

8.2 Für die Verbindlichkeiten des Vereins haftet ausschliesslich das Vereinsvermögen. Eine persönliche Haftung und Nachschusspflicht der Vereinsmitglieder ist ausgeschlossen.

9. Statutenänderungen und Auflösung

9.1 Statutenänderungen und die Auflösung des Vereins erfordern die Anwesenheit von mindestens zwei Dritteln aller Mitglieder sowie die absolute Mehrheit der abgegebenen Stimmen.

9.2 Wird eines der Quoren nicht erreicht, ist eine zweite Vereinsversammlung mit den gleichen Traktanden innerhalb von 6 Wochen einzuberufen. Diese Versammlung ist ohne Rücksicht auf die Zahl der anwesenden Mitglieder beschlussfähig.

9.3 Im Falle der Auflösung bestimmt die Vereinsversammlung über die Verwendung des Liquidationserlöses.

10. Inkrafttreten der Statuten

Diese Statuten wurden an der Gründerversammlung vom DD-MM-YYYY genehmigt und treten sofort in Kraft.

____

Ort und Datum

______________________________ ______________________________

[Unterschrift Gründerpräsident:in] [Unterschrift Protokollführer:in]