State of the art and technology inventory 2

06/11/2023

Grant Agreement No.: 101093274  
Call: HORIZON-CL4-2022-HUMAN-01  
Topic: HORIZON-CL4-2022-HUMAN-01-03  
Type of action: RIA

D3.2 State of the art and technology inventory 2

Report-2

|  |  |
| --- | --- |
| Work package | WP 3 |
| Task | Task 3.1 |
| Due date | 28/02/2024 (M13) |
| Submission date | 27/02/2024 |
| Deliverable lead | ALA |
| Version | 0.3 |
| Authors | Alexander Herranz (ALA)  Pablo Vela (ALA) |
| Other contributors | Veniamin Boiarkin (ICS) |
| Reviewers |  |
| Abstract | Analysis of the current state of the art, as input for the open call specifications and mentoring support. |
| Keywords |  |

Document Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Description of change** | **List of contributor(s)** |
| **0.1** | **06/11/2023** | **Initial deliverable structure** | **Alexander Herranz (ALA)**  **Pablo Vela (ALA)** |
| 0.2 | 21/12/2023 | Structure rediscussed | Pablo Vela (ALA) |
| 0.3 | 03/01/2024 | More information added | Pablo Vela (ALA) |
| 0.4 | 24/01/2024 | More information added | Veniamin Boiarkin (ICS) |
| 0.5 | 15/02/2024 | Contributions from partners | XXX |
| 0.8 | 20/02/2024 | Finalization of the report | Pablo Vela (ALA) |

Disclaimer

The information, documentation and figures available in this document are written by the TRUSTCHAIN project’s consortium under EC grant agreement 101093274 and do not necessarily reflect the views of the European Commission. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein. The information in this document is provided “as is” and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. Moreover, it is clearly stated that the TRUSTCHAIN Consortium reserves the right to update, amend or modify any part, section or detail of the document at any point in time without prior information.

The TRUSTCHAIN project is funded by the European Union’s Horizon Europe Research and Innovation programme under grant agreement no. 101093274.

Copyright notice

© 2024 TRUSTCHAIN

This document may contain material that is copyrighted of certain TRUSTCHAIN beneficiaries and may not be reused or adapted without prior permission. All TRUSTCHAIN Consortium partners have agreed to the full publication of this document. The commercial use of any information contained in this document may require a license from the proprietor of that information. Reproduction for non-commercial use is authorised provided the source is acknowledged.

executive summary

Analysis of the current state of the art, as input for the open call specifications and mentoring support.

Table of Contents

[1. Introduction 5](#_Toc740144599)

[2. METODOLOGY 6](#_Toc1931349511)

[3. USE-CASE DRIVEN USER Privacy and data governance: REAL PROBLEMS AND EXISTING APPLICATIONS 6](#_Toc1905471560)

[3.1. Healthcare 7](#_Toc957245638)

[3.2. Finance 10](#_Toc2022299101)

[3.3. RETAIL 13](#_Toc463206162)

[3.4. EDUCATION 16](#_Toc1533020458)

[3.5. ENTERTAINMENT and SOCIAL MEDIA 19](#_Toc748923081)

[3.6. ENERGY AND SUSTAINABILITY 22](#_Toc453105876)

[3.7. Supply Chain Management 25](#_Toc1423311782)

[3.8. Public Sector and Government 28](#_Toc1052259735)

[3.9. Automotive and Transportation 30](#_Toc407917291)

[4. DECENTRALISED USER-CENTRIC USER Privacy and data governance FRAMEWORKS 33](#_Toc1519963858)

[4.1. GDPR: Blockchain-based Solutions for GDPR Compliance 33](#_Toc479902908)

[4.2. CCPA: Decentralized Identity Management using Blockchain 34](#_Toc1763357451)

[4.3. HIPAA: Blockchain-based Secure Storage and Sharing of Medical Records 35](#_Toc1145206692)

[4.4. Privacy by Design: Blockchain-based Self-Sovereign Identity Management 35](#_Toc1179952208)

[4.5. Blockchain-based solutions: Zero-Knowledge Proofs 41](#_Toc2121998567)

[5. DECENTRALIZED DATA MARKETPLACES 44](#_Toc738284967)

[6. privacy-aware data processing (FEDERATED LEARNING AND Privacy Enhancing Techs) 45](#_Toc628104037)

[7. Cross-border data transfer technologies 45](#_Toc820815383)

[8. Data identification, data provenance, data tracking mechanisms 47](#_Toc1612972852)

[9. Data certification/verification methods should be developed to verify the trustworthiness of the data. 50](#_Toc1186963186)

[10. Challenges and Limitations 50](#_Toc963621285)

[11. Regulatory and Compliance Aspects 50](#_Toc681694854)

[12. Future Trends and Innovation Potential 50](#_Toc1151350126)

[13. REFERENCES 50](#_Toc1691950307)

# Introduction

In this era of digital transformation, efficient and secure data management has become a key pillar for both organizations and individual users. This state-of-the-art review focuses on analyzing data governance frameworks, with a particular emphasis on the innovative use of blockchain technology. Blockchain, known for its immutability and decentralization, offers a revolutionary perspective in data protection and management. By examining practical cases in critical sectors like healthcare, finance, and public administration, we aim to understand how blockchain implementation is reshaping data governance strategies. This paper highlights the advantages of blockchain, such as enhanced security and transparency, while also addressing the challenges and limitations inherent in its application. It provides a comprehensive view of the potential and obstacles faced in integrating this technology into complex data management systems. This thorough analysis aims to lay a solid foundation for future research and development in the field, emphasizing the importance of technological innovation in protecting user privacy and efficiency in data governance.

# METODOLOGY

This research adopts a mixed-methodological approach, combining qualitative and quantitative analysis. It begins with a comprehensive literature review, selecting publications from academic databases like IEEE Xplore, PubMed, Zenodo, and JSTOR, using keywords such as 'blockchain', 'data governance', and 'user privacy'. To ensure a current perspective, the search is focused on publications from the last five years. Then, a thematic analysis is conducted to identify trends, challenges, and innovative solutions. Quantitative data from case studies and relevant statistics complement the qualitative analysis, providing a comprehensive view of blockchain implementation in data governance.

A long search has also been carried out for non-academic articles that provide different points of view, which, although not scientific papers, add great value to the context of this document.

# USE-CASE DRIVEN USER Privacy and data governance: REAL PROBLEMS AND EXISTING APPLICATIONS

In the current information era, user data privacy and governance have become critical issues, particularly as innovative technologies change how we interact with data. Blockchain implementation offers significant potential for improving privacy and governance through a decentralized, secure data management framework. Additionally, data handling ethics have become a focal point as tech companies balance business practices with social responsibility and regulation. This context explores how current technological developments, like blockchain, and ethical considerations in data governance, can address real issues and enhance existing privacy and data management applications. These advancements, alongside evolving regulatory frameworks, present an opportunity to develop more robust, user-centric systems that not only protect individual privacy but also promote ethical and responsible data management.

It's essential to explore specific use cases illustrating how technologies like blockchain and ethical data management practices are applied across various contexts. In healthcare, protecting sensitive medical records is crucial; in finance, transaction security and transparency are critical. Retail requires balancing customer data management and personalization with privacy. Education faces unique challenges in protecting student and faculty data. Entertainment and social media need careful handling of personal data and user preferences. In energy and sustainability, transparency in data management is key for environmental initiatives. Supply chain management benefits from blockchain's traceability and efficiency. For public sector and government, trust and security in citizen data management are vital. Lastly, in the automotive and transportation industry, secure and efficient data integration is essential for innovation and safety. These use cases underscore the need for privacy and data governance solutions that are both robust and tailored to each sector's specific needs.

# Healthcare

In the healthcare sector, user privacy and data governance are of paramount importance due to the sensitivity of medical information. A notable challenge faced is the secure sharing of medical records. Patients often need their medical histories and records accessible by various healthcare providers, which raises concerns about data security. Robust data governance mechanisms are critical to ensuring that sensitive medical information remains confidential and is accessible only to authorized personnel, enabling secure and seamless data sharing while safeguarding patient privacy. Similarly, ensuring patient data privacy is a significant concern. With the vast amount of personal and health data generated and stored in electronic health records, it's crucial to have stringent data governance practices in place. These practices help protect patient data, control access, and ensure compliance with stringent data protection regulations. By establishing data governance frameworks that address these concerns, the healthcare sector can continue to provide high-quality care while upholding patient confidentiality.

* **Secure sharing of medical records**: In the healthcare sector, maintaining the security and privacy of patients' medical records is paramount. Robust data governance ensures that sensitive medical information remains confidential and accessible only to authorized personnel.
* **Patient data privacy**: Protecting patient data is a significant challenge in the healthcare industry. Comprehensive data governance mechanisms are essential to safeguard the personal information of individuals and ensure compliance with privacy regulations.
* **Clinical Trials and Research Data Management:** used to secure and streamline the management of clinical trial data, ensuring integrity and transparency in research.
* **Pharmaceutical Supply Chain Integrity:** solutions are applied to track the production and distribution of pharmaceutical products, helping to combat counterfeit drugs.
* **Remote Patient Monitoring:** blockchain facilitates secure and reliable data sharing in remote patient monitoring systems, enhancing patient care and data privacy.

In healthcare, data governance involves managing policies, procedures, and standards to ensure data's accuracy, consistency, and security. Its key objectives include patient safety, improved clinical outcomes, and compliance with industry regulations. Trust is acknowledged as crucial for the acceptability of data sharing and the adoption of new health technologies such as AI, though there's reported distrust in this domain. Data privacy regulations in healthcare are evolving, and it is important to understand regulations like HIPAA, GDPR, and CCPA. See more on these regulations in Section 4.

The paper "A Privacy-Preserved and User Self-Governance Blockchain-Based Framework to Combat COVID-19 Depression in Social Media" introduces a blockchain-based framework to manage COVID-19 related depression on social media. It proposes a solution using the DPoS (Delegated Proof of Stake) consensus protocol to share depression-related information on social media platforms while maintaining user information privacy and autonomy. The approach focuses on identifying and assisting users with depression through blockchain, emphasizing the importance of privacy and self-management of data in the context of a pandemic.

MediLedger

MediLedger is a project that employs blockchain technology to enhance the pharmaceutical supply chain. It focuses on ensuring drug authenticity and reducing counterfeit risks. MediLedger offers a secure, decentralized platform for supply chain participants like manufacturers, distributors, and pharmacies to track and verify drug origins and journeys. This system not only improves patient safety but also aids pharmaceutical companies in regulatory compliance and operational efficiency enhancement.

Patientory

Patientory is a blockchain-based electronic health records management platform. It focuses on providing patients with secure and easy control of their medical information. The platform utilizes blockchain to ensure data privacy and security, allowing users to store, manage, and securely share their health information. Patientory also aids in connecting patients with healthcare providers, facilitating more effective communication and better care coordination.

BurstIQ

BurstIQ is a platform that leverages blockchain technology for secure health data management. It allows users to control and share their medical information securely. BurstIQ's blockchain solution addresses data security and interoperability challenges in healthcare, enabling seamless data exchange while maintaining privacy and compliance with regulations. The platform supports a range of health data services, from personal health management to clinical research, emphasizing user empowerment in health data control.

Guardtime Health

Guardtime Health specializes in securing health data using blockchain technology. It focuses on enhancing interoperability among various healthcare systems while ensuring data integrity and security. Guardtime Health's blockchain solution is designed to facilitate secure and efficient exchange of medical information, adhering to privacy and regulatory standards. This approach aims to streamline healthcare processes, improve patient outcomes, and foster trust in digital health data management.

Embleema

Embleema is a virtual trial and regulatory analytics platform designed to fast-track drug development. Users are recruited to digitally consent to secure, untampered medical data collection, which is then stored on Embleema’s blockchain and analyzed.

Embleema’s platform allows patients to assist in speeding up treatment availability and improving safety, all through the company’s Virtual Studies Suite.

# Finance

The finance sector relies heavily on user data and transactions, making data governance and user privacy essential components. In the context of fraud prevention, finance institutions face constant threats from fraudulent activities, be it credit card fraud, identity theft, or other financial crimes. Effective data governance strategies are vital to implementing robust fraud prevention measures. These measures include transaction monitoring, anomaly detection, and the secure storage of financial data, enabling institutions to safeguard their assets and protect their clients. Additionally, secure transactions are central to the finance industry. From online banking to mobile payments, data governance practices help maintain the integrity and confidentiality of financial data, protecting it from unauthorized access and manipulation. Encryption, access controls, and authentication mechanisms are just a few examples of the tools used to ensure secure financial transactions. By upholding strict data governance standards, the finance sector can enhance customer trust and protect the financial system's integrity.

* **Fraud prevention**: The finance sector faces constant threats from fraudulent activities. Data governance plays a crucial role in implementing fraud prevention measures, ensuring the security of financial transactions, and safeguarding customer assets.
* **Secure transactions**: Enabling secure transactions is a fundamental goal in the finance industry. Data governance strategies are employed to maintain the integrity and confidentiality of financial data, protecting it from unauthorized access and manipulation.
* Risk Management: Financial institutions use advanced data analytics for risk assessment, analysing patterns and trends within vast datasets to identify potential risks. This proactive approach aids in minimizing financial losses and maintaining system stability.
* Personalized Banking Services: Banks and financial services companies harness customer data to offer personalized banking solutions and financial advice. By analysing spending habits, investment history, and financial goals, they can tailor products and services to individual needs, while adhering to strict privacy and data protection regulations.

In finance, data governance is critical for ensuring regulatory compliance in a heavily regulated environment. Data governance also supports data-driven decision-making by creating reliable, well-managed data pipelines. Data quality is essential for accurate decision-making, and data governance can improve this by setting standards for data entry, storage, and usage.

Ripple

Ripple is a blockchain-based platform designed to revolutionize financial transactions, particularly in the realm of international transfers. It offers a faster, more cost-effective alternative to traditional banking systems for cross-border payments. Utilizing its own cryptocurrency, XRP, Ripple enables secure, instant, and low-cost international transactions. This innovation in blockchain technology has positioned Ripple as a key player in financial technology, offering transformative solutions for global finance challenges. Ripple's technology is particularly notable for its scalability, speed, and the reduced transaction costs it offers.

Circle

Circle is a fintech company that utilizes blockchain technology to facilitate peer-to-peer payments, making financial transactions simpler and more accessible. Their platform allows users to transfer money as easily as sending a text, streamlining the process of sending and receiving payments. Circle's use of blockchain technology enhances security, reduces transaction costs, and improves efficiency, making it an innovative solution in the financial services sector.

Chainalysis

Chainalysis is a blockchain data analytics platform that provides insights to financial institutions and government agencies. It specializes in monitoring and investigating cryptocurrency transactions to enhance transparency and security in the digital finance world. Chainalysis plays a crucial role in detecting and preventing illegal activities in cryptocurrency markets, such as money laundering or fraud. Their tools help ensure compliance with financial regulations, aiding organizations in navigating the complex landscape of digital currencies.

BitPesa

BitPesa is a digital foreign exchange and payment platform focused on the African market. It uses blockchain technology to facilitate business-to-business transactions across borders. BitPesa simplifies the process of sending and receiving payments in different currencies, reducing the cost and increasing the speed of transactions. This service is particularly beneficial for businesses operating in Africa, offering a more efficient and cost-effective solution for handling cross-border payments.

Digital Asset Holdings

Digital Asset Holdings is a financial technology firm that develops blockchain-based solutions for various financial services, including transaction settlement and record-keeping. The company focuses on creating tools that streamline and enhance the efficiency of operations within financial markets. Their solutions aim to reduce risk, ensure compliance, and improve the overall functioning of financial systems. Digital Asset Holdings is notable for its efforts to integrate blockchain technology into traditional financial infrastructures.

Veem

Veem is a global payments platform that utilizes blockchain technology to facilitate small business payments worldwide. It simplifies international transactions, making them more transparent and cost-effective. Veem's service is particularly beneficial for small and medium-sized enterprises involved in international trade, offering an efficient alternative to traditional banking methods for cross-border payments.

BlockFi

BlockFi is a financial services company that leverages blockchain technology to offer products like interest-earning accounts and loans using cryptocurrencies. BlockFi's platform provides a secure and efficient way for users to access financial services using digital assets. They focus on bringing traditional banking services to the cryptocurrency market, enhancing the utility and accessibility of cryptocurrencies for a broader range of financial activities.

# RETAIL

Retail businesses collect extensive customer data for marketing and operational purposes, making data governance critical in maintaining customer trust and regulatory compliance. The challenge of customer data privacy arises as retailers gather vast amounts of data about their customers, including purchase history, preferences, and personal information. Effective data governance practices are essential to protect customer privacy and comply with data protection regulations like the General Data Protection Regulation (GDPR) in Europe. Retailers must ensure that customer data is stored securely, and that access is restricted to authorized personnel. Additionally, supply chain transparency is a growing concern in the retail sector. Consumers increasingly demand information about the products they purchase, including details about their origins, sustainability, and quality. Data governance solutions can facilitate the collection and dissemination of accurate and trustworthy data related to product origin, manufacturing processes, and ethical sourcing. By embracing data governance, the retail sector can not only meet consumer demands but also improve supply chain efficiency and transparency.

* Customer Data Privacy: Retail businesses, which collect extensive customer data, need robust data governance to ensure customer privacy, regulatory compliance, and consumer trust. Blockchain can offer secure, decentralized storage and management of customer data, enhancing privacy and control.
* Supply Chain Transparency: Blockchain solutions in retail enable transparent tracking of product origins and quality, addressing concerns about sustainability and ethical sourcing. This transparency fosters consumer trust and brand loyalty.
* Loyalty Programs and Rewards: Blockchain enables secure and transparent loyalty programs, allowing customers to earn and redeem rewards seamlessly. It enhances the integrity of loyalty systems and prevents fraud.
* Anti-Counterfeiting Measures: Retailers can use blockchain to verify the authenticity of products, particularly in luxury goods, reducing the prevalence of counterfeits.
* Real-time Payments and Settlements: Blockchain can revolutionize the payment process in retail, enabling real-time transactions and settlements. This technology reduces transaction times and costs, streamlining the payment process for both retailers and consumers.

Walmart

Walmart's blockchain project focuses on improving food traceability and safety in the retail sector. By utilizing blockchain technology, Walmart can track the origin and journey of food products through its supply chain. This initiative enhances transparency, ensuring that consumers have access to detailed information about the food they purchase, including its source and processing history. The project aims to improve food safety standards, reduce waste, and build consumer trust in Walmart's food products. This application of blockchain in retail showcases how technology can innovate traditional supply chain practices.

De Beers

De Beers' blockchain project, known as Tracr, is focused on enhancing the traceability of diamonds in the retail industry. This initiative ensures that each diamond's journey from the mine to the retail point is tracked, guaranteeing authenticity and ethical sourcing. Tracr provides a transparent record of a diamond's origins, characteristics, and ownership history, helping to prevent the trade of conflict diamonds and reassure consumers about the ethical standards of their purchases. This application of blockchain in retail illustrates how technology can bring transparency and trust to luxury goods.

Maersk

Maersk, in collaboration with IBM, has developed a blockchain solution called TradeLens for the maritime shipping industry. This platform enhances the efficiency and transparency of shipping logistics by providing a secure and real-time sharing environment for shipping data among various stakeholders in the supply chain. TradeLens allows participants to track cargo and manage documentation, significantly reducing the time and cost associated with maritime transport. This application of blockchain demonstrates its potential to revolutionize logistics and supply chain management in the retail industry.

LVMH

LVMH's blockchain project, AURA, is designed to verify the authenticity and trace the origin of luxury goods. This initiative combats counterfeiting in the luxury retail sector by providing a secure digital ledger that tracks each product's lifecycle, from production to sale. Consumers can access detailed product history and proof of authenticity, ensuring the legitimacy of their luxury purchases. AURA exemplifies how blockchain can enhance consumer trust and transparency in the high-end retail market.

Alibaba

Alibaba's blockchain initiative is focused on protecting intellectual property and ensuring authenticity within its e-commerce platform. The technology is used to track and protect the rights of creators and brands, ensuring that copyrights are respected, and counterfeit goods are minimized. This application of blockchain in retail is crucial for maintaining the integrity of online marketplaces and safeguarding the interests of both sellers and consumers in the digital economy.

Starbucks

Starbucks' blockchain initiative, in partnership with Microsoft, is aimed at tracing the journey of coffee beans. This program allows Starbucks to track its coffee products from the farm to the final cup, ensuring transparency and sustainability in its supply chain. Customers can access detailed information about the origin of the coffee beans, the farmers who grow them, and the environmental impact. This application of blockchain technology enhances consumer awareness and supports Starbucks' commitment to ethical sourcing and sustainability in the coffee industry.

JD.com

JD.com's blockchain initiative is focused on enhancing inventory management and quality assurance in their e-commerce operations. By implementing blockchain technology, JD.com can effectively track the origin, authenticity, and quality of products listed on its platform. This initiative ensures a transparent supply chain, providing customers with reliable product information and fostering trust in the e-commerce marketplace. This application of blockchain technology demonstrates its potential to revolutionize retail operations, particularly in online shopping.

# EDUCATION

In the education sector, institutions handle a vast array of sensitive data, from student records to research findings and intellectual property. Secure student records management is a top priority, considering that educational institutions are entrusted with personal and academic information of students. Data governance practices are essential to secure these records, control access, and protect the academic and personal information of students. This includes implementing access controls, data encryption, and regular security audits to ensure that student data remains confidential. Moreover, intellectual property protection is crucial for educational organizations engaged in research and innovation. Universities and research institutions generate valuable intellectual property in the form of research findings, patents, and innovations. Data governance mechanisms help safeguard these intellectual assets by controlling access to research data and ensuring that patents and copyrights are protected. By adhering to comprehensive data governance practices, educational institutions can maintain data security, protect student privacy, and safeguard valuable intellectual property.

* Secure Student Records: Educational institutions use data governance to protect sensitive student records. This includes controlling access and safeguarding academic and personal information.
* Intellectual Property Protection: Ensures the security of research findings and innovations, maintaining the integrity of intellectual property rights within educational organizations.
* Credential Verification: Blockchain technology is employed to authenticate academic credentials, reducing fraud and ensuring the validity of qualifications.
* Online Learning Platforms: Data governance is applied to protect personal information and learning progress on e-learning platforms.
* Research Collaboration: Facilitates secure data sharing in research projects across institutions, protecting sensitive information and intellectual property.
* Attendance and Academic Achievement Records: Utilizing blockchain to securely record and store student attendance and academic achievements.

Student1

"Student1" is a specific platform that provides cloud-based systems for the management of student data and records in educational institutions. It emphasizes the secure handling of sensitive student information, integrating advanced data governance and privacy measures. This platform ensures controlled access and safeguards academic and personal student data, streamlining data management while maintaining compliance with privacy regulations. "Student1" represents a practical application of technology solutions in the education sector, addressing real-world needs for data security and efficiency.

IPfolio

IPfolio is a platform that helps universities and educational organizations manage and protect their intellectual property (IP) rights. It provides tools for tracking, managing, and safeguarding various forms of IP, including innovations and research findings. IPfolio's services are designed to streamline the IP management process, ensuring that educational institutions can effectively secure their intellectual property.

Blockcerts

Blockcerts is an initiative that utilizes blockchain technology for issuing and verifying academic credentials. It provides a decentralized system for creating, issuing, and storing certificates and badges, ensuring authenticity and reducing the risk of fraud. Blockcerts enables educational institutions to issue tamper-proof records, and recipients can easily share their verified credentials with employers or other institutions. This application of blockchain in education enhances the security and portability of academic achievements.

Coursera

Coursera, an online learning platform, offers a wide range of courses from various educational institutions globally. While Coursera itself does not specialize in blockchain technology for its operations, it focuses on providing secure and accessible e-learning experiences. It implements robust data protection measures to safeguard user information, ensuring privacy and security for its learners. Coursera's platform is a significant example of how technology can enhance educational accessibility and quality in the digital age.

Artifacts

Artifacts is a blockchain-based platform designed for academic research collaboration. It provides a secure environment for researchers to share data and findings while maintaining intellectual property rights and data integrity. The platform facilitates transparent and verifiable tracking of research contributions, enhancing collaboration among scientists and institutions. This application of blockchain technology aims to revolutionize how academic research is conducted and shared, ensuring credibility and secure data management in scholarly work.

Accredible

Accredible is a platform that leverages technology to provide digital badges and certificates for academic and professional achievements. It enables educational institutions and organizations to issue verifiable and secure digital credentials. These credentials can be easily shared by recipients across digital platforms, enhancing their visibility and credibility. Accredible's approach modernizes the certification process, making it more accessible and convenient for both issuers and recipients.

# ENTERTAINMENT and SOCIAL MEDIA

Social media platforms have become central to modern communication, and with that comes a responsibility to protect user data and maintain the quality of content. One of the key challenges in this sector is user data privacy. Social media platforms gather extensive user data, including personal information, preferences, and online behaviour. Robust data governance practices are crucial to protect user privacy, ensure responsible data handling, and comply with data protection laws, such as the European Union's General Data Protection Regulation (GDPR). These practices include transparency in data collection, robust consent mechanisms, and user data access controls, allowing users to have control over their information. Content moderation is another significant issue in the social media domain. Maintaining the quality and safety of content shared on social platforms is essential to foster a positive online community and prevent the spread of harmful or inappropriate content. Data governance mechanisms, such as content filtering algorithms and user reporting systems, play a pivotal role in identifying and removing content that violates community guidelines. By implementing effective data governance, social media platforms can create a safer, more responsible, and user-centric online environment while respecting user privacy and content quality.

* **User data privacy**: Social media platforms handle extensive user data. Robust data governance is crucial to protecting user privacy, ensuring responsible data handling, and complying with data protection laws.
* **Content moderation**: Maintaining the quality and safety of content shared on social media is a significant challenge. Data governance mechanisms are used to implement content moderation, filter out harmful content, and protect the online community.
* Targeted Advertising: Balances personalized advertising with user privacy. Data governance ensures relevant content delivery while respecting privacy norms and user preferences.
* Royalty Distribution in Streaming: Manages the fair allocation of royalties to content creators in streaming platforms, using data tracking to accurately measure content usage and viewership.
* VR/AR Data Management: Safeguards personal data in virtual and augmented reality experiences, ensuring immersive interactions are secure and privacy compliant.

In the realm of social media, companies are increasingly focused on regulating user privacy. This involves developing and implementing robust data governance measures to protect user data and comply with various data protection laws. The use cases include employing blockchain technology for secure data management, enhancing privacy in targeted advertising, and ensuring ethical handling of personal data. These efforts are crucial for maintaining user trust and adhering to legal standards in the rapidly evolving digital landscape of social media.

* Facebook’s Data Privacy Initiatives: Implementing robust data governance to enhance user privacy and comply with regulations like GDPR.
* YouTube’s Content Moderation System: Utilizes advanced algorithms and user reporting to filter harmful content, ensuring a safe online environment.
* Spotify’s Royalty Distribution Model: Employs data analysis to track music streaming for fair royalty payments to artists.
* Snapchat’s AR Filters: Provides secure and privacy-compliant augmented reality features, ensuring user data protection.

In the entertainment industry, mainly in gaming, the application of technology for privacy and data management shows a distinct approach compared to social media companies. The focus is on integrating blockchain to enhance in-game asset management and protect player data. Unique cases like the use of blockchain in virtual real estate trading in gaming environments or managing music rights in gaming platforms illustrate the industry's innovative use of technology. These applications reflect the specific needs of gaming environments in safeguarding user interactions and digital assets.

In 2017, Spotify acquired Mediachain, a blockchain startup, with the intention to explore blockchain technology for potential applications in music streaming. Mediachain's technology was aimed at addressing issues like proper attribution and royalty payments for artists. The acquisition indicated Spotify's interest in leveraging blockchain to improve and streamline music rights management and royalty distribution, although specific details on the implementation and progress of this initiative within Spotify's platform have not been widely publicized.

Audius

Audius is a blockchain-based music streaming service that seeks to disrupt the traditional music industry model by giving more power to artists and listeners. It allows artists to upload their music directly and engage with their audience without intermediaries. This model offers potentially fairer compensation and more control over their work for artists. For listeners, it provides access to a diverse range of music and direct interaction with creators. Audius exemplifies how blockchain technology can be used to foster a more direct and equitable music ecosystem.

Enjin

Enjin is a blockchain platform that focuses on the gaming industry. It enables developers to create and manage virtual goods on the Ethereum blockchain. Enjin's platform is used for integrating blockchain-based assets like in-game items and currencies, providing a secure and transparent system for their management. This approach offers gamers true ownership of their in-game assets and enhances the overall gaming experience. Enjin demonstrates how blockchain can be innovatively applied in gaming for asset management and user engagement.

Decentraland

Decentraland is a virtual reality platform powered by the Ethereum blockchain. It allows users to create, experience, and monetize content and applications in a virtual world. In Decentraland, users can purchase and develop virtual land, creating an immersive experience that blends gaming with digital asset ownership and management. This project showcases the innovative use of blockchain in virtual environments, offering a unique combination of entertainment, digital real estate, and community-driven development.

# ENERGY AND SUSTAINABILITY

In the energy and sustainability sector, the integration of robust data governance and privacy practices plays a pivotal role. Efficient resource management is a key application area, where data analysis is used to optimize energy consumption and reduce waste. This involves strategies for sustainable energy use and identifying areas for improvement in resource utilization. Another crucial application is in monitoring and reporting carbon emissions. Accurate data collection and analysis are essential for assessing environmental impacts, complying with international environmental standards, and supporting sustainability initiatives. These use cases underline the importance of data governance in driving sustainable practices and promoting a greener future.

In the context of energy and sustainability, the integration of data governance and privacy practices is becoming increasingly important. This sector is at the forefront of addressing global environmental challenges, where efficient management and analysis of data play crucial roles. From optimizing energy consumption and integrating renewable sources to monitoring carbon emissions and developing smart grid technologies, data governance is pivotal in enhancing sustainability efforts. Additionally, it aids in regulatory compliance and informs policy development, paving the way for technological advancements in energy storage and distribution. This introduction sets the stage for exploring the multifaceted impact of data governance in driving sustainable and innovative practices in the energy sector.

* Resource Management and Efficiency: Leveraging data to optimize energy consumption patterns, organizations implement strategies to enhance energy efficiency, reduce waste, and promote sustainable energy use.
* Emissions Monitoring and Reporting: Accurate data collection on carbon emissions enables organizations and governments to monitor and report their environmental impact, ensuring compliance with international standards and promoting sustainability initiatives.
* Smart Meter Data Analysis: Utilizing smart meters for detailed insights into energy usage patterns, helping both providers and consumers optimize energy consumption.
* Demand Response Systems: Implementing systems that adjust energy consumption based on real-time data to balance supply and demand, reducing strain on energy grids.
* Predictive Maintenance in Energy Infrastructure: Leveraging data for predictive analytics to maintain energy infrastructure, enhancing efficiency and reducing downtime.

Germany's Energiewende

Germany's Energiewende ("Energy Transition") is a significant national project aimed at transforming the country's energy system to be more sustainable and environmentally friendly. It involves a comprehensive shift from traditional fossil fuels and nuclear energy to renewable sources like wind and solar power. The initiative emphasizes not only the integration of renewable energy into the power grid but also energy efficiency and reduction in carbon emissions. This transition is a key example of how data governance and innovative policies can drive sustainable changes in the energy sector.

EnergyHub

EnergyHub is a technology company that provides a platform for managing energy use, particularly in residential areas. Their system allows consumers and utilities to effectively monitor and control energy consumption, contributing to more efficient use and cost savings. EnergyHub's solutions are focused on smart home devices and demand response programs, helping to balance energy supply and demand, and promoting sustainable energy practices in households.

Siemens

Siemens is actively involved in developing smart grid technologies, focusing on enhancing the efficiency and reliability of energy distribution. Their smart grid solutions incorporate real-time data analysis, which allows for better management of energy flow, integration of renewable energy sources, and improved responsiveness to changes in energy demand. Siemens' smart grid technology is a key component in modernizing electrical grids to be more adaptive, efficient, and sustainable.

Irena

The International Renewable Energy Agency (IRENA) is an intergovernmental organization that supports countries in their transition to sustainable energy. It provides policy advice and facilitates knowledge sharing and capacity building. IRENA plays a crucial role in gathering and analyzing data to help formulate energy policies that are environmentally sustainable and economically viable. Their efforts are instrumental in guiding international standards and practices for renewable energy implementation and regulation.

Tesla

Tesla, known for its electric vehicles, also develops advanced energy storage and distribution solutions like the Tesla Powerwall. The Powerwall is a rechargeable home battery system that stores energy from solar panels or the grid. It's designed to provide energy backup, enhance energy independence, and facilitate the use of renewable energy sources. Tesla's approach to energy storage exemplifies the application of innovative technologies in managing and optimizing energy use, aligning with sustainable energy practices.

These cases in energy and sustainability demonstrate how data governance and privacy practices, aligned with the objectives of OC2, are vital in advancing sustainable energy solutions. From Germany's Energiewende to Tesla's Powerwall, each example reflects the importance of data management in achieving energy efficiency, integrating renewable sources, and complying with regulations. These projects illustrate a broader trend towards sustainable practices, supported by innovative technologies and data-driven strategies, contributing valuable insights to the OC2 framework.

Smaller-scale projects, such as local community solar initiatives, startups focusing on microgrid technology, and university-led sustainability research, play a crucial role in the energy and sustainability ecosystem. These projects, while smaller in scope, contribute significantly by innovating at a grassroots level, serving as practical models for sustainable energy practices and data governance. They complement larger-scale initiatives, demonstrating that contributions towards a sustainable future can be effective at various scales and in diverse contexts, aligning seamlessly with the broader objectives outlined in the initial section of OC2.

Some examples of smaller-scale, tangible projects in the energy and sustainability sector include:

* Local Community Solar Projects: These projects involve community-driven efforts to establish solar power installations, providing renewable energy to local residents.  
  "Co-op Community Solar" in various regions, where local communities collectively invest in and benefit from solar energy.
* Startups Focused on Microgrid Technology: Small companies developing innovative microgrid solutions to enhance energy distribution and reliability in remote or under-served areas.  
  "LO3 Energy" develops microgrids using blockchain to enable local energy trading.
* University-Led Sustainability Initiatives: Academic research projects exploring new methods of energy conservation, storage, or efficiency.  
  MIT's "Sustainable Energy Initiative" conducts research on advancing sustainable energy technologies and solutions.

These projects are more directly relatable and demonstrate practical applications of sustainable energy practices and data governance at a community or regional level.

# Supply Chain Management

In the realm of supply chain management, the integration of sophisticated data governance and privacy protocols is pivotal. This field encompasses complex networks of production, shipment, and distribution, where efficient and transparent data management is crucial. Effective data governance in supply chain management not only enhances operational efficiency but also ensures the integrity and security of data across various stakeholders. This approach is instrumental in optimizing logistics, managing inventory, and ensuring timely delivery, thereby fostering a more resilient and responsive supply chain ecosystem.

* Product Traceability and Authenticity: Data governance enables the tracking of product origins and movements, ensuring authenticity and preventing counterfeiting. This is crucial for maintaining quality standards and consumer trust.
* Logistics Optimization: Strategic data analysis improves logistics efficiency, encompassing inventory management and transport route optimization. Efficient data management in this area ensures timely deliveries and cost-effective operations, vital for competitive advantage in the market.
* Sustainable Sourcing: Implementing data governance to monitor and ensure sustainable practices in sourcing materials, crucial for eco-friendly supply chains.
* Blockchain for Supply Chain Transparency: Utilizing blockchain technology to create a transparent and tamper-proof record of transactions, enhancing trust and traceability.
* Predictive Analytics for Demand Forecasting: Using data analysis to predict market trends and customer demand, improving inventory management and reducing overstock or shortages.
* Cold Chain Monitoring: Implementing IoT and data governance for monitoring temperature-sensitive products during transportation, ensuring quality and compliance.
* Supplier Relationship Management: Leveraging data to evaluate and manage supplier performance, ensuring efficient collaboration and supply chain resilience.

Provenance and Woolmark

Provenance and Woolmark collaborated on a project using blockchain technology to trace and verify the production and sourcing of wool. This initiative was designed to ensure that wool products are sourced sustainably and ethically. The project aimed to provide transparency in the wool supply chain, allowing consumers to have detailed information about the origin and production practices of wool products. This use of blockchain in the textile industry serves as an example of how technology can enhance sustainability and ethical practices in supply chains.

Walmart

Walmart's blockchain project focuses on enhancing food safety and traceability within its supply chain. They implemented a blockchain system to track the journey of food items, from their origin at the farm to the store shelves. This initiative aims to improve transparency in the food supply chain, allowing for quicker identification and resolution of food safety issues. The project highlights how blockchain technology can be used to ensure product safety and build consumer trust in food quality.

DHL and Accenture

DHL, in collaboration with Accenture, developed a blockchain-based project to enhance the tracking and security of pharmaceutical products. This initiative aims to combat counterfeit drugs in the supply chain, ensuring the integrity and authenticity of pharmaceuticals. The project uses blockchain technology to create a secure, transparent record of each medicine's journey from manufacture to delivery, contributing significantly to the safety and reliability of pharmaceutical distribution.

Maersk and IBM

Maersk and IBM collaborated on TradeLens, a blockchain-based digital shipping solution designed to promote global trade transparency. This platform enables various stakeholders in the shipping industry, such as cargo owners, shipping lines, and customs authorities, to access real-time shipping data and documentation. TradeLens enhances efficiency and accuracy in global shipping operations by improving information sharing and reducing the time spent on document processing.

Sensitech's Cold Chain Monitoring Solutions

Sensitech's Cold Chain Monitoring Solutions focus on monitoring temperature-sensitive products during transportation. They provide technology to ensure that goods such as pharmaceuticals and perishable foods are transported within safe temperature ranges. This is crucial for maintaining product quality and compliance with health and safety standards. Sensitech's solutions use sensors and data analysis to provide real-time monitoring and alerts, helping prevent spoilage and ensuring the integrity of temperature-sensitive shipments.

IBM Food Trust

IBM Food Trust is a blockchain-based system designed to enhance food safety and traceability. It provides a transparent platform for tracking food products throughout the entire supply chain, from the farm to the consumer. This system enables improved food safety standards, reduces waste, and ensures fresher produce. By leveraging blockchain technology, IBM Food Trust enhances the ability of different stakeholders in the food supply chain to share information quickly and accurately, improving overall food quality and safety for consumers.

# Public Sector and Government

In the public sector and government, effective data governance and user privacy are critical for ensuring transparent and efficient public services. These institutions handle vast amounts of sensitive information, making data security and ethical management paramount. The application of robust data governance frameworks in this sector is key to maintaining public trust, enhancing service delivery, and ensuring compliance with legal and ethical standards. This integration facilitates a more accountable and responsive government, paving the way for innovative and citizen-centric public services.

* Citizen Data Management: Governments utilize data governance for managing citizen information, ensuring accuracy, privacy, and security. This includes personal identification data, tax records, and health information.
* Policy Making and Public Services: Data-driven decision-making in policy formulation and public service delivery. Effective data governance allows for the analysis of large datasets to inform policy decisions and improve public service offerings.
* Digital Identity Management: Implementing secure digital identity systems for citizens, enhancing access to government services and protecting identity data.
* Public Health Data Analysis: Utilizing data governance to manage public health data, crucial for disease tracking, health policy development, and crisis response.
* E-Government Services: Offering citizen-centric online services, requiring effective data management for service delivery and data protection.
* Infrastructure and Urban Planning: Using data analytics for sustainable urban development and infrastructure planning, improving living standards and efficiency.
* Transparency and Anti-Corruption Initiatives: Employing data governance tools to increase transparency in government operations and combat corruption.

In the public sector and government, the effective integration of data governance and privacy is crucial, as seen in the examples of managing citizen data and policymaking. Projects like Estonia's e-Residency, Singapore's Smart Nation Initiative, and India's Aadhaar program demonstrate practical applications of these principles. They showcase how innovative data governance can enhance public services, improve efficiency, and foster transparency.

Estonia's e-Residency Program

Estonia's e-Residency program is a groundbreaking initiative that offers a digital identity to non-residents, allowing them to access Estonian services online. The program enables entrepreneurs globally to establish and manage an EU-based company online. This forward-thinking approach to digital identity and business administration has positioned Estonia as a leader in digital governance.

Singapore's Smart Nation Initiative

Singapore's Smart Nation Initiative is a comprehensive effort to leverage technology to improve living standards and create economic opportunities. This initiative includes projects ranging from smart urban solutions, like intelligent transport systems and smart buildings, to digital government services. It aims to transform Singapore into a leading global city powered by digital innovation

India's Aadhaar Program

India's Aadhaar program is a massive biometric ID system, the largest in the world, providing a unique 12-digit identity number to Indian residents. It's linked to biometric data and used for various government and non-government services, aimed at improving efficiency in service delivery and reducing fraud. Aadhaar has been pivotal in streamlining processes in welfare distribution, banking, and public services.

The "Open Government Partnership" initiative

The Open Government Partnership (OGP) is a multilateral initiative that aims to secure concrete commitments from governments to promote transparency, empower citizens, fight corruption, and harness new technologies to strengthen governance. Launched in 2011, OGP provides a platform for reformers inside and outside of government to develop open government reforms and foster a global culture of transparency and accountability in public administrations. This initiative reflects a growing global movement towards open governance and participatory democracy.

Sidewalk Labs

Sidewalk Labs, an Alphabet (Google's parent company) subsidiary, focuses on urban innovation. It aims to develop technologies to address urban challenges, improve infrastructure, and enhance the quality of life in cities. The company's approach combines forward-thinking urban design with cutting-edge technology, such as IoT and AI, to create smarter, more sustainable urban environments. Their most notable project was the proposed development in Toronto's Quayside area, intended to be a model of a smart, sustainable neighbourhood. For detailed information about Sidewalk Labs and its projects, you can visit their official website or look for news articles covering urban development and smart city initiatives.

The "European Centre for Disease Prevention and Control" (ECDC)

The ECDC exemplifies data governance in public health by collecting, analysing, and disseminating critical health data across EU member states. Their role in coordinating disease surveillance and response strategies hinges on the effective management of health data. This approach underscores the importance of data governance in public health policymaking and crisis management, aligning with the overarching themes of effective data usage and privacy concerns in the public sector, particularly in the context of public health management.

These projects underscore the transformative impact of data governance in the public sector. They highlight the potential of technology to drive more accountable, efficient, and citizen-centric government services, reflecting the objectives and themes central to OC2.

# Automotive and Transportation

In the automotive and transportation sector, data governance and privacy are paramount for innovation and safety. This industry is rapidly evolving with the advent of connected vehicles and smart transportation systems. Effective data management is essential for optimizing vehicle performance, enhancing safety features, and facilitating efficient transportation networks. The integration of robust data governance in this sector not only improves operational efficiency but also ensures the security and privacy of user data, which is increasingly critical in an era of connected and autonomous vehicles.

* Vehicle Performance Optimization: Leveraging data to enhance vehicle efficiency, maintenance, and safety features. This involves analyzing performance metrics to identify areas for improvement.
* Traffic Management and Smart Cities: Utilizing data for efficient traffic flow management and the development of smart city infrastructure. This includes the use of real-time data to optimize traffic signals and reduce congestion.
* Autonomous Vehicle Data Analysis: Collecting and analyzing data from self-driving vehicles to improve safety features, navigation, and traffic interaction.
* Emission Monitoring and Eco-Driving: Utilizing vehicle data to monitor emissions, promoting eco-friendly driving practices and compliance with environmental regulations.
* Connected Vehicle Services: Implementing data-driven services in connected cars, such as predictive maintenance alerts and real-time navigation updates.
* Public Transport Optimization: Analyzing data from public transit systems to optimize routes, schedules, and passenger flow, enhancing the efficiency and reliability of public transportation.

Tesla

Tesla optimizes vehicle performance using data analytics and software updates. Their electric cars are equipped with sensors and computing systems that gather data on vehicle usage, performance, and environmental conditions. Tesla analyzes this data to improve features like battery life, driving range, and autopilot capabilities. Through over-the-air software updates, Tesla continuously enhances vehicle performance and introduces new features, ensuring their vehicles remain technologically advanced. This approach to data-driven performance optimization is a hallmark of Tesla's innovative strategy in the automotive industry.

Google Waymo

Waymo's self-driving technology has significant implications for smart city management. By collecting and analyzing vast amounts of data from its fleet of autonomous vehicles, Waymo contributes to more efficient traffic flow and reduced congestion in urban areas. The insights gained can inform city planning and traffic management decisions, leading to smarter, safer, and more sustainable urban environments. Waymo's technology demonstrates the potential of autonomous vehicles to integrate seamlessly into the broader ecosystem of smart city infrastructure.

Toyota

Emission Monitoring and Eco-Driving involve using technology in vehicles to track and reduce emissions, promoting environmentally friendly driving habits. Modern vehicles, especially hybrids and electric cars, are equipped with sensors and systems to monitor real-time emissions. This data is used to inform drivers about their driving patterns and how they can adjust to be more eco-friendly, like optimizing fuel efficiency and reducing idling. These technologies are crucial for meeting environmental regulations and contribute to reducing the overall carbon footprint of transportation.

BMW Connected Drive

BMW ConnectedDrive is a suite of digital services and features designed to enhance the connectivity and functionality of BMW vehicles. It includes various services such as real-time traffic information, remote control functions, and maintenance alerts. The system integrates the vehicle with the driver's digital lifestyle, offering seamless access to information and entertainment. ConnectedDrive also focuses on safety features, providing assistance services and the ability to call for help in emergencies. This technology represents BMW's commitment to advanced automotive connectivity and user convenience.

City of London  
The Oyster card system in London is an electronic ticketing scheme used for public transport across the city. It allows travelers to use a single, rechargeable card for access to various modes of transportation like buses, underground trains, trams, and some river services. The system simplifies payment and reduces the need for paper tickets. It also collects travel data, which helps in optimizing routes and schedules for London's public transport, contributing to more efficient and effective transit management.

# DECENTRALISED USER-CENTRIC USER Privacy and data governance FRAMEWORKS

Blockchain technology is emerging as a powerful solution in data governance and user privacy, especially in compliance with regulations like GDPR, CCPA, and HIPAA. It offers innovative approaches to regulatory compliance and decentralized identity management. Applications include secure storage and sharing of medical records under HIPAA and self-sovereign identity management, aligning with Privacy by Design principles. This text explores how blockchain can be a key tool in establishing user-centric data governance frameworks while respecting existing data privacy and security regulations.

# GDPR: Blockchain-based Solutions for GDPR Compliance

The General Data Protection Regulation (GDPR) has introduced stringent requirements for data protection and privacy. Blockchain technology plays a significant role in ensuring GDPR compliance. Blockchain's immutable ledger ensures that data transactions are recorded transparently and cannot be altered, enhancing transparency and accountability. Moreover, blockchain's decentralized nature reduces the reliance on central authorities, empowering individuals to have more control over their personal data.

A notable example of blockchain-based GDPR compliance is the use of smart contracts for consent management. These contracts allow individuals to specify how their data is used and shared. When data usage aligns with the terms defined in the smart contract, the system automatically grants access, ensuring that data processing adheres to user consent. If there is any deviation from the agreed-upon terms, the smart contract prevents unauthorized access, thereby preserving privacy.

In addition to blockchain, cryptography-based privacy-preserving techniques are instrumental in achieving GDPR compliance. Techniques like homomorphic encryption allow computations on encrypted data without revealing the underlying information. This ensures that sensitive data remains confidential while still enabling data analysis and processing. By combining blockchain and cryptography, organizations can navigate the complexities of GDPR while empowering users with greater control over their data.

# CCPA: Decentralized Identity Management using Blockchain

Examples:

* Decentralized identity management using blockchain
* Cryptography-based data anonymization techniques

The California Consumer Privacy Act (CCPA) gives consumers more control over their personal information. Decentralized identity management, facilitated by blockchain technology, aligns well with the spirit of CCPA. Blockchain allows individuals to maintain control over their digital identities and determine who has access to their personal information. This decentralized approach grants users the ability to revoke or grant permissions, offering a more user-centric and compliant way of handling personal data.

For instance, users can establish a self-sovereign identity on a blockchain, creating a secure and immutable record of their identity. This identity can be used to control access to various services and platforms. Users have the autonomy to share or withhold their information, ensuring compliance with CCPA's privacy regulations.

Additionally, cryptography-based data anonymization techniques are critical in CCPA compliance. These techniques allow organizations to process data in a way that removes personally identifiable information while still retaining data utility. This not only protects consumer privacy but also allows businesses to gain valuable insights while adhering to CCPA's requirements.

# HIPAA: Blockchain-based Secure Storage and Sharing of Medical Records

Examples:

* Blockchain-based secure storage and sharing of medical records
* Cryptography-based access control mechanisms

The Health Insurance Portability and Accountability Act (HIPAA) mandates strict standards for safeguarding medical data. Blockchain-based solutions offer a secure and compliant approach to the storage and sharing of medical records.

Blockchain's immutability ensures that medical records are tamper-proof. Once information is recorded on the blockchain, it cannot be altered or deleted. This feature is vital for maintaining the integrity of medical records and ensuring that they remain accurate and unaltered.

Cryptography-based access control mechanisms further enhance HIPAA compliance. Only authorized personnel can access specific medical records, and this access can be controlled and audited using cryptographic techniques. This ensures that sensitive medical data is only accessible to those who have a legitimate need, helping organizations adhere to HIPAA's stringent regulations.

# Privacy by Design: Blockchain-based Self-Sovereign Identity Management

Privacy by Design is an approach that emphasizes the integration of privacy into the design and development of systems and practices. Blockchain-based self-sovereign identity management is a prime example of this approach in action. With self-sovereign identities, users have ultimate control over their digital identities, deciding who has access to their personal information and under what circumstances.

Users can create and manage their digital identities on a blockchain, allowing them to selectively share identity attributes and personal information as needed. This aligns with Privacy by Design principles, as it ensures that privacy is a fundamental aspect of the system's architecture, and users have the power to determine how their data is used.

Cryptography-based privacy-enhancing technologies are integral to this approach. They enable secure and confidential data sharing. Through cryptographic techniques, data can be encrypted, and zero-knowledge proofs can be used to verify certain attributes without revealing the underlying data. This not only enhances privacy but also aligns with the core principles of Privacy by Design, ensuring that privacy is a fundamental consideration in the development of digital systems.

Examples:

* Blockchain-based self-sovereign identity management
* Cryptography-based privacy-enhancing technologies

--- NEW CONTENT ---

Considering the user-centric approach in the domain of SSI, it is of monumental importance to ensure that the users' privacy is not disclosed, namely that an adversary cannot access sensitive user information or link any data to a particular user.

There are different techniques that may be utilized to preserve the privacy of the users, where the primary concern when choosing a privacy-preserving method is to balance the trade-off between privacy and data utility. For example, by using encryption, a high level of data utility may be achieved, while privacy may be disclosed in the case of a key leakage attack. Another disadvantage of encryption-based privacy-preserving mechanisms is that they are heavy in terms of computations and may not be suitable for resource-constrained devices. On the other hand, noise addition mechanisms are lightweight, but the utility of the noisy data may be lower.

There are a number of data transformation techniques that can be used for the purpose of data anonymization. For example, **Generalisation** may be utilized to replace the values in the original dataset with more generic representation following a predefined generalisation hierarchy. Another technique called **Suppression** implies the erasure of some values from the original dataset. Suppression can be performed on different levels. Record Suppression is used to delete the entire record from the dataset, while Value Suppression is utilized to remove a particular value from the entire dataset. **Bucketisation** may be used to publish publicly known and sensitive attributes separately, whereas both datasets contain a common attribute (group id) that may be used to link a record containing sensitive data to a record with publicly known attributes. A **permutation** mechanism may be used to separate public data and sensitive attributes, where sensitive records are shuffled. However, in the presence of logical links between the attributes, the privacy guarantee may be poor. **Perturbation** mechanisms modify the records in the original dataset in a way that they do not correspond to the original values. One of the most popular data perturbation techniques is related to noise addition, where **Differential Privacy** (DP) (Cynthia Dwork, 2014) is one of the most popular approaches. The data in the original dataset are perturbed by injecting controllable noise using the Laplace or Gaussian mechanism, where the trade-off between privacy and utility is controlled by the parameter called privacy budget, which reflects the amount of noise to be injected. Thus, the higher the privacy level, the lower the data utility.

**Uber**

Uber uses DP to protect the privacy of users' geospatial data. By injecting noise into location data, Uber provides aggregated insights for analysis while ensuring the privacy of individuals.

**Facebook**

Facebook uses the concept of Local Differential Privacy (Teng Wang, 2020), which means that the noise is injected on the end-user's side before sharing data with a central server. This allows Facebook to perform aggregated analyses, while preserving the privacy of individuals.

Some of the most popular data anonymization methods include but not limited to **k-anonymity**, **l-diversity**, and **t-closeness** (De Capitani di Vimercati, 2023). The idea of K-anonymity is to ensure that no entry in a dataset can be identified using the combination of available attributes and external data. Thus, any record in a dataset should be indistinguishable from at least k-1 other records with respect to particular attributes. L-diversity is another approach that is used to protect sensitive information. The main idea of L-diversity is to make sure that any group of indistinguishable entries contains a sufficient variety of values of the sensitive attribute. T-closeness is another privacy-preserving concept that measures the distributional similarity of the sensitive attribute within each group of indistinguishable entries and compares it to the overall distribution in the entire dataset. The idea of T-closeness is to prevent outliers (values) of the sensitive attribute from being over-represented within any group of indistinguishable entries.

**Social Media Platforms** may utilize k-anonymity techniques to preserve the privacy of individuals when sharing aggregated user behavior or demographic data for analytics.

**Educational Institutions** may utilize l-diversity techniques when sharing aggregated data for analytics. This allows the protection of students' privacy by ensuring a diversity of sensitive attributes within indistinguishable groups.

**Financial Institutions** may employ T-closeness techniques to make sure that the distribution of sensitive financial attributes in aggregated data does not reveal too much information about individuals.

One of the most popular and advanced encryption techniques that is used to ensure privacy is **Homomorphic encryption** (HE) (Abbas Acar, 2019). Using HE, the computations can be performed on encrypted data without decrypting it. In contrast, in traditional encryption schemes, computations can only be performed on decrypted data, which exposes it to cyber-security risks. The key advantage of HE is that the result of a computation on encrypted data, when decrypted, is the same as if the same operations have been performed on the unencrypted data. There are three types of HE mechanisms offering different degrees of functionality. **Partially Homomorphic Encryption** (PHE) supports addition or multiplication operations on encrypted data but not both operations. **Somewhat Homomorphic Encryption** (SHE) supports both addition and multiplication but the number of operations is limited. **Fully Homomorphic Encryption** (FHE) supports an unlimited number of both addition and multiplication operations, which is computationally intensive compared to PHE and SHE.

**Tenfold**

Tenfold utilizes HE to provide a platform for secure and privacy-preserving computation on sensitive data, where data confidentiality is paramount.

**Enveil**

Enveil utilizes HE to enable secure and private search operations on encrypted data, which allows organizations to perform analytical tasks without revealing raw data.

**Secure Multiparty Computation** (SMPC) (Chuan Zhao, 2019) is a cryptography-based technique, which can be used when multiple parties want to collaboratively perform a computation without revealing sensitive information to each other. Traditionally, the parties would need to share their data with a central authority to perform computation. However, in the scenarios when parties need to perform computation on sensitive data, and they are not willing to share sensitive information with other parties, a traditional approach may not be suitable. Moreover, a central server is a single point of failure in this case, as well as it may be curious and want to access the sensitive information of participants. SMPC allows parties to jointly perform a computation while not revealing their confidential information to other parties. Some of the applications of SMPC include but are not limited to collaborative financial analysis and voting processes. During the SMPC process, parties need to communicate with each other a number of times to complete the computation, which may add additional communication overhead. However, there are some advanced SMPC mechanisms recently proposed aiming at reducing the number of communication rounds (Ran Canetti, 2020).

**Partisia**

Partisia utilizes SMPC to enable privacy-preserving machine learning in decentralized AI environment, which allows multiple parties to collaboratively build machine learning models without revealing their raw data.

**ScaleOut Software**

ScaleOut uses SMPC for privacy-preserving geospatial analytics, where multiple parties compute aggregate results without exposing their raw geospatial data.

**Blockchain** technology (Afif Monrat, 2019) has become one of the most popular technologies for privacy preservation and has attracted a lot of attention across industry and academia. Blockchain enables secure and tamper-resistant record-keeping of transactions across a number of participating nodes. The transactions in the blockchain are packed in blocks, where each new block is chained to a previous one. There is no central authority in the blockchain network where each participant stores a copy of the entire blockchain. To prevent malicious activities, participants utilize a consensus mechanism to agree on the validity of a new block before it is written to the blockchain. This ensures that all participants have access to the same information, eliminating a single point of failure, reducing the risk of fraud, and creating transparency. One of the outstanding features of blockchain is related to the concept of **smart contracts**. Smart contracts are self-executing contracts that contain the terms of the agreement written into codes. When the conditions are met, a smart contract executes automatically without any third party.

**Everledger**

To provide transparency and traceability in the diamond industry, Everledger utilizes blockchain to create a digital ledger for diamonds, which helps combat the trade of conflict diamonds and provides consumers with information about the ethical and environmental impact of their purchases.

**Accenture**

Accenture uses blockchain to enhance security and efficiency in interbank transactions, namely to streamline cross-border payments, reduce settlement times, and improve the overall security of transactions.

# Blockchain-based solutions: Zero-Knowledge Proofs

Examples:

* Zero-knowledge proofs for privacy-preserving transactions
* Decentralized data marketplaces for secure data sharing

Blockchain-based solutions play a pivotal role in enhancing privacy and data governance. For example, zero-knowledge proofs provide a powerful tool for privacy-preserving transactions. They enable parties to prove a statement is true without revealing the underlying data. In the context of financial transactions, this technology allows for secure and private transactions, as no sensitive data is exposed during the verification process. This not only safeguards user privacy but also ensures compliance with regulations like GDPR and CCPA.

Zero-knowledge proofs (ZKPs) are a cryptographic concept that plays a crucial role in blockchain-based solutions, enhancing privacy and security. In a zero-knowledge proof, one party (the prover) can prove to another party (the verifier) that they possess certain information without revealing the actual content of that information. This cryptographic technique has various applications in blockchain technology, promoting privacy, security, and efficiency. Here's an overview of how zero-knowledge proofs are used in blockchain-based solutions:

* **Privacy in Transactions:**

In blockchain networks, transactions are recorded on a public ledger. However, the details of these transactions, such as the sender, recipient, and transaction amount, are often visible to all participants. Zero-knowledge proofs can be employed to prove the validity of a transaction without revealing these details. This ensures financial privacy while maintaining the integrity of the blockchain.

* **Identity Verification:**

Zero-knowledge proofs can be applied in identity verification processes. Users can prove they possess certain credentials without disclosing the actual details. This is particularly useful in scenarios where individuals want to prove they meet certain criteria without revealing unnecessary personal information.

* **Smart Contracts and Computation:**

Zero-knowledge proofs enable the verification of computations without revealing the actual data being computed. This is valuable in smart contracts, where parties can execute complex computations without disclosing the input data. For example, a party can prove it possesses certain data without revealing the data itself, allowing for confidential and secure execution of smart contracts.

* **Password Authentication:**

Zero-knowledge proofs can be utilized in password authentication processes. Instead of sending a password, a user can prove they know the password without disclosing it, adding an extra layer of security to authentication processes.

* **Supply Chain and Asset Tracking:**

Zero-knowledge proofs can enhance privacy in supply chain and asset tracking systems. Participants can prove the authenticity of certain data (like the origin of a product) without revealing sensitive information, ensuring data integrity without compromising privacy.

* **Decentralized Finance (DeFi):**

In decentralized finance applications, zero-knowledge proofs can be employed to validate transactions and financial operations without disclosing specific details to all network participants. This helps in maintaining financial privacy and security.

* **Scalability and Efficiency:**

Zero-knowledge proofs can be used to improve scalability by reducing the amount of data that needs to be processed and stored on the blockchain. This can lead to more efficient and faster transaction processing.

Examples of zero-knowledge proof systems include zk-SNARKs (zero-knowledge succinct non-interactive arguments of knowledge) and zk-STARKs (zero-knowledge scalable transparent arguments of knowledge).

It's important to note that while zero-knowledge proofs offer enhanced privacy, their implementation requires careful consideration to ensure security and proper cryptographic design. As blockchain technology evolves, zero-knowledge proofs are likely to play an increasingly important role in addressing privacy concerns and improving the overall functionality of decentralized systems.

Several companies and projects were actively working on zero-knowledge proofs (ZKPs) and blockchain privacy. Keep in mind that the blockchain space is dynamic, and new developments may have occurred since then. Here are some notable companies and projects that were involved in ZKPs and blockchain privacy:

* [**Zcash (ZEC)**](https://z.cash/)**:**

Zcash is a cryptocurrency that focuses on privacy and uses a zero-knowledge proof system called zk-SNARKs. It allows users to make transactions with shielded addresses, where transaction details are encrypted.

* **Ethereum (ETH):**

Ethereum has been exploring the integration of zero-knowledge proofs to improve privacy and scalability. Projects like Aztec Protocol and Matter Labs are working on zero-knowledge rollups to enhance the privacy and efficiency of Ethereum transactions.

* [**Monero (XMR):**](https://www.getmonero.org/)

Monero is a privacy-focused cryptocurrency that uses a different technology called ring signatures and stealth addresses for enhanced privacy. It doesn't specifically use zero-knowledge proofs, but it is noteworthy in the privacy-focused blockchain space.

* [**zkSync (Matter Labs):**](https://matter-labs.io/)

Matter Labs is a company that focuses on layer 2 scaling solutions for Ethereum. They have developed zkSync, a layer 2 scaling solution using zk-rollups, which leverages zero-knowledge proofs for enhanced privacy and scalability.

* [**StarkWare:**](https://starkware.co/)

StarkWare is a company working on scaling solutions for blockchains, including zero-knowledge proof systems like zk-STARKs. They aim to improve scalability and privacy in blockchain networks.

* [**Quorum (Consensys):**](https://consensys.io/)

Quorum, an enterprise-focused blockchain platform developed by ConsenSys, has integrated privacy features, including a privacy implementation called Constellation. It allows for private transactions on a permissioned blockchain network.

* [**Enigma (Secret Network):**](https://www.media.mit.edu/projects/enigma/overview/)

Enigma focuses on privacy for decentralized applications (DApps). They have been working on solutions to bring privacy to public blockchains, allowing developers to build privacy-preserving applications.

* [**Oasis Labs:**](https://www.oasislabs.com/)

Oasis Labs is a blockchain platform that emphasizes privacy and security. They utilize a combination of techniques, including secure enclaves and zero-knowledge proofs, to provide confidential smart contracts and data privacy.

* [**Aztec Protocol:**](https://aztec.network/)

Aztec Protocol is a privacy-focused protocol for Ethereum that uses zero-knowledge proofs. It allows users to make confidential transactions on the Ethereum blockchain.

**Decentralized Data Marketplaces**

Decentralized data marketplaces are another innovative use of blockchain. These marketplaces enable secure data sharing while empowering users to maintain control over their data. Users can selectively share data and decide who has access, how long the data can be used, and for what purpose. This aligns with the principles of user-centric data governance and privacy, allowing individuals to share data securely and transparently while still maintaining control.

In summary, blockchain-based solutions and cryptography-based techniques are at the forefront of user privacy and data governance. They provide the tools and mechanisms necessary to comply with stringent regulations and empower users to have greater control over their personal data, ensuring that privacy remains a fundamental consideration in the design and management of digital systems.

# DECENTRALIZED DATA MARKETPLACES

In this section, the main characteristics of the data sharing and trading platforms (i.e., marketplaces) are studied in the vertical sectors of the industry where they are located and an overview of the dominant trading platforms is provided.

Based on literature (Spiekermann, 2019), (Koutroumpis, Leiponen, & Thomas, Markets for data, 2020), (Fricker & Maksimov, 2017), a data market can be understood as a digital platform on which data products are traded. Such a platform may be owned and operated by a company or organization, which participates in the platform as the sole provider or purchaser of the data. On the other hand, such a platform can act as a neutral intermediary and allow third parties to trade (sell or buy) data products. The study (Stahl, Schomm, Vossen, & Vomfell, 2016) introduces a framework for classifying the business models generally followed in electronic platforms and by extension in data trading platforms, as shown in Figure 1. This framework consists of three dimensions: orientation, ownership and business model model.

Diagram

Description automatically generated

*Figure 1:* *Business models of online shopping that are distinguished into three types of ownership.*

Platforms are placed on a scale that indicates the orientation of trading on the platform. Specifically, a distinction is made between two types of transactional relationships that can potentially be adopted on a platform, hierarchical relationships and market-based relationships. Hierarchical trading relationships are characterized by predetermined price levels for buyers or suppliers. On the other hand, when trading market relations are adopted, the competition between the participants on the platform determines the price and quantity of the goods. Also, platforms can be categorized based on their ownership model, private (a single company/organization owns the platform, buyer or seller), consortium (a small number of companies/organizations co-own the platform, either buyers or sellers) or independent (the owner of the platform has nothing to do with buyers or sellers). Based on the dimensions of the framework in Figure 2 six types of business models can be distinguished: buy-side system, sell-side system, buy-side platform, consortium platform, sell-side platform and two-sided marketplace. Examples of data trading platforms that adopt any of the above types of business models are:

* IOTA[[1]](#footnote-2)– two-sided marketplace
* Caruso[[2]](#footnote-3)– consortium marketplace
* Ocean Protocol[[3]](#footnote-4)– two-sided marketplace
* INRIX[[4]](#footnote-5)– sell-side system

Additionally, based on the study (Koutroumpis, Leiponen, & Thomas, 2017) data trading platforms can be classified based on the number of parties on each side of the market and the matching mechanism adopted. The four types of matching models and the corresponding market design, terms of exchange and some examples of platforms adopting each type are shown in Figure 2

Table

Description automatically generated with medium confidence

*Figure 2:* *Types of platforms by matching mechanism.*

Some of the existing platforms focus on trading and trading general-purpose datasets (e.g., AWS Data Exchange [[5]](#footnote-6), Dawex [[6]](#footnote-7), and DataRade [[7]](#footnote-8)), while others focus on trading data that is exploitable by specific industries (e.g., Otonomo [[8]](#footnote-9), Caruso , Wibson [[9]](#footnote-10)). Specifically, the sectors in which data markets are located and the relevant platforms are active include the following sectors:

* industry,
* health,
* automotive,
* financial,
* energy,
* Georgia,
* Public administration.

Also, there are several platforms that simply focus on indexing publicly available datasets (e.g. Google Cloud DM[[10]](#footnote-11), Azure Data Catalog[[11]](#footnote-12), Advaneo[[12]](#footnote-13), Data Intelligence Hub[[13]](#footnote-14)). A very small number of the platforms studied (<10) use data encryption or distributed ledger technologies Ledger Technologies - DLT, such as Blockchain, to store information about data exchanges (e.g. IOTA, Airbloc[[14]](#footnote-15), Wibson , Meeco[[15]](#footnote-16)). Only 4 of these platforms, Datum[[16]](#footnote-17), AMO[[17]](#footnote-18), BurstIQ[[18]](#footnote-19)and Dataeum[[19]](#footnote-20), store the actual data sets in Blockchain. An interesting case is BurstIQ , which adopts Blockchain technology to store health data shared between health organizations and health professionals. It is worth noting that some of these platforms, such as Datum, failed due to the high cost of processing the data stored on the Blockchain. So, the long-term viability of these architectures that store data on-chain is questionable.

Data trading platforms can provide both static data and (dynamic) data streams and allow access through various types of access, such as individual file downloads, APIs , custom web interfaces or specialized applications. As such, these platforms provide standardized licensing models, as well as regulations on data access and use. The format of the data exchanged on these platforms varies. Primarily, they exchange:

* *Sensor and mobility data*, i.e. data collected from IoT sensors , such as smart city data, traffic data, parking data and car data. This data is usually provided in *near real-time*, as sensors collect the data and send it directly to related platforms. Examples of platforms that utilize this type of information are the Caruso and Otonomo platforms that manage and trade data from cars. This data can be used by auto parts manufacturers, auto parts suppliers and insurance companies.
* *Geographical data*, i.e. data related to a specific location on Earth. This data is usually used by geographic information systems ( geographical information systems - GIS ). Examples of such data are environmental data, housing data, road data, weather data, business locations, and static and dynamic map data.
* *Personal data and health data*, i.e. data about patients or personal data, which is collected against a reward. Health data typically includes information about patient names, dates of birth, medical treatments, and health conditions of individuals or populations. Examples of personal data are name, gender, age, home address and income. Some dimensions of health data and personal data overlap.
* *Financial and alternative data*, i.e. data that provides information about a company's financial situation, such as data about a company's assets, liabilities and equity. Alternative datasets provide information about a company but come from sources outside the company. Examples of alternative data are credit card transactions, website usage, product reviews and price trackers. Both financial and alternative data are used by investment professionals.
* *Audience data*, i.e. data derived from the interaction of users (i.e. the "audience") with online advertisements or from the online purchase of products. This data is linked to a group of customer users and can be exploited for commercial purposes, e.g. targeted promotions. In many cases, audience data is collected by a data provider through tracking user behavior on websites and mobile apps.

The most prevalent platforms studied are described in detail, emphasizing their business models and specifically the following characteristics:

* **Value proposition:** easy access to data and provision of tools for their synthesis/analysis, high quality data, secure access to data, etc.
  + *Transaction - centric*: The platform focuses on establishing the infrastructure and necessary mechanisms to facilitate the interaction of sellers and buyers.
  + *Data - centric*:The platform provides tools for data synthesis, analysis and visualization.
* **Customer sectors:** Individuals orcompanies operating in the sectors where the provided data can be exploited, e.g. agriculture, health, automotive , etc.
* **Key Partners:** Individuals or businesses that provide data or services (e.g., software) to the Platform. A partner may provide data or services that are either industry specific or general purpose.
* **Ownership Model:** Private, Consortium or Independent. A Federated Platform makes cross-platform synergy possible. Federation falls into a different dimension than the ownership model. Typically federated platforms follow the Consortium or Independent Platform model.
* **Platform Architecture**: Centralized or Decentralized *(emphasis on transaction control rather than resource allocation).*
* **Data format:** Static and/or dynamic, which can be sensor and mobility data, geographic data, personal data and health data , etc. Also, this data can be raw , enriched or metadata.
* **Data Source:** Data collected by the Platform, data provided by customers or data provided by third parties.
* **Market access:** Open, Closed or Hybrid, via APIs, individual file downloads or specialized software.
* **Business Type:** Business - to - Business (B2B), Consumer-to- Business ( C2B ) , Anyone .
* **Correlation** **Mechanism:** One-to- one , One - to - many , Many - to - one , Many - to - many to - many )
* **Pricing:** Free, Freemium (a subset of services/data is free – the rest are paid), pay-per-use, subscription payment, service/data bundle payment, tiered payment based on usage.
* **Revenue Streams:** Transaction Fees, Subscription Fees, Storage Fees, Service Fees.
* **Price setting:** By the seller, by the buyer, by the platform, after negotiation between the parties involved.
* Currency**:** Fiat Currency, Cryptocurrency.
* **Consent mechanism:** Proof of Work (PoW), Proof of Stake (PoS), Proof of Authority (PoA), Proof of Verified Source, etc.

In the next subsection, existing platforms that manage and trade various types of data, i.e., general-purpose data markets, are presented. Afterwards, we present special-purpose data marketplaces, i.e., platforms that specialize in (but in many cases are not limited to) trading and making available data from IoT devices or personal data. In the sequel, federated data markets are described, which combine data from many different platforms.

## General Purpose Data Markets

**DAWEX**

Aiming to transact/exchange, share and exploit datasets between businesses in a safe and easy-to-use manner and with a global turnover, the Dawex market is aimed at a multitude of businesses of independent size and industry, e.g. automotive, smart city & IoT, health, finance, energy, transport, insurance, retail. As an open market it supports any data format such as files, APIs , raw or rich data. Acts as a trusted neutral broker for transactions between buyers and data providers adopting the Ethereum Blockchain on the one hand **for the direct matching of buyers/providers with sellers, on the other hand** for the management of the smart contracts concluded between them, thus making it possible to verify the integrity and authenticity of the data (Dawex, 2022), (Bounie & Quinn, 2018).

What makes the case of this market unique in terms of the network architecture is that while Blockchain is used for transactions, the control of access and origin of the data as well as services related to search, discovery and payment remain centralized. Combined therefore, Dawex's centralized system verifies the identity of interested parties through the validation process provided by Ethereum Blockchain before the data set can be purchased or exchanged. This identification process takes place both in individual transactions and when activating a registration/subscription feature offered to data providers by the Dawex market to promote and create new data offers to prospective buyers (Bounie & Quinn, 2018). Regarding this subscription feature, it is noted that Dawex allows data providers to offer prospective buyers access to their existing and future data sets through monthly subscriptions. It also allows providers to highlight the value of their data, offer discounts on published data, and adjust their prices based on the subscription period of future buyers. In this way the data providers earn revenue (Dawex, 2022), (SynchroniCity, 2022).

From the above it is obvious that the providers are not limited by a central entity in the way in which the data will be distributed. In the Dawex market the seller/provider is the one who configures the usage rights, defines which transactions will be free or for a fee as well as with which frequency-time intervals and price they will be carried out. The role of the Dawex market is to generate alerts on new data offering opportunities, facilitate businesses to find relevant data sources and advise providers (for a fee) on pricing for that data. In addition the Dawex market charges for services related to the development of a strategy to monetize this data. Similar functional features in terms of consulting services are also found in Caruso [[20]](#footnote-21), a data market related to the vertical sector of the automotive industry that we will see next.

At this point it should be pointed out that the Dawex market creates and publishes the offers of datasets of other platforms, acting only as a reseller. Storing, partaking in, or providing and reselling the datasets from the marketplace itself is not possible, as it does not support a data directory that could provide input to host actual data (Van de Ven, Abbas, Kwee, & de Reuver, 2021). Data\_Exchange is the Dawex market tool to provide comprehensive trading capability and create new revenue streams for businesses through the sharing, distribution and commercialization of data sets (Dawex, 2022).

Through the sophisticated digital tools available on this platform, users can visualize, edit, multi-criteria search, discover and explore a dataset, create alerts, test consumer engagement for datasets with thematic content as well as define contract parameters ( e.g. duration, usage rights, coverage area), which makes it easier for consumers to assess the suitability of this data for use in their own applications. Although data harmonization is not supported through the platform (data harmonization)[[21]](#footnote-22), however the “data sourcing”[[22]](#footnote-23). In addition to creating and quickly identifying the attractive data set offers offered through the use of the specific functional capabilities of the platform and related digital tools, the value proposition is also related to the following elements:

* Flexible data sharing and utilization which means data distribution is open. Data offers are visible on and off the platform (usually except by invitation from a company). The creation and publication of data offers is based on subscriptions, files or interfaces (APIs). Regarding the interfaces (APIs) for accessing the data sets, it is noted that the platform environment itself does not have a ready-made interaction and communication interface (API) between the contracting parties, including machine-to-machine communication for buying and selling data. Instead, data providers must register external APIs with the help of which they connect to the Blockchain to initiate transactions. The platform then proceeds to verify the interested buyers, which it reroutes to the provider's API to consume the data (Gindle, 2020). In the whole process of consuming “revocable distributed ledger decryption keys” (Revokable DLT decryption keys) are implemented allowing buyers to decrypt encrypted data sets for the duration of the contract with the seller/provider.
* Advanced transaction governance through: a) counterparty-tailored smart contracts and invoicing, b) tracking of all transactions c) transaction traceability and d) alert and search based data source utilization recommendations.
* Compliance with the General Data Protection Regulation for each geographic region regarding the commercial exploitation of data (Bounie & Quinn, 2018) and security. The users of the Data platform Exchange retain full control over their visibility and decide what information is shared and with whom. Personal data and activity-related data are encrypted (“privacy by design”, “privacy by default”). Also, the payment data is transmitted directly through the APIs encrypted between the trading parties as mentioned in the paragraph above, without the mediation of the Dawex market servers (Dawex, 2022).
* Achieving strong partnerships between trading parties around the exploitation of data, as well as creating new revenue streams at the enterprise level from the sale of datasets and raw data and the commercialization of insights around data. (Bounie & Quinn, 2018), (Van de Ven, Abbas, Kwee, & de Reuver, 2021)(Dawex, 2022).
* Finally, through the Dawex market, various pricing mechanisms are offered models[[23]](#footnote-24) which are differentiated according to the functions they serve, the type of data (bid price, volume, duration) and the business models they support. The individual mechanisms are divided into: a) free, b) freemium where a membership fee is provided with a monthly recurring fee for access to the market in combination with c) personalized programs that are distinguished for the authorizations regarding user quota, transactions, geographical coverage, etc. Also, supported transaction model with fee/commission for each transaction.

In Table 1 an attempt is made to present the main elements that distinguish the Dawex data market (Bounie & Quinn, 2018), (Spiekermann, 2019), (Dawex, 2022), (Fruhwirth, Rachinger, & Prlja, 2020), (Van de Ven, Abbas, Kwee, & de Reuver, 2021).

*Table 1: Summary of main business features of the Dawex platform*

|  |  |
| --- | --- |
| **Value proposition** | * Dataset trading broker that offers:   + Data Exchange as a Service   + Direct data licensing between sellers and customers/buyers - use (simplifying data trading process)   + The seller determines the conditions, frequency, way of using the data, the organization of the ecosystem regardless of the geographical origin as well as which data exchanges will be free or for a fee. * Emphasis on the development of features to comply with the different data protection regimes in the EU (GDPR), USA, Brazil, Japan with the aim of strengthening the legal security of platform users when exchanging data * Secure data / resource sharing (Secure Data sharing )   + Smart contract (Ethereum Blockchain for storing digital signatures - GPG signatures)à integrity in data file licensing agreements (licensing contracts) between buyer and provider   + " Data privacy by design "," privacy - by default platform "- use for encrypted data transactions * Dashboard for users of the platform with editing capabilities, export map & visualization diagram, selection of parameters for user license contracts, such as duration, geographic coverage, usage, sub-license right, etc. |
| **Dear Partners**  **( Key Partners)** | * Global ecosystem of strategic partnerships with international public and private organizations, multinational companies developing software products and solutions, business development services, consulting and services, research institutes and research and technology companies (e.g. OECD, PWC, Microsoft, Deloitte, Accenture, smart city insights , IDC, Gartner etc.) * Participation in GAIA -X, World Economic Forum, HORIZON -2020, DXA, Cyber Security4Europe, Consortium (Data pitch innovation program) |
| **Customer Segments** | Any business, coming from a wide field of productive activities, e.g. agriculture, energy, health, public services, environment, tourism, economy, banks, automotive industry, retail, space, etc. |
| **Business types** | Business to Business (B2B) association |
| **Data Access** | Hybrid type (possibility of creating closed user groups) à(external) APIs |
| **Data Type \_** | * Support for any format - dataset (eg, IoT data, one-time license data, financial data, user data, product data ), raw data ( raw data ), APIs , files * Static ( fully accessible) and dynamic data sets without harmonization |
| **Platform Architecture** | Centralized - transaction-centric |
| **Income Streams**  **( Revenue streams )** | * Levels of tiered access (subscription) to the platform based on a monthly subscription and depending on the terms and conditions of use (transactional services) * Freemium * Free open data availability |
| **Pricing**  **( Pricing )** | * Pay - per - use data / data bundle pricing * Individual organization-by-organization pricing agreement on request |
| **Payment in currency** | Fiat currency (fiat money) |

**OCEAN PROTOCOL**

The Ocean Protocol[[24]](#footnote-25) is a decentralized data exchange protocol and network that incentivizes the sharing of data for use by Artificial Intelligence (AI) applications. The Ocean Protocol is based on Blockchain technology and facilitates the storage, distribution and consumption of digital data goods (data assets) and data services in a secure, transparent, traceable and reliable way, while ensuring full control of data sets by their providers. This is achieved with "Service Performance Agreements" (Service Execution Agreements - SEA), which are a version of smart contracts that run on the Ethereum public network (Ethereum mainnet).

The data provided is used by data consumers and service providers to develop applications. The buying and selling of digital data goods or IT services is achieved through the *Ocean market* which brings together data providers, data consumers and service providers. The data shared through the Ocean platform are primarily datasets made publicly available and freely provided to the Ocean community (public commons data). However, there is a plan to future support the trading of private digital data goods that will be provided for a fee (priced data assets). These goods can be raw data, “cleansed” data or modeled data. Data providers can control who will have access to their data and how it will be used. The Ocean ecosystem consists of (Ocean Protocol Foundation, 2019), (Ocean Protocol Foundation; BigchainDB GmbH; Newton Circus, 2019):

* *Data Providers*. Businesses, governments, data custodians who hold high value untapped data.
* *Data Consumers.* Enterprises, governments, startups or universities that need data to analyze and train AI and Machine Learning (ML) models.
* *Ocean Community.* The general public, legislators, supervisory agencies/authorities, Internet defenders who wish to contribute to monitoring how data is used and ensuring its proper use (as defined by Data Providers).
  + *Keepers.* Part of the community are the ocean keepers nodes, which run the ocean protocol and form a Proof of Authority (PoA) network, where each node runs an Ethereum Parity client. The keepers community collectively governs the addition/removal of keepers and smart contract upgrades.
  + *Curators.* They discover digital data goods and services and signal their value through some evaluation mechanism (e.g. upvoting or downvoting) or through curation markets by buying "shares" of a dataset they believe will have high future value with the goal of making a profit.
  + *Verifiers.* They verify transactions between providers and consumers, the delivery of a data good or service to a consumer (at the required quality), and the commitment of reward to the provider.
* *Ocean shopping.* An Ocean marketplace is an electronic portal that acts as a marketplace and brings together buyers and sellers, where computing infrastructure, algorithms, storage, digital data goods and other related services are provided. Beyond the Ocean reference market, it is possible to create multiple markets from market providers.
* *Programmers.* Anyone wishing to develop a service or marketplace on top of Ocean Protocol.

Ocean Protocol incentivizes data providers by rewarding them with cryptographic tokens called ocean tokens and datatokens (ERC-29 Ethereum tokens). Ocean tokens are an internal currency for the Ocean community and can be leveraged to access Ocean data or services. Since the Ocean platform supports the existence of multiple markets, it also supports the existence of multiple datatokens each of which has its own pool of *"automated market makers" (AMM).* Anyone can inject liquidity into a market by buying "stakes" of a data good in the form of datatokens. The number of shares of a good is a metric for its value (curation markets).

Ocean smart contracts running on the Ethereum public network (Ethereum mainnet) facilitate the exchange of data and services using datatokens. Table 2 (Ocean Protocol Foundation, 2019) presents the players of the Ocean ecosystem, the services or functionality each player provides, and how they are rewarded in ocean tokens. The value network of the Ocean ecosystem is illustrated in Figure 3 (Ocean Protocol Foundation, 2019).

*Table 2: Services provided and rewards of Ocean ecosystem players.*

Table

Description automatically generated

Diagram

Description automatically generated

*Figure 3:* *Value network in the Ocean ecosystem*

Table 3summarizes the business features of the Ocean platform (Ocean Protocol Foundation, 2019), (Ocean Protocol Foundation; BigchainDB GmbH; Newton Circus, 2019), (Ocean Protocol Foundation; BigchainDB GmbH, 2020), (Myadam & Patnam, 2020), (Aliance of IoT Innovation, 2018).

*Table 3: Summary of main business features of the OCEAN platform.*

|  |  |
| --- | --- |
| **Value proposition**  **(Value Proposition)** | * Distributed open-source platform for trading and consuming data goods and related services for the purpose of training algorithms and developing distributed AI and machine learning applications. * Peer-to-Peer, no middleman/third party * Exposure of data, storage, computing and algorithms for consumption * Saving the transaction history * Data control by their providers: transparency and traceability of transactions * Transparency, compliance with legal framework e.g. GDPR * Integration of functional features of the Decentralised protocol Finance (DeFi) ("Decentralized Finance") for the digital service transactions of data goods ( data assets ) such as: * Automatic pricing utilization “automated market maker (AMM)” * "Curation", i.e. staking ("stake") to provide liquidity. Curators are rewarded with ocean tokens ("curation rewards") for assurance quality, data relevancy and with a percentage of data transaction fees. * Smart contracts ("Keeper contracts"), on multi-chain Ethereum to ensure access control. * Compute - to - Data for private data- data persistence assets owned by the owners àremote execution or development of machine learning algorithms. |
| **Dear Partners**  **(Key Partners)** | * Businesses and multinational companies e.g. insurance AVIVA, Unilever, Johnson & Johnson * Providers of AI/M software applications/tools e.g. Newton Circus, DEX * Accounting tax advice service providers e.g. PWC, MESSARI, KPMG * Third party data markets e.g. VERV, WEEVE, ConnectedLife (IoT) * Providers DLT /Blockchain platform e.g. Parity, SingularityNet, Enigma, EnergyWeb, Zeppelin, Ethereum, TrueBit, Aragon, Fetch * Consortiums (Consortium) MODI, TrustedIoT * MKOs, public and regulatory authorities, legislators, internet advocates * Academic community and research institutes * GAIA-X * Investors of capital. FabricVentures, Outlier Ventures, BIV |
| **Categories of Customers**  **(Customer segments)** | AI researchers , developers, start - ups , SMEs, multinationals, organizations, academia. |
| **Gentlemen actors** | Data Consumers, Service Providers, Data Purchases, Curators, Verifiers, Keepers |
| **Business types** | Business to Business (B2B)? Many-to-Many association |
| **Data Access** | * Open (and open source) - every access equals a service * *“ Fine - grained permissions ” (Detailed rights)* when publishing-consuming data goods in the created third-party markets. |
| **Data Format**  **(Data Type)** | Static datasets? Metadata - rich data sets (e.g., autonomous vehicle data, anonymized medical data from measurement devices; census data) |
| **Platform Architecture** | Decentralized |
| **Income Streams**  **(Revenue stream)** | * OceanDAO grants of program and project proposals from the Ocean DAO Community for Web3 development, machine learning models and applications * Disburse curation rewards over time |
| **Pricing**  **(Pricing)** | * Fixed value (fixed pricing) - simple smart contract with datatoken transfer * Automatic determination of data value through AMM pools * Transaction fees |
| **Consent Mechanism**  **(Consensus Mechanism)** | Evidence Proof-of- Authority  Mining |
| **Payment in currency** | * Rewards, transactions in cryptocurrency Ocean tokensor datatokens (ERC-29 Ethereum tokens) |

**ADVANEO**

The primary goal for ADVANEO[[25]](#footnote-26), and by extension for the software company of the same name, which is based in Germany and has taken over its management and operation, is the development of solutions that ensure the ownership of data sovereignty) to their owners. This platform creates a marketplace that allows trading, processing and analysis of datasets (Azkan, Iggena, Krotova, Spiekerman, & Otto, 2021). It also supports scalable solutions for data-driven business models and artificial intelligence applications aimed at the digital transformation of enterprises. To meet these requirements, the datasets are continuously updated and updated. Focused on promoting innovative solutions across different business sectors, this approach handles every data set format. Two formats of these datasets are distinguished, the open and the commercial type. The latter are available as metadata for purchase and sale. For their quick retrieval, data can be categorized (IoT data, environmental data, industrial production data, etc.). In addition to trading, the marketplace hosts a large number of datasets that are open source and freely available.

Peer - to - peer (P2P) topology network that has features of a decentralized system. This means that data providers (data providers) to keep raw data (raw data), stored locally, i.e. in their "ownership". Access and viewing of data on the platform are only possible in the form of metadata which provides information about the raw data offered or requested for transaction.

The transmission of the actual data can only take place after a contractual agreement has been reached between the buyer and the seller. For the secure, encrypted data transaction between trading parties, ADVANEO uses an application programming interface (API) but mainly the certified Intrusion Detection System (IDS). More specifically responsible for the secure, encrypted data transaction is the SAE link (" IDS connector ") implemented by each trading party. Via SAE connectors the seller transmits the raw data directly to the seller without using the platform, thus ensuring automation in transactions and mutual authentication between these connectors. It should also be noted that the SAE connector can be used by the end user/data buyer to integrate different data sources, coming from multiple providers, which is the basis for developing innovative solutions.

All e-commerce models, i.e. Business-to-Business, Business-to-Consumer and Consumer-to-Consumer, Business-to-Government, in which multiple sellers and multiple buyers transact, are supported for data buying and selling transactions. in the many-to-many association (Many - to - Many) (Van de Ven, Abbas, Kwee, & de Reuver, 2021).

The **value proposition** offered through this service platform is the secure exchange and transaction of data of any format (open or commercial) through SAE links, as well as the combined exploitation of data from different providers, analysis tools and algorithms to provide new services based in metadata to end users/buyers (Azkan, Iggena, Krotova, Spiekerman, & Otto, 2021). Pursuit of the market in question is the processing of transactions in an impartial manner regardless of the role of the participants (buyers/users, partners). This platform also guarantees the authentication of participants during business transactions, while ensuring the ownership of data for each trading party separately (Kramberg & Heinzl, 2021).

The virtual environment of the ADVANEO platform provides additional data governance (" data governance ") to facilitate the management of data in a controlled and clear way. This is achieved with the help of data analysis, visualization and processing tools for the presentation of digital data goods (data assets customizable presentations). For example, databases (DB), websites, files generated by applications (application output files), text files, etc. It is therefore a **data-centric data trading platform**[[26]](#footnote-27)(Spiekermann, 2019)**.** At the same time, it supports the purchase of metadata from various thematic categories of datasets with the aim of updating and gathering new data as well as the development of new business models or services by users.

**Diagram

Description automatically generated**

*Figure 4**: Advaneo market value network.*

In terms of revenue streams, the main source of income comes from using a subscription model, which differs depending on the type of business user, for example whether it is an SME, a multinational company, a sole proprietorship, etc., and varies in services provided (e.g., technical support), virtual tools and features supported from time to time. In this case, the premium and small subscription plans are offered business and enterprise, while it is possible to pay the entire amount of the subscription price or in installments using a credit card.

Advaneo also supports a "freemium" payment model, which covers free trial access allowing users to test data collection, create basic data views using visualization tools such as charts, graphs, maps, etc., as well as use "data workbenches" as 15 GB. In this model, a monthly basic fee is provided (basic fees) for providing access to the underlying platform as well as a transaction fee/fee. Finally, a personalized agreement is provided for the charge prices based on the usage model of the service (usage model) per individual business upon request, maintaining fixed pricing (Van de Ven, Abbas, Kwee, & de Reuver, 2021).

**Data Intelligence Hub (DIH)**

Aimed at buying and selling data - on an unstructured Business-to-Business (B2B) basis, the Data platform Intelligence Hub ( DIH )[[27]](#footnote-28) is a *data-centric* platform that supports the idea of an independent centralized data exchange and trading of various market sectors, also providing the appropriate tools for data collection, integration, analysis (Van de Ven, Abbas, Kwee, & de Reuver, 2021), (Spiekermann, 2019), (Deutsche telekom IoT, 2019). Its success proves the importance for these platforms of providing additional value-added services.

This solution is designed to act as a neutral broker in an open data market, mediating between providers and buyers of data. It aims to offer all kinds of data, the use of which is to contribute to the optimization of the internal processes of the participants' value chain, or to the development of innovative business models based on IoT and Artificial Intelligence - AI ). Deutsche Telekom is the operator of this cross-sector data sharing platform, and as its owner, determines and maintains its copyright (Van de Ven, Abbas, Kwee, & de Reuver, 2021).

The economic viability of the DIH platform results from the multitude of companies participating in it. The main partners of DIH are (among others) KPMG, Detecon, and T-Systems Multimedia Solutions, which are active in providing consulting services to clients for the implementation of their projects through this particular platform. These partners come from a wide range of market sectors, with typical examples being the public sector, health, finance, transport, logistics and the automotive industry. DIH's main partners also include companies specializing in data analytics and data visualization tools, location database service providers, platform providers, Microsoft Azure, as well as Cloudera [[28]](#footnote-29)as a cloud computing platform provider for machine learning.

DIH data is extracted from large internal or external data sources, and can be open or closed, commercial or public, etc. The data retrieved by DIH is published on the platform, which then invites commercial data providers to sell the data through it. Therefore, the ecosystem provided by the DIH platform for data exchange and processing includes the following (Aktas, 2019):

* Various supporting data analysis tools so that a company or organization can structure and process its data and combine it with that of the platform, using machine learning, deep learning and artificial intelligence methods. Data processing may include collection, normalization, aggregation, quality control , etc.
* Directory of over 200,000 open datasets ( data sets ) of various categories and formats.
* "Workspaces" with applications for developing database business models through the use of 50+ related tools.

The cloud-neutral DIH platform claims to be the first platform to meet International Data Spaces Association (IDSA) security standards (Van de Ven, Abbas, Kwee, & de Reuver, 2021). Data is accessed through specialized software that provides certification of participants in accordance with IDSA procedures. The approach taken is decentralized, that is, data is transferred directly from sellers to buyers through a secure connection, without passing through the platform itself. This ensures the reliability of the data, which includes the certification of the participants and the preservation of the ownership of the data sovereignty). For these purposes, the Umati approach of the VDMA (The Mechanical Engineering Industry Association) (European Round Table for Industry, 2021), which allows the collection, hosting and sharing of data produced by machines inside or outside the production environment. It enables the integrated communication of engineering tools and systems in customer-specific IT ecosystems, such as: an ecosystem of manufacturers, with 110 machines in 10 countries interacting with 30 different data users, such as machine manufacturers. They thus have the possibility to analyze data about the use and operation of the machines and to improve their products. Standardized semantics are also provided through an information model based on the OPC UA standard[[29]](#footnote-30) (Aktas, 2019).

An additional field of application of the DIH platform is the hosting of open data (Open-Data) on the platform by public organizations and local government bodies. This service eliminates the need for public bodies to have their own data delivery infrastructure. The desired data is published via web upload and an interface used for automated data upload. For the implementation of the specific service, the market in question proceeded to comply with specific legal requirements related to the provision of public data, such as ensuring the availability and free access thereof, as well as the agreement with GDPR regulations (General Data Protection Regulation of Data - GDPR) concerning data privacy (European Round Table for Industry, 2021).

The DIH platform is monetized through the use of a hybrid freemium model, whereby several features are offered for free while users are charged additional commission fees for storage, computing power usage, and transactions, as well as for selling additional infrastructure resources (Linder, Straub, & Kühne, 2022). In addition, the DIH platform earns revenue from providing internal solutions for organizations, while operating as a service provider at the Platform as a Service level as a Service ) to customers from the public sector. Payment is made with fiat currency (fictitious money).

The main features of the *value proposition* from the DIH platform are the following: a) peer - to - peer communication between nodes/users and transactions with the advantages of an open market, b) maintaining data ownership sovereignty ), with data owners maintaining control over its flow and use (Linder, Straub, & Kühne, 2022), c) neutrality for all key players to prevent the promotion of organizational silos, d) secure data exchange and storage in accordance with the IDSA standard, e) self-service access to easy-to-use data for the development of business models by the participants f) provision of data in a manner compliant with the regulations of all European data protection regulations, g) reliability, as the data will not be stored or exploited for themselves purposes of the DIH platform (for federal data storage ) promoting the certification of participants before the start of buying and selling.

The DIH platform facilitates users to combine a wide range of information from different market sectors (Figure 5), and achieve new insights based on it as well as added value. For example, international meteorological data can be used in the context of the insurance market or route information related to logistics. Other applications concern the areas of real estate management ( real estate ), smart cities, and food production, and are presented in the figure below (Aktas, 2019). Even competitors within a market can share their data safely, due to the neutrality of the platform (European Round Table for Industry, 2021). With the increasing range of data available on the DIH platform, economies of scale, scope and network effects are also created . effects ), which can be used in combination with existing analysis tools.

A city with many buildings and a road

Description automatically generated with medium confidence

*Figure 5: Use cases in the DIH market.*

* **OriginTrail**

OriginTrail is a decentralized open peer-to-peer platform that indexes knowledge [[30]](#footnote-31)(processed data with semantic information) and facilitates access to it (Rakic, Levak, Drev, Savic, & Veljkovic, 2017). This platform integrates Blockchain technology with digital supply chains to provide data integrity. The basic idea is to ensure product standards and consumer safety using a standard Blockchain-based solution through an incentivized protocol. OriginTrail addresses two key factors disrupting data collection and sharing across supply chains: a) Data fragmentation—Different data storage structures across the supply chain lead to data silos and poor data interoperability across both single and multiple supply chains . b) Data Aggregation—Most of the existing supply chains rely on a trusted intermediary or central authority that provides information about product authentication and origin.

All information in the system is stored on the Blockchain in the form of data fingerprints at the time of arrival. The data layer (OriginTrail Data Network (ODN) in the peer-to-peer network) is primarily responsible for performing all data management and connectivity functions between different data sets throughout the supply chain. To leverage data relationships efficiently, ODN incorporates a decentralized graph database that provides the most efficient solution for interconnected data in terms of interoperability, performance, and availability. Also, a number of decentralized applications (“*decentralized applications”* or “*DApps”*) act as an interface between users and the system to provide data entry facilities.

Nodes in ODN can be classified into two types according to their interaction with the supply chain: a) data creators (DC) and b) data holders (DH). DC nodes are mainly responsible for integrating supply chain data into the network and replicating it to a certain number of DH nodes, while DH nodes must ensure the immutability and storage of the corresponding data. To defend against the collusion attack, for every n DC nodes in a supply chain, an additional number of at least n + 1 DH nodes must be selected to store the replicated data. Since a DC node is also a DH node, the minimum ODN replication factor is 2n + 1 in this case. OriginTrail implements a data distribution protocol within the ODN.

The current version of OriginTrail implements PoW running on top of Ethereum. The consensus mechanism works in three steps: a) to maintain the chain of accountability, each stakeholder must be approved by the previous and following stakeholders in the supply chain, transactional compliance and sensor data, and b) the data could also be verified by control and compliance organizations by providing their confirmations.

From a business perspective, OriginTrail incentivizes network nodes in the form of cryptocurrencies (TRAC) for their resources. TRAC is also used as a guarantee fund for the honesty of data creators and the nodes that store them, but also as a means of exchange for the exchange of data between Data Providers, Data Creators and Data Customers.

***Special Purpose Data Purchases***

In this section follows an analysis of special purpose platforms, i.e. platforms that trade/trade data for specific industry sectors and/or data formats. Below, Section 3.2.1 presents platforms specializing in IoT data , while Section 3.2.2 analyzes platforms specializing in personal data trading. Some of them enable the exchange/trading of data for multiple IoT industry sectors , while some focus on data that can be leveraged by actors/companies operating in specific sectors. For example, the Otonomo and Caruso platforms are active in trading vehicle data to B2B service providers as well as the automotive aftermarket, focusing mainly on auto parts and component manufacturers, insurance companies and vehicle repair shops.

**AUTONOMOUS**

It is a cloud computing provider that has platform services [[31]](#footnote-32). It was founded in 2015 as a start - up company in Israel by Ben Volkow and Avner Coher , with the aim of creating an interconnected service ecosystem to the commercial partners of the automotive industry on a cloud computing platform computing ). The object of transactions for the participants is the vehicle/car data either in near real time or by recalling historical records (Bounie & Quinn, 2018).

The data formats are distinguished as follows: Vehicle data (near real-time or historical data), traffic data with geo - coding information information), road sign data, real-time accident data, construction equipment, historical commercial cargo and fleet data (commercial fleet data).

Otonomo focuses on the level of data set collection (unstructured or structured) from vehicles to create applications and services. Specifically, data sets are combined into data packets ( data bundles ) including data from third-party sources (e.g. application providers), aggregated metrics as well as aggregated events and metrics ). The data packets are then associated with the services offered for sale to customers.

The innovative tools it uses are all about guaranteeing data security and user privacy. More specifically, the service “Consent Management hub” provides a simple process for drivers to grant or revoke permission to access various services that consume data from their vehicles. At the same time, the service “Otonomo Data Blurring engine” uses techniques that make it impossible to verify the details of vehicles and their owners, while keeping the data usable for specific use cases such as road safety and parking (Otonomo, 2022).

Otonomo's cloud platform supports and provides the exchange, collection, aggregation (aggregate), normalization, enrichment/modification, cleansing and disposal of vehicle/car data, while ensuring the safety of data sharing, the protection of personal data and access to anonymized (transactional) data. Main design choices of the platform are easy access through a (RESTful) API of a visualized query builder (visual query builder) and a geo-fencing type user interface. Added to these is the ability to provide an application - ready & enriched data set for use by application cases and application developers, thus facilitating the reduction of cost and time to implement new applications and services related to vehicle data.

Indicative use cases / types of services are mentioned in relation to: Traffic management for the Smart City, Parking solutions (e.g., free parking space detection), Vehicle fleet management, Electric mobility services, Insurance supporting the “ pay - as - you ” business model - drive ”, remote diagnosis, maintenance forecast, location intelligence to support drivers during traffic or in cases of extreme weather conditions (Otonomo, 2022).

Regarding the business model, Otonomo is a platform service provider, which acts as a data server in a neutral manner server platform provider ”). This means, on the one hand, that the conclusion of B2B agreements between Otonomo and KPE (Standard Equipment Manufacturers) is mandatory for access to vehicle data information, on the other hand, that KPE does not manage or control the underlying platform (and infrastructure), but has control over the selection of applications and services developed by third parties to access the sharing of its data through the platform (Kao, et al., 2021), (Bounie & Quinn, 2018).

In addition, it is important to point out that the participants have a Many-to-Many relationship with each other as this particular platform acts as an intermediary between "clients" and "partners". Therefore, each user of the platform can upload and maintain data sets while access to secure transaction and interaction between them is allowed through a variety of licensing models (Otonomo, 2022).

Otonomo 's partners include user-drivers, car manufacturers (e.g. BMW, Mitsubishi Motors , Mercedes Benz ) the EV infrastructure providers provides ), KΠE, Telematics providers and private investors while customer sectors include car manufacturers, local authorities, parts and spare parts retailers, insurance companies, traffic management and infrastructure design consultancy providers ( Traffic Management consulting companies ), providers of (smart) parking services ( Parking service providers ), fleet management companies ( Fleet Management Service providers (Kao, et al., 2021), (Otonomo, 2022)..

Based on the above, therefore, the value **proposition** for Otonomo briefly includes the following elements (Bounie & Quinn, 2018), (Spiekermann, 2019):

* It is an innovator in leveraging the data format for marketing.
* It offers cooperation models to big data vendors who are the main partners, an ad-hoc agreement to share vehicle data used by data providers while ensuring their privacy and security.
* It supports the development of services and business models from third-party B2B service providers.
* It is also a neutral server that allows connected vehicle owners to take control of how their data is shared with third party service providers.
* It is an example of *a currency-centric* data trading platform.

**Revenue streams** arise from fixed pricing transaction fees and are distinguished into: a) pay-as- you - access pricing ( e.g. $ 60 per 20k trips) b) Custom plans ” based on personalization depending on the application case / type of service (Kao, et al., 2021). Another source of income comes from entering into partnership models models ) with the main partners. In this case, the excess profit is distributed at least partially among all the beneficiary parties while it is also considered a profit for Otonomo as the partners create this excess profit through the participation incentive. Finally, it should be noted that the market offers a free 10-day trial period, during which any application created has access to a limited number of countries and use cases as well as the production of reports from aggregated/enriched data.

**CARUSO**

Caruso [[32]](#footnote-33)platform is funded by TecAlliance , which provides the automotive industry with vehicle-related data. In addition to TecAlliance , shareholders of the Caruso platform are also some multinational companies such as Bosch and Continental . This is a closed data platform where only *consortium members and partners are allowed to trade* . The neutrality of the platform is guaranteed by the participation of a broad group, more than 20 shareholders, from the automotive industry. It is a *data-centric* trading platform, which is the "broker" between multiple players, just like the ADVANEO and Data Intelligence platforms Hub (Section 3.2). The Caruso platform enables the harmonization and enrichment of on-vehicle data (e.g. brakes, battery, engine, front wheel, distance travelled, speed, etc.), their trading, and access to them, while providing advisory services to the conclusion contracts, invoicing, costing to the users of the platform. Therefore, the platform can be used to serve multiple purposes, such as car sharing through appropriate applications, offering personalized car insurance contracts from insurance companies based on the driver's driving profile, etc. In addition, a software development portal is offered ( Developer portal ), with the aim of facilitating on a technical level the access of business customers' systems to data on said platform, even if they have limited technological resources or human resources. This software solution provides the following additional features: an access interface ( API ) to the data through the platform, communication with the catalog of data by thematic content, as well as the subscription process and error handling.

the platform is financially supported by the following: (a) the conclusion of a partnership *agreement* agreement) with the buyers on the condition of their participation as shareholders if they wish to use the platform and (b) a *member subscription scheme* , which consists of four levels. These include a three-month subscription to trial use of the platform, and two additional levels which allow a higher number of vehicle-level data accesses with an annual subscription of €1,500 or €15,000, depending on the member's requirements. Finally, the "On Demand" category offers additional customized options, which include *exclusive access* to a manufacturer's data as well as technical assistance services for data integration or data-driven service design (Linder, Straub, & Kühne, 2022).

Table 5 summarizes the key business features of the Caruso (Naab & Knodel, 2018)platform , (Linder, Straub, & Kühne, 2022), (Spiekermann, 2019).

*Table 5 : Summary of main business features of the Caruso platform*

|  |  |
| --- | --- |
| **Value proposition**  **( Value Proposition )** | * Broker to provide vehicle data exchange services * Harmonization and enrichment of disparate proprietary automotive vehicle data sets * Neutral data server - anonymized data * Editing (organization and presentation) in predefined Data Catalogs, for use in the development of telematics solutions in the automotive sector by customers of the platform Privacy Assurance and Data Consent Management for compliance with the GDPR " Authentication & Authorization" via API Keys - encryption ( TLS ) Big support Data analytics from third parties Consulting services in contracting , pricing , billing as a service ) |
| **Dear Partners**  **( Key Partners )** | car manufacturers (e.g. BMW), OEMs, providers of specialized software products and services for the automotive industry (e.g. Electrobit) |
| **Customer Domains**  **( Customer Segments )** | parts and spare parts retailers, insurance companies, automotive Telematics providers, IoT service and platform providers (e.g. AVAX-IoT), Leasing and Fleet Management Systems Providers |
| **Business types** | Business-to-Business (B2B) |
| **Data Access** | Open source, API |
| **Data Format**  **(Data Type)** | Aggregated vehicle data (aggregated vehicle data ) and events ( vehicle events )  *Examples:* basic information (battery, engine, brake status, tire pressure), user - based data (km traveled, geolocation), events (accident, damage) |
| **Platform Architecture** | Data Centered |
| **Income Streams**  **( Revenue stream )** | * Entering into a corporate agreement with purchasers acting as shareholders * broker transaction fees when trading data on a Business-to-Business (B2B) basis * "direct charge" for advice and services * payment of fees for API keys license agreements |
| **Pricing**  **( Pricing )** | * basic commission/fee (Membership fee) *in combination with:* * pay-per-use data / data bundle pricing based on monthly payment options: flat rate, per VIN, per call |
| **Payment in currency** | Fiat currency (fiat money) |

**IOTA**

The IOTA platform is a purpose-built platform for data transactions and micropayments (less than £0.01) between IoT devices using cryptocurrencies. IOTA adopts distributed ledger technology (DLT) specially designed for IoT device communication called Tangle. MIOTA is the cryptocurrency used for transactions on the platform. Based on an open source (open source) protocol, the platform aims to facilitate machine-to-machine communication and establish it as a service of the economy of things, without the need for human intervention.

High scalability, speed in transactions and the possibility of validating (validation) an innumerable number of transactions at the same time are some of the key advantages that make this technological approach suitable for use cases related to IoT. IOTA *has no fees, so for the recording and validation of transactions it relies on a different consensus approach that does not include a* mining data validation process and therefore exclusive miners, but also relevant consensus mechanisms of the existing Blockchain technology.

The founding team of IOTA proposed the implementation of the Tangle , which makes transactions between devices and sensors fully decentralized, and without blocks, thus allowing the network to operate asynchronously. Blockchain technology has previously identified weaknesses in managing the growing IoT crowd in real time. These concern the slow speed of transaction validation, the difficulty of scaling the network, and the "Proof-of-Work" consensus mechanism itself, which is energetically expensive (due to hashing ) , disincentive for conducting micropayments, while requires high computing power to run the block reward system (Chowdhury M. J., Ferdous, Biswas, Chowdhury, & Muthukkumarasamy, 2020).

A Tangle is a set of independent unique transactions which are recorded in a directed acyclic graph ( DAG – Directed Acyclic Graph ) where to verify the legality of a transaction, the user/validation of two previous transactions is required (Chowdhury M. J., et al., 2019). It uses Markov algorithms Chain Monte Carlo (MCMC) (IOTA, 2018), which is a stochastic process combining Monte Carlo methods with Markov chains, in which new vectors representing transactions are connected to the graph based on a Poisson distribution. In this way it becomes more resistant to security breaches than platforms based on Blockchain technology.

The graphical representation of the Tangle is given in Figure 6 (Higgins & Sandner, 2019).

Diagram

Description automatically generated

*Figure 6**: IOTA TANGLE*

The purple node symbolizes a transaction on the IOTA network. A graph is a collection of nodes that acts as a ledger to store transactions and has replaced the chain of blocks in Blockchain technology . Nodes are called IRIs. Green lines represent direct or indirect validations from subsequent trades. Transactions are always sequenced from the first to the next and so on, but never the other way around. The consensus mechanism works by assigning weights to transactions. Verification by the network is done through the network of indirect and direct validations. An algorithm increases transaction weights as scalability is satisfied in order to strengthen network verification. In case the weight criteria are not met in a node then it is removed. The more activity there is in the Tangle , the faster transaction validations are (Higgins & Sandner, 2019), (Popov, 2018).

In this way, the Tangle solves a number of important issues related to the execution of real-time micropayments, the authenticity and provenance of transactions, i.e. participants know where the assets came from and whether their ownership has changed. in time or if the transactions have received manipulations and security breaches by malicious actions of third parties.

At this point it is worth mentioning two features of IOTA. The first concerns the authenticity of the data streams which is achieved both publicly and privately, making use of Masked Authenticated Messages (MAM). For public authentication, only the address of the message is sufficient to allow publication to participants. When the message is private then decryption is possible only by the parties involved and is done with the help of an encryption key) to ensure data integrity and privacy (Higgins & Sandner, 2019).

The use of Flash channels (IOTA, 2017) is IOTA's proposal for instant token transactions and data streams (data streams) with no initial involvement of the Tangle in order to limit the number of validations. It is a two-way payment channel and works as follows: Participants deposit equal amounts of MIOTA in a Multi - signature Address[[33]](#footnote-34) jointly owned by the users. Then the participants carry out transactions while the Flash channel records the balances. Upon completion of transactions, the final balance is validated by only two transactions instead of multiple through the Tangle (Higgins & Sandner, 2019), (IOTA, 2017).

Recently IOTA implemented the Qubic protocol, which is a technological solution for quorum - based computing techniques computations), in outsourced computations (outsourced computations), e.g. in the cloud, as well as in smart contracts, with the latter "running" at Layer-2 on top of the Tangle (Layer-1) of the network. The ultimate goal is to automate contracts between IoT devices thus satisfying the requirements for scalability and low transaction costs(Ferdous, Biswas, Chowdhury, & Chowdhury, 2019) (IOTA, 2022).

Table 5 summarizes the main business features of the IOTA platform (Van de Ven, Abbas, Kwee, & de Reuver, 2021), (Fruhwirth, Rachinger, & Prlja, 2020), (Chowdhury M. J., et al., 2019), (Chowdhury M. J., Ferdous, Biswas, Chowdhury, & Muthukkumarasamy, 2020), (Spiekermann, 2019), (IOTA, 2019), (IOTA, 2017), ( IOTA Data marketplace ., 2017).

*Table 5: Summary of main business features of the IOTA platform.*

|  |  |
| --- | --- |
| **Value proposition**  **(Value Proposition)** | * cryptocurrency platform based on distributed ledger technology (Distributed Ledger) * IOTA micro-transactions in (near) real time to use less resources (total active nodes) for IoT/M2M communications * execution of smart contracts on top of IOTA-Tangle (from 10/2021)   + Interoperability and ease of use due to Ethereum Virtual Machine (EVM) support * Block validity and quorum-based computations using the Qubic protocol for:   + Automation of autonomous contracts between IoT devices   + Promoting Machine-to-Machine (M2M) communication (encouraging the creation of license-free open innovation ecosystems) * Secure transaction, privacy, high data integrity - Tamper proof data * IOTA tokens |
| **Dear Partners**  **(Key Partners)** | Platform Partners (e.g. Dell Technologies, Everything, Ensuresec), building facilities management companies (e.g. ENGIE), consultants and business and technology solution providers (IT security, Distributed ledger technologies, edge cloud computing, AI, robotics systems, IoT, Blockchain), financial services exchanges, working groups (e.g. FIWARE, Mobility Open Blockchain Initiative, Tangle EE, Chamber of Digital Commerce, TMforum) |
| **Categories of Customers**  **(Customer Segments)** | Businesses belonging to the vertical sectors ( connected automated mobility , MMM, smart city, digital identity , industry 4.0, health ) |
| **Business types** | Business to Business (B2B) |
| **Data Access** | Open type (distributed ledger) (via API) |
| **Data Format**  **(Data Type)** | A processed data of IoT devices e.g. sensors (and financial data from API) |
| **Platform Architecture** | Decentralized (Distributed ledger technology) - transaction centric |
| **Income Streams**  **(Revenue stream)** | * Donations (Community Donations) * Program and project grants for Research and Development |
| **Pricing** | * Zero transaction fees (and for smart contracts from 10/2021) |
| **Consent Mechanism**  **(Consensus Mechanism)** | * “The Tangle” - structure of Directed acyclic graphs to store and support a large number of transactions per sec in the distributed ledger without miner fees * No mining support * Prerequisite for reaching consensus: the validation of 2 other transactions that are considered costs for IOTA |
| **Payment in currency** | Cryptocurrency - micropayments in MIOTA (Coin Limit: 2 779 530 283 MIOTA)  *(exchanges for IOTA trading are* [*Binance*](https://coinmarketcap.com/el/exchanges/binance/) *,* [*OKEx*](https://coinmarketcap.com/el/exchanges/okex/) *,* [*Huobi Global*](https://coinmarketcap.com/el/exchanges/huobi-global/) *,* [*Upbit*](https://coinmarketcap.com/el/exchanges/upbit/) *)* |

**STREAMR**

Streamr, a decentralized **publish - subscribe system for IoT data transactions** that supports the scalability of IoT systems in near real-time while ensuring confidentiality of the data powering blockchain-based decentralized **applications** (*DApps*) (Streamr, 2017). Network-level communication between nodes is conducted through Blockchain 's peer-to-peer P2P network thus meeting the implementation's needs for fault tolerance, easy scalability, and decentralization.

To create the P2P overlay network and index trackers are used such as those in BitTorrent , which are responsible for coordinating the network's peer - to - peer connections. In essence, trackers are servers of information about users and connections. User-nodes periodically communicate with the t racker and in return get information (in the form of time-stamped messages) about other user-nodes with whom they want to connect to transact. Streamr nodes are also access points (access points) of the network, i.e. the points of connection of the decentralized applications with the data flows from the IoT devices (Savolainen, Juslenius, Andrews, Pokrovskii, & Tarkoma, 2020).

In contrast to IOTA which has developed Tangle as an alternative DLT architecture design in which the creation of blocks is not necessary, Streamr uses the Ethereum platform (Buterin, 2013), a public permissionless Blockchain , for the following purposes: a) to support the development of personalized smart contracts that will handle payments and data access control in case the latter are not freely available, and b) to securely maintain public indexes (*registries*) made up of data streams, "*trackers*" and access permissions (Savolainen, Juslenius, Andrews, Pokrovskii, & Tarkoma, 2020).

At this point it should be pointed out that in the Streamr network the storage, complex calculations and analysis of the data are done exclusively off-chain. Only certain functions such as user authentication and payments are outsourced to the Blockchain. The actual raw data remains off-chain of blocks. In the Streamr market, the sequences of time-blocked messages that record events *from* new offers, "bids" and "trades" of data are also published and recorded on the chain of blocks (on-chain).

As a general finding, this solution serves decentralized messaging and event processing (eg real-time alerts and notifications based on criticality). It also guarantees low latency, immutability, tamper proof in data delivery, event persistence and fault tolerance achieved by storing events, offering all the benefits of multicast encryption cryptography. The basis of the latter is the public and private "keys". A transaction is essentially the recipient's address, i.e. the public key attached to the amount being transferred and cryptographically signed by the sender's private key in order to allow direct access to the data only by the people involved in the transaction. Data streams in Streamr are not stored.

The owners of the streams *decide* which trackers will be used, choosing either from a set of "default trackers" maintained in the smart contracts, or by assigning their own ("custom") trackers by registering their address in the smart contract of each stream. When a node wishes to publish or subscribe to a stream, it must first find a tracker to connect to the smart contract and then connect to the rest of the nodes that make up the stream's "*topology*"*.* The tracker is exclusively responsible for controlling the connections between the nodes, which additionally applies algorithms to optimize the network paths for shorter delays (Savolainen, Juslenius, Andrews, Pokrovskii, & Tarkoma, 2020).

An overview of the functionality just described in the Streamr network is given in Figure 7 (Savolainen, Juslenius, Andrews, Pokrovskii, & Tarkoma, 2020), where “Bob” is the user-node that wishes to participate in the stream overlay.

A picture containing timeline

Description automatically generated

*Figure 7**: Example of using the Streamr platform.*

In order to replace platforms such as Azure EventHub and Azure [[34]](#footnote-35)Stream Analytics [[35]](#footnote-36)the architectural structure of the platform includes 5 levels (Voulgaris, et al., 2019), (Streamr, 2017):

* Streamr Editor: A graphical user interface that is essentially a toolbox for non-experienced users of Blockchain technology to develop Dapps and microservices.
* Streamr Engine: decentralized machine for event processing and automated decentralized analytics.
* Blockchain 's peer-to-peer network by combining features of a cloud - based and real-time data transfer system such as Kafka (Garg, 2013). In this way, high throughput is achieved for real-time data applications such as peer discovery, flow routing, etc. Core software of the network is the Broker Node (broker node). It is responsible for delivering digital goods, i.e. events, to subscription smart contracts and consumers. Also part of its job is to contribute uplink bandwidth to transfer messages/events, which significantly increases the system's bandwidth consumption. The Brokers Nodes connect to decentralized applications via APIs. Figure 8 (Streamr, 2017) shows the process of transferring events through Broker Nodes from Publisher to Client.

Diagram

Description automatically generated

*Figure 8**:* Streaming events based on the Streamr protocol .

* Streamr Data Market: On the Streamr market Ethereum platform any node/user can publish and record events from new offers, bids, trades (timestamped data) and participate in processes on it (e.g. node creation, block validation ). Is the supply of data streams primarily from connected devices, IoT sensors with reference to weather, transportation, finance, social media? conducted by Data producers and implemented in the Blockchain peer network. Consumers of the goods are, for example, car manufacturers such as Rolls-Royce, the outdoor billboard company JCDecaux, from the media sector the Finnish Alma Talent, application developers and also individuals subject to Confidentiality Agreements (Non-Disclosure Agreements - NDAs) while responsible for the performance of the data streams are the "broker nodes" (Broker nodes).
* Smart contracts: as mentioned above they are used for payments (tokenisation) and access to data streams applications. Additionally, they verify the identity of IoT devices. The terms of use are communicated to the trading parties through three processes: Ethereum Smart Management, Data Authorization and Payment contract (Sharma, Lawrenz, & Rausch, 2020).

Graphical user interface

Description automatically generated

*Figure 9**: Streamr technology stack*

*Use Cases.* Typical is the case of using a network of interconnected vehicles connected to several streams, where each of them is assigned to geographical areas that indicate the position. The network of interconnected vehicles can receive data from other vehicles belonging to the same geographic area, thus allowing drivers to receive alerts about traffic or accidents in the immediate area. Another use case supported by Streamr is smart city open data. For example, individual users can publish data from sensors measuring atmospheric air quality or temperature changes with the aim of using them to provide new decentralized services related to the microclimatic conditions of the urban space.

In general, the publish-subscribe feature in this marketplace allows for multiple types of associations between trading parties, such as one-to-many (e.g. a news channel), many-to-many (group chat or online game), one- one-to-one (analytics pipeline) and many-to-one (voting system).

The data streams are available either free of charge or for a fee from the data producers, while access to them requires a ***limited time subscription*** to them.

The function of the Streamr market has been taken over by the encrypted coupon (token) DATA coin token, an ERC20 token, which is used (a) for subscribing to "streams", (b) for rewarding with micropayments for the contribution of data publishers to the expansion of the community for the benefit of all trading parties and c) to reward the miners (Brokernodes) after the end of the bid confirmation if they have spent computing power and network bandwidth. Therefore, DATA has a dual use: it is a means of payment but also a measure of reputation. In addition to the existing token, Ether has been included as a native cryptocurrency of the Ethereum protocol (Buterin, 2013). It is mainly used for processing gases fees, i.e. the cost charges for storage, bandwidth commitment imposed by the operation of the Ethereum Blockchain as well as for performing calculations to make transactions, smart contract or applications on the platform. Ether can also be used when subscribing to said market.

Streamr leverages the time-based subscription model to generate **revenue streams.** Access to digital data goods that are free does not require the use of tokens (DATA or Ether). For those categories that are paid, it is first required to create a new Ethereum Wallet, i.e. a private-public key pair from which the address will be generated. The message is then generated to transfer Ethers to the user's account to pay the gas fees related to the consumption of computing resources on the Streamr network. Native DATA is used for transactions. The Streamr platform does not withhold transaction fees but fees are expected from the Ethereum network. Gas fees are the charges on the Ethereum network.

When the user-client now wants to gain temporary access to a decentralized application or service, then the owners of the data streams using smart contracts offer their space. Payment in the owner's DATA precedes access to them. For this purpose, Streamr offers an access interface with the help of which the client's wallet reproduces the message with information about recipient/owner (address) and for transfer of amount through the generation of the signature from the protocol between the two parties and the final transfer of the amount is confirmed by the Ethereum Blockchain. Upon completion of the transaction, the permitted access period is also defined (Bajoudah, Dong, & Missier, 2019).

**DATABROKER GLOBAL**

The Databroker Global[[36]](#footnote-37) (formerly Databroker DAO) platform which in its early stages was self-described as the “eBay” for trading real-time IoT sensor measurements has now evolved into a “Platform as a Service” (PaaS) which is not strictly connected to the buying and selling of IoT sensor data but refers to data formats of a wide variety of devices such as for example autonomous vehicles with integrated sensors, cameras, lights, security systems, etc (Databroker Global, 2020).

The platform consists of a distributed network of peer nodes (peer-to-peer/P2P) that uses Blockchain technology as a way of decentralization to ensure the authenticity and integrity of data transactions by allowing " *the direct sharing and availability of resources for each transaction, whether public or private, between the interconnected nodes-users that are autonomously organized in online topologies, without the need for a central authority* ", having the role of a broker (Van Nierkerk & Van der Veer, 2017), (Databroker Global, 2020).

The goal is to maximize the utility of IoT and any data format, to offer new sources of revenue to their owners/providers, and at the same time to make data sets available to buyers at a low cost. By supporting various types of data covering a wide range of business sectors such as environment, transport, agriculture, energy, supply chain, human resources, agriculture, economy, etc. Databroker Global hopes to enable entrepreneurs, researchers and organizations from around the world to further explore a multitude of use cases related to these activities (Databroker Global, 2020). It should be noted that at Databroker Global these eight branches of activity are grouped into so-called "communities" in order to make the bidding process and the publication of the different types of data easier to use.

Ethereum is used as a platform for the implementation of this market, on which smart contracts are executed to automate transactions and DApps can be created and run securely in order to provide accessibility and authentication for the sharing of IoT sensor data and other types (of data). The security of transactions between buyer/provider and seller is ensured in Ethereum by applying elliptic curve digital signature algorithms to encrypt them. Encrypted tokens called DTX tokens are also used for transactions, which serve as credit and at the same time as an incentive for buying and selling data/resources. Reputation mechanisms *are not* supported by this market (Van Nierkerk & Van der Veer, 2017).

By effectively combining Ethereum and the Platform-as-a-Service “white label” offering, Databroker Global guarantees its trading parties the support of large contracts and thus a larger volume of data and transactions. This means that it does not rely on microtransactions like IOTA but seeks to leverage developers' familiarity with Ethereum to develop decentralized and distributed applications on it, which will be offered as services (Databroker Global, 2020). The combined use of the platform with "smart contracts" offers an additional benefit. It helps to enhance interoperability with other platforms. Along with the adoption of the "white label" technology, important operational and commercial requirements are met which mainly have to do with a) speeding up the publication of the data set in more markets, and b) with the creation of data markets adapted to various vertical sectors of the industry , i.e. the "communities" of Databroker Global, in an easy and efficient way. Other complementary services offered through Databroker Global include the provision of consulting services as well as the free service "DataMatch" i.e. the service of matching and matching buyers/customers with potential partners and sellers. Both of these services aim to help e-auction participants discover data trading opportunities (Databroker Global, 2020), (Van Nierkerk & Van der Veer, 2017).

Regarding the *pricing process*, data owners/providers reserve the right to propose the price at which they wish to sell their data, however this sale price is variable and affected by supply and demand at each time. Specifically, providers are given the following options through Databroker Global: setting a fixed price (with or without the right to participate in an auction process), free data availability and participation in an electronic auction[[37]](#footnote-38) (Databroker Global, 2020). After each sale is completed owners receive information about who the buyers are by email as well as through the transaction record, which is maintained for each owner/provider separately, in the section “My Sales” of their personal account. The revenue streams for Databroker Global come from the *commission on the sale price* of the data, which is 10% (Databroker Global, 2020).

In this particular market, a distinction is made between the "data offers" (data offers) to be auctioned and the data itself, the so-called "products" (products) that are published. The first refers to descriptions of a data set that is for publication and sale in one of the eight supported Databroker communities, but without including prices or details of the file format in it e.g. weather data for Greece to be published to the geospatial community. Products, on the other hand, refer to the type and content of the data, e.g. temperature measurements for the period 2010-2020. Included in the data offering, which may include any number of specific data products (Databroker Global, 2020).

In the existence of these different types of products lies the differentiation with the rest of the markets under consideration and more specifically with the Streamr market. Data providers in the Databroker Global marketplace may offer static data products in addition to the real-time data streams exclusively supported on Streamr (Databroker Global, 2020). This fact adds value to the existing solution, at least at the moment considering that the majority of participants do not have the infrastructure and technological solutions for real-time data exchanges.

Regarding the transmission of data through the Databroker platform Global it is worth mentioning that the data can be provided in the form of a file, stream or API within a predetermined period of time. Additionally, it is pointed out that the data derived from the products are not stored on the platform. Instead, they are securely transferred directly from the provider to the buyer thanks to a software called “Data eXchange Controller” (DXC), which is installed on the provider. The provider, in this way, can declare all data sources in the DXC environment, which then assigns unique product IDs (Databroker Global, 2020).

DXC acts as a "proxy gateway" between the data sources and the user/buyer, where access is granted depending on the format of the data (files, data streams and APIs). So for example for the files the user receives a hyperlink while for the APIs the access is through an authenticity key. Access to the data sources through DXC is blocked after the time period defined by the providers (Databroker Global, 2020).

Table 6 (Van Nierkerk & Van der Veer, 2017) summarizes the key business features of the Databroker Global platform.

*Table 6: Summary of key business features of the Databroker Global platform*

|  |  |
| --- | --- |
| **Value proposition**  **(Value Proposition)** | * Dataset Provider / Seller & Buyer Matching: * Peer-to-Peer Network * Personalized DataMatch Service àproviding consulting services and guidance for discovering and matching data partners (free of charge) * private platform as a Service (PaaS) with White-labelled web portals offers: * launch of a new " w hite-label" data set exchange platform for each vertical * “out-of-th e -box” interoperability with other data markets for the scalability of data exchange beyond the original use case, business or organization, vertical * quick publication of the data ("only a few clicks") on the platform àprofits for the provider, development of new business models, services ("real business value from data") * Active search of users and data - transmission of anonymized data under aggregation * Dataset Thematic Content * Smart contracts using DTX tokens and Ethereum platform (Data eXchange Smart Contracts (DXS) platform)   + interoperability with other data markets through the DXS platform. * Secure transaction, privacy, high integrity and verification of private and public data (Blockchain) - without storage on the Databroker platform (use of digital signature and public key) * Static data products |
| **Main stakeholders**  **(Key Stakeholders)** | * Data Provider/Seller: organizations, businesses, natural persons (e.g. insurance, DEKO, transport sector, research institutes, MNOs, factories, automotive industry and car drivers, OTAs, telematics companies, health care providers, farmers, cooperatives, manufacturing , social networks, machine manufacturers) * Data Publisher : IoT, big data platforms that collect data on behalf of data providers and can create new business models. * Data Buyer : organizations, businesses, R&D, local authorities, press, pharmaceuticals, contractors, chemical industries, insurance, land planners, health administration, Environmental organizations, energy investors, electricity companies, educational institutes, e-commerce,, data analysts, companies, manufacturers drills |
| **Business types** | Business to Business (B2B), Consumer-to-Business/C2B  Many-to-Many association |
| **Data Access** | Open - install a RESTful API (Data Exchange Controller) on a data provider's device or server to initiate data transfer |
| **Data Format**  **(Data Type)** | Support of any format - thematic content (geographic/geospatial, environmental, transport related, agricultural, economic, energy, supply chain and anonymous data for natural persons or groups of citizens )   * data streams e.g. temperature data, events (alerts) * files e.g. zip * API flows (reducing data storage on multiple servers), e.g. API queries to retrieve a list of cities that experienced a temperature increase of more than 20% in the last 10 years |
| **Platform Architecture** | * Distributed (Ethereum Blockchain Technology)- transaction centric |
| **Income Streams**  **(Revenue stream)** | * supply of license to use the "white label" platform * Transaction fees * indirect revenue from increased market availability and attracting interested parties to join the public platform |
| **Pricing**  **(Pricing)** | * 10% commission on the sale price of the data |
| **Consent Mechanism**  **(Consensus Mechanism)** | Information not available |
| **Payment in currency** | DTX token (as data transaction medium only)  (Databroker uses Fiat currency to buy DTX when buyer buys using Fiat currency)  30 days warranty for purchase control by data buyer |

**DATASPACE**

Based in Nokia Bell Labs (France), Datapace[[38]](#footnote-39) is a decentralized data marketplace for online auctioning and trading of IoT device and sensor data streams, without meaning that any data format of any origin cannot be used.

The Datapace platform optionally offers various types of data stream processing and analysis for building IoT applications, leveraging tools such as Google and a prototype platform developed by Nokia Bell Labs, World Wide Streams (WWS) (Van Raemdonck, et al., 2017)that aims to facilitate of IoT application development by connecting geographically distributed data sources even in real time. Specifically, users of the D at a pace platform have the option of enriching their data streams, thus giving them additional value, through an optional combination of a) data analytics applications developed on the WWS platform, b) machine learning algorithms (AI/ML) offered by the Datapace marketplace itself and c) connectivity with Google's big data warehouse (the so-called BigQuery-Cloud Data Warehouse). The goal of the market is to create a global network of complex IoT sensor ecosystems as well as decentralized telemetry.

Based on Hyperledger technology Fabric the Datapace platform implements a private & permissioned Blockchain network for buying and selling IoT sensor data streams by conducting "sealed bid" auctions. In order to register in the network, it is necessary to verify the hardware of the sensors belonging to the trusted partners of Datapace through a certification software granted by the market itself. This proves the reliability, validity and source of the sensor data (Draskovic & Saleh, 2017).

Hyperledger Fabric uses encryption technique and anonymization when accessing data streams published by providers, thus achieving secure and immutable transactions. It also supports the use of a smart contract (chaincode)[[39]](#footnote-40) to verify transactions every time the ownership of data streams changes. The automated contract validates that a transaction meets the standards of the client/buyer and is executed by authorized nodes or users based on certain policy rules. Upon completion of each transaction, encrypted coupons are used which are distributed among the transacting parties according to the agreed prices. This whole process will be described in detail below.

In addition to the aforementioned, this solution makes use of the Practical Byzantine Fault Tolerance (PBFT) protocol (Baliga, 2017), a consensus mechanism used in distributed systems to combat byzantine failures [[40]](#footnote-41). This consensus mechanism has to demonstrate low latency and high throughput. It works on a majority basis where all participants are involved in validating a block and when a limit of nodes agree on a block, it is added to the network. In the Blockchain network in question, this practically means that transactions are validated only by a closed consortium (consortium) made up of entities (predefined nodes) that have a validator role, which does not allow increased scalability in terms of the number of entities but leaves room for parallel performing checks. Therefore, necessary conditions for accessing and reading the data are the consent of the members of the consortium as well as the validation of the specific node that makes the request based on the contract. This structure of the Blockchain consortium combined with the requirement for at least 2/3 consensus of the nodes of the system (as a result of the operation of the PBFT) contributes to the protection against malicious ΄Byzantine΄ nodes. Other advantages stemming from the choice of PBFT are low energy costs and almost non-existent latencies (around 1000ms on Datapace) (Cachin, 2016).

In addition, due to the implementation of the Hyperledger Fabric, a high throughput and fast finality of the transactions is achieved, so by extension the phenomenon of forking does not appear, [[41]](#footnote-42)so it is possible to perform with a certain number of users, up to 3500 transactions per second (Cachin, 2016).

Datapace uses two additional consensus algorithms to enhance network security, Proof of Stake and Proof-of-Verified-Source (Draskovic & Saleh, 2017). The first allows one of the users with the right to validate the transactions to undertake to create and sign the validity of each block while the second acts as a miner to determine from which network equipment the data flows originate. Authenticity is confirmed both in terms of time sequence and by the time stamp on each block. In this way, this particular solution attempts to offer high security, validity, privacy and confidentiality.

The interoperability of the private Blockchain, i.e. the communication with other established Blockchains such as Ethereum, is done through the Cosmos hub[[42]](#footnote-43). The platform also offers connection and management of sensors via API.

The process of trading, i.e. the online auction of the data streams within the market is followed as follows: The owners/sellers of the data streams publish a valid data source in the form of URL that is not accessible to the buyer accompanied by a detailed description of the data in JSON format and smart contracts. It should be noted at this point that in the Hyperledger Fabric architecture different smart contracts describing the different contracts of each Blockchain node are created by the vendors. Subsequently, the parties participating in the auction are informed about a specific contract and if they agree, they sign it digitally in order to complete the first stage of the transaction.

To ensure the reliability of the data, a combination of different techniques is applied, including, for example, a vendor reputation mechanism and verifying whether the hardware and firmware of IoT gateways and sensors are certified. Given that access and transfer of confidential data is required with complete security, cryptographic techniques for the digital signature of the data and access to the platform are applied - but without compromising the anonymity of the transacting members/voters.

Buyers interact with the automated smart contract (smart contract) running as code within the Blockchain to source data streams and process transactions. Each of the buyers must have a wallet with a sufficient number of TAS (Tarush) tokens , which are the encrypted tokens issued during the creation of Datapace. When performing the purchase of the data streams, a temporary "*proxied HTTP URL*" address is assigned to the buyer for the consumption of the data for the duration of the leasing period. Upon completion of the auction and without any further interventions between the buyer and other traders, the amount of TAS tokens is registered in the seller's wallet as "pay-per-use" transaction revenue. Validators are also rewarded with TAS tokens for validating the block.

It is worth noting that the Datapace platform stores references to the data, but not the data itself. No node has access to the entire data set. Due to Blockchain technology micropayments for microtransactions between interested parties are also encouraged.

Summary information of the technological solution as well as the main business characteristics of the Datapace platform is presented in Table 7.

*Table 7: Summary of key business features of the Datapace platform*

|  |  |
| --- | --- |
| **Value proposition**  **(Value Proposition)** | * Hyperledger Fabric technology cryptocurrency platform * IoT platform as a service through Mainflux technology àfor management and connectivity and IoT devices and * management of data flow collection, connection of alternative IoT devices to network gateways with the platform as an access link * (near) real-time for IoT/M2M communications and micro-transactions * Automated smart contract, introduces ERC-20 token payment system (TAS Tokens) and Secure Wallet for data trading * Cosmos hub àfor scalability, interoperability with Ethereum, Bitcoin and token interchangeability * Security (Hyperledger Fabric technology) with PBFT consensus algorithm, P2P state replication & transaction traceability * Immutability àkeeping records of executed transactions * data privacy (no storage, real-time consumption mechanism) and access control (transactions visible only to members of the same organization, encryption and anonymization of initial access) * verification and validation/verification of IoT data through HW & firmware certification program of IoT gateways, sensors, GPS modules, edge computers and partnerships * data validity (historical, timestamped cryptographic hashes) * Graphical User Interface for Smart Contract & Geofencing (interactive maps to select data streams) |
| **Main stakeholders**  **(Key Stakeholders)** | * Data Seller: organizations, businesses, natural persons * Validators : to verify and validate the block * Data Buyer : organizations, businesses - purchase data via TAS Tokens |
| **Collaborators**  **(Key Partners)** | NOKIA, HYPERLDEGER, NKN, MAINFLUX |
| **Business types** | Business to Business (B2B) |
| **Data Access** | * Open type - access via app installation on smartphone |
| **Data Format**  **(Data Type)** | * Measurements of IoT sensors, drones , autonomous cars , physical assets (e.g. geolocation data) * Data streams - read only in SeML, JSON, CBOR |
| **Platform Architecture** | Decentralized - Hyperledger Fabric (private permissioned Blockchain) - transaction centric |
| **Income Streams**  **(Revenue stream)** | Information not available ( probably low fees when managing micro-transactions) |
| **Pricing**  **(Pricing)** | * Pay-per-use data transactions (price determined by buyer) * TAS Tokens |
| **Consent Mechanism**  **(Consensus Mechanism)** | * PBFT * Proof-of-Verified-Source - mining * Proof-of-stake |
| **Payment in currency** | TAS Token  Cryptocurrency – micropayments |

* **NOKIA DATA MARKETPLACE (NDM)**

It is a data marketplace that follows the same design philosophy as the Datapace platform, due to the common business concept at Nokia of implementing a private and permissioned Blockchain network on an IoT platform. Using the open source HyperLedger Fabric as the Blockchain back-end, the present solution aims to support a wide range of use cases with different requirements for validation times, decentralization, confidentiality and security of transactions. Capable of running machine learning (AI/ML) algorithms and coupled with HyperLedger Fabric, it targets service providers and enterprises by offering reliable real-time data sets to trade and transact.

A key advantage of this implementation is that it seeks to meet the growing need for data processing and decision-making at the edge of IoT networks by running AI/ML algorithms close to IoT devices. AI/ML algorithms running on multiple decentralized IoT devices use the Federated Learning ( FL) technique to keep data local i.e. at (McMahan, Moore, Ramage, Hampson, & y Arcas, 2017)each edge and therefore limit the potential attack of malicious nodes and exchange information between each device and the NDM platform.

The FL technique works as follows. Decentralized IoT nodes train through AI/ML algorithms statistical models using their data while avoiding sharing it. Then, these models are sent to the NDM platform that plays the role of "aggregator", which updates the general statistical model which is sent back to the IoT nodes to result in the retraining of the algorithms and renewed models.

Each AI/ML algorithm has an API to communicate with and run the models on the individual devices as well as to help interact with the NDM platform. The platform supports heterogeneity in algorithms. In this context and because of the HyperLedger Fabric the models are encrypted so the problem of security and privacy between the IoT nodes in the IoT ecosystem of the market in question is solved. In addition, scalability is provided by connecting the IoT to the target device in real time, verifying the identity and therefore validity and reliability before the blocks enter the Blockchain network.

At this point it is worth highlighting how the hosting and execution of the AI/ML algorithms is carried out in infrastructure services of the commercial cloud provider Equinix (NOKIA; Equinix, 2022)(specifically in dedicated “private cages”), **in order to satisfy the high energy and computational requirements** as well as to make better handling of large datasets. This approach removes the limitations in computing resources (computing power and memory) of IoT edges, while it is very convenient for developers and contributes to easy development and application interface (API).

The process for the authenticity of transactions as well as the processing of requests in order to execute them is carried out through smart contracts that are based on micro- transactions . Smart contracts are used to automate processes when i.e. IoT devices generate revenue, data producers automatically receive a share of the profits.

From a business point of view, NDM stands out for the following innovations (NOKIA; Equinix, 2022):

* It offers "Software as a Service/SaaS" to customers whether they are Data Center providers, SMEs, organizations based on their own data. It is also responsible for the management of the various policies (policy rules) defined in the system, such as the policy of support, approval and consent but also of the personalized agreements with the interested parties such as the algorithm developers or the data service providers.
* It provides support for pre-selected business models that address specific use cases around the IoT, without excluding their expansion such as supply chain, vehicle parking, environmental metrics, digital health, smart city, intelligent energy networks, systems electric vehicle charging and intelligent ports.
* Does it act as a broker, thus implementing a brokerage business model responsible for processing requests in order to execute transactions, maintaining a list of auctioned resources, managing payments based on encrypted tokens while determining whether such transactions are valid from a business perspective. In this business model although there is no relevant reference to (NOKIA, 2021), the charges are usually made in fee or commission for each transaction .

Finally, it includes a token, approach agnostic, without supporting any specific currency or cryptocurrency. It is a private permissioned network that allows only authorized participants to access data.

The main architectural elements and business features of the NDM platform are depicted in Figure 10 and Table 8.

Table

Description automatically generated

*Figure 10**: Architecture of NDM.*

*Table 8: Summary of main business features of the NDM platform.*

|  |  |
| --- | --- |
| **Value proposition**  **(Value Proposition)** | * technology platform Hyperledger Fabric technology for secure transaction, trading of data and machine learning models àCSPs & enterprises become data marketplace providers   + Automated smart & trusted contract, introduces Digital Catalogs and token payment system * Secure sharing (private, permissioned Blockchain), transaction traceability and validity of vertical market data * data privacy and security - no storage and control only during data transaction between buyer and seller * Software as a Service (SaaS) - offered over the cloud infrastructure of Data center Providers , SMEs, Consortiums, CSPs and each vertical as their own service to end customers and partners. * use Federated AI/ML learning at the edge of the network - insights through open APIs without transferring data centrally   + Implement a machine learning model with input training datasets, for use case analysis   + Federated, geodistributed, heterogeneous analytics support |
| **Dear Partners**  **(Key Partners)** | Data Center Ecosystem partner (Equinix), Logistics partners such as Analysis Mason, Marlin), Nokia Worldwide IoT Network Grid |
| **Categories of Customers**  **(Customer segments)** | Enterprises & Joint Ventures (C.P.E., Aviation, Transport, Network Enterprises in the Electricity, Gas, Oil Sectors ), Financial Institutions, Electronic Communications Service Providers, Public Authorities |
| **Business types** | Business to Business (B2B) |
| **Data Access** | Not available info |
| **Data Format**  **(Data Type)** | Metadata - enriched data sets (rich datasets) of any type  (environmental metrics, IoT sensor metrics, drones, autonomous vehicles, traffic/transportation data, automotive supply chain data, medical data for research, equipment maintenance data, etc.) |
| **Platform Architecture** | Decentralized |
| **Income Streams**  **(Revenue stream)** | levels of tiered access to the platform based on subscription depending on the use of marketplace features & operations support |
| **Pricing**  **(Pricing)** | * “PAYG” (pay-as you grow/ pay as you go) business model (pay according to demand of resources and growth of customer needs) * 2-tier subscription model for cover of marketplace features & operations support * Tokens |
| **Consent Mechanism**  **(Consensus Mechanism)** | Information not available  (the use of "Proof-of-Authority" or the Apache Kafka mechanism is assumed) |
| **Payment in currency** | Tokens  Fiat currency (fiat money) |

**ZENODYS**

The Zenodys platform[[43]](#footnote-44) is an IoT data marketplace based on Ethereum Blockchain technology (as a smart contract platform) to create, collect and trade IoT data digital goods. The main objective of this market is to act as a data value chain by generating revenue in an automated way and therefore increasing the value of the trading medium it uses. To this end, the Zenodys platform offers multiple trading options with recipients for all trading parties. In the Zenodys ecosystem any of the interested parties can create and exchange digital data goods for exploitation by algorithms, decentralized and non-applications (DApps/apps), Augmented/Virtual Reality – AR / VR ), digital twins etc. Referring to the interested parties, these are mainly (Zenodys, 2017):

* Data buyers: companies or other legal or governmental entities that need data to improve their services, products, business models or other entities (companies of any kind, such as for example consulting, research, NGOs, marketing).
* Data Sellers/Providers: anyone willing to make data available on the market, eg private individual, business.
* Solution and product providers: companies/businesses that directly offer their services and products on the platform. They can also be data buyers or sellers who wish to increase their profit margin through the added value of data.
* Refiners: data scientists, databrokers, or other entities through which data is enriched and then made available for trading by all participants. Enrichment is usually done at the request of data buyers or in response to a business opportunity.
* Developers, such as refiners, offer services, connectors and decentralized applications ( DApps ) to the market to generate revenue, thus enhancing the value of the platform. They can also help other interested parties develop workflows and new nodes, upload them to the platform and with the help of the hashing mechanism promote them together with a price and a description.

The Zenodys platform is designed to be used by non-programming users. Combining ease of use with visualization and data analysis tools (visual development tools and analytics), offers fast and near-automated conversion of IoT data goods into applications for buying and selling. Marketing of digital data goods and applications is conducted through the use of virtual workflows ( visual workflows ) that provide a set of commands ( script ) for how and what data is collected, how it is processed (based on standardized data formats) and what is sent. The whole process is performed in the computing engine engine ) stored locally or remotely.

More specifically, the platform uses at its core graphical tools with integrated functionality for IoT which are available for their creation. These tools support the necessary interoperability, standardization and reuse of digital data goods to facilitate the development of applications, as they are considered value-added products. The individual graphical tools included in the platform are the following (Zenodys, 2017), (Rahmad, 2018):

* Connectors , objects that represent data sources for collecting, binding and interconnecting IoT devices, APIs and services. Other data sources used as connectors are databases and data files such as Microsoft Excel, SQLServer and MySQL, IoT protocols, and sensors, robots, AI/ML algorithms, applications (APIs), wearables, etc. The solution in question is designed to free the user from using code. Connectors are the most basic component of Zenodys' design philosophy (Krishna, Le Pallec, Mateescu, Noirie, & Salaün, 2019).
* A "Workflow builder" for the implementation of complex support applications (back-end) with the help of a drag-and-drop interface, where the connectors are placed while for the flow of information/work between them graphically represented connections (wires) are used.
* A user interface called “BUILDER” for creating a graphical environment for displaying IoT data (dashboards).
* The combination of distributed platform, decentralized application (DApps) and computing engine [[44]](#footnote-45)offers easy transfer of data from any node for visualization as reporting, graphical representation or data analytics or for application development.

For programmers, a visual debugger is provided as well as tools for programming, including "automatic code generation", the utilization of third-party libraries as well as the ability to develop code from any operating system. This distributed approach of the platform serves at the level of the Blockchain network as a basis for the creation of a peer-to-peer network (P2P), through which the exchange of all kinds of digital goods, data and services is allowed, without also excluding communication from Device-to-Device (D2D).

**Smart contracts** handle the processing of electronic auctions-exchanges of digital data goods reliably and securely without the intervention of third parties or the involvement of a central administrator. Running smart contracts on the Zenodys Ethereum platform ensures the authenticity of transactions in terms of data and services during the agreement between seller and buyer. This facilitates transparency and confidentiality in the revenues arising from them (transactions). The hashing mechanism is applied to each virtual workflow (visual workflow), as part of the contract, to preserve the inviolability. The license attached to each workflow is also hashed. This license defines the relationship and the parameters that must be met for the transaction to take place. From the activation of the contract it is further noted that the providers of the digital data goods receive a special cryptocurrency as a reward for their contribution. This happens whenever a digital data asset is used either individually or as part of a project.

As a **special cryptocurrency,** the CoinZZ encrypted token is used, which is based on the ERC20 standard and emerged as a result of writing special currency smart contracts on the Ethereum platform. With an initial price of 0.01 USD and a stock code (ticker) called “ZZ”, the CoinZZ token allows interoperability with digital wallets and other tokens as well as the execution of transactions on the Blockchain network.

The main transactions achieved through the specific implementation of smart contracts and coupons are summarized as follows: (a) access to the platform, DApps, the market in question as well as special discounts or data offers, (b) the payment in tokens for depositing a data offer, purchase of data through visual templates, purchase of exclusive nodes for use by developers, or qualifications and services from third parties that belong for example to the e-commerce sector, DEKOs, (c) purchase of exclusive type of services provided by providers (Service Providers) regarding Base nodes Data, distributed services etc. and (d) value when enriching visual templates by data scientists and brokers.

Zenodys marketplace and eponymous platform supports interactions between raw data (e.g. sensor data (IoT), personal, social, medical, public), metadata (e.g. files, video, images, audiovisual), digital assets (algorithms, code, AR/VR APIs) and applications (e.g. for services, DApps). Users are also given the possibility to trade historical data stored in Databases or services, real-time data transmitted directly from their sources and future data.

Regarding the function of trading in the Zenodys ecosystem, it consists of the following steps (Figure 11):

* the seller publishes the digital data goods and sets a selling price.
* users such as developers or business analysts can combine digital data assets with others that are part of the applications they have developed using the drag-and-drop interface.
* Sellers move into the market. The distribution of the coupons follows the purchase of the digital goods. After the purchase is completed, the smart contract takes care of distributing the tokens in real-time to the authors of the digital goods.

Diagram

Description automatically generated

*Figure 11:* Zenodys value chain.

**Use cases.** Zenodys marketplace and eponymous platform supports interactions between raw data (e.g. sensor data (IoT), personal, social, medical, public), metadata (e.g. files, video, images, audiovisual), digital data goods (algorithms, code, AR/VR APIs) and applications (e.g. for services, DApps). Users are also given the possibility to trade historical data stored in Databases or services, real-time data transmitted directly from their sources and future data. Some data management use cases that could benefit from Zenodys specific services are related to preventive maintenance, real-time monitoring systems, production line automation, implementation of machine learning (AI/ML) models, reduction of waste, data lakes, mass data consumption.

**Main partners** include multinational companies such as Dell, Accenture (professional services consultants), IOT valley (cluster of IoT startups) and MetaGroup (IT Services and IT Consulting). It also has a team of expert Blockchain consultants while being a member or partner of investment schemes (accelerators) such as Rockstart Accelerator (Netherlands), Techfounders (Germany), Starburst (France), WeXelerate (Austria) and Slovenian Business Angels Association H-Farm (Italy). Customers are multinational enterprises that make extensive use of IoT devices such as Alliander (energy - smart grids), Festo (industrial control and automation), RWE (energy - tendering) and KPMG (accounting ).

Regarding the model of associations between participants, all types of associations within Zenodys marketplace are eligible: many-to-many, one-to-many and many-to-one as well as one-to-one (One-to-One).

From a business perspective, this solution summarizes the benefits to customers through an 8-point value proposition that includes the following: a) standardization of digital data goods (standards) by leveraging artificial intelligence and facilitating their reuse in combination with other applications, b) confidentiality based on a smart contract, c) available graphical tools for collecting and trading digital data goods such as content, code, etc. - no programming knowledge is required from users, d) possibility to develop applications through synthesis of digital data goods and creation new revenue streams, e) simple user interface that allows developers to extend the functionality of the tool by creating their own nodes and accelerate the creation of digital data goods and applications, f) easy interoperability and simplification of transactions making use of virtual wiring (visual wiring) between devices, software, services and the Blockchain protocol, g) leveraging an advertising model to promote transactions in data, nodes and services, h) an evaluation system that detects, rewards or punishes any violation of a smart contract.

The **revenue streams** come from a small fee as an access fee to the Zenodys virtual environment and from the transaction fee, which amounts to a 10% fee for each transaction made. To reduce transaction costs, billing is monthly. Also considered as a source of income is the “*Initial Coin Offering (ICO)*”which is essentially a popular way of raising capital in the Blockchain network through the sale of encrypted token/coin to investors gaining market share. Capital raising predicts future investment returns to stakeholders. The main part of the financial support is granted for the development of the said market, infrastructure and platform technologies as well as for the sales circuit and the attraction of strategic partnerships and customers. Operating costs cover 15% of capital while around 15% is reserved for network security and the legal framework.

Finally, transaction pricing is done by applying an algorithm, which searches for and matches similar digital goods and then offers a price, thus simplifying the process of deciding on the pricing of a digital asset by interested parties (i.e. sellers, buyers, refiners, developers or service providers).

***Personal Data Markets***

**WIBSON**

The Wibson platform[[45]](#footnote-46) is active in the personal data market (Van de Ven, Abbas, Kwee, & de Reuver, 2021). It is based on Blockchain technology and provides the infrastructure of a distributed data marketplace for the secure and anonymous buying and selling of private information and data, which is checked for validity. It solves the problem of controlling personal data, at the same time giving financial incentives for their exchange (Fernandez, Futoransky, Ajzenman, Travizano, & Sarraute, 2020). It is built on a set of core principles: transparency, anonymity, fairness, resistance to censorship and the individual's absolute control over the use of their personal information. It adopts a one-to-one matching mechanism where Wibson acts as a neutral intermediary providing the Blockchain infrastructure for data exchange. Users install a smartphone app and can share some personal data, such as their Facebook account and mobile device location, and sell the data in exchange for Wibson tokens (Fernandez, Futoransky, Ajzenman, Travizano, & Sarraute, 2020).

Businesses can buy personal data directly from consumers, paying – and thus rewarding consumers – using Wibson's encrypted token. Consumers get paid for sharing access to their data when the transaction is confirmed (Stallone, Wetzels, & Klaas, 2021). It uses smart contracts to allow people to securely and anonymously share data in a trusted environment (Fernandez, Futoransky, Ajzenman, Travizano, & Sarraute, 2020). Market prices are set by data buyers, and buyers are matched with individual data sellers who are willing to sell their personal data for the set price. Wibson has implemented smart contracts to handle payments and transactions are paid in cryptocurrencies, which can be exchanged with Wibson in exchange for fiat money ( (Fernandez, Futoransky, Ajzenman, Travizano, & Sarraute, 2020).

In Table 9, a summary of the key business features of the Wibson platform follows(Van de Ven, Abbas, Kwee, & de Reuver, 2021) (Tufiș, 2020), (Travizano, Sarraute, Dolata, French, & Treiblmaier, 2020), (Yansen, 2019), (Panarello, Tapas, Merlino, Longo, & Puliafito, 2018).

*Table 9: Summary of main business features of the WIBSON platform.*

| **Value proposition**  **(Value Proposition)** | * Ethereum technology cryptocurrency platform offers legal advice and financial incentives to individuals to safely sell their personal data * validation without violating personal privacy * transparency, anonymity, fair competition (Fairness), censorship-resistant data in terms of how it is used, removal of personal privacy only after consent from the parties involved * maintaining control over personal data by natural persons * determination of purchase price by buyers as a function of changes in demand and supply * Smart contract (named “Data Exchange”) that introduces: * batch payment system ("Batch payments") through one transaction to multiple users at the same time to reduce gas costs * payment of the “ Notary ” role * Compile contracts to Ethereum Virtual Machine (EVM) digital code |
| --- | --- |
| **Main stakeholders**  **(Key Stakeholders)** | * Data Seller: natural persons * Notary : to verify the quality and trustworthiness of the data - Delegate : transaction submission service on behalf of the Data Sellers against an agreed fee/compensation (in tokens) * Data Buyer : organizations, businesses |
| **Business types** | Consumer-to-Business/C2B  One-to-One Correlation (one-to-one) |
| **Data Access** | Open type - access via smartphone app installation (Wibson DApp) |
| **Data Format**  **(Data Type)** | * Personal data (e.g. location data, LinkedIn, Facebook, Google accounts, device information, bank card transactions) * Medical data |
| **Platform Architecture** | Decentralized - Ethereum platform ("public, permissionless Blockchain")   * Saving only the transaction file (buyer-seller) |
| **Income Streams**  **(Revenue stream)** | Tiered platform access pricing packages based on monthly or annual subscription and depending on usage specifications (units of measurement: available domains & views) and response time for support (e.g. 24h, <12h, <2h)   * free pricing plan ("Up to 11k unique views per month") |
| **Pricing**  **(Pricing)** | pay-per-use data pricing (price determined by buyer)   * Wibson tokens (for each format) (eg 0.02 USD / pts. where position data 15 pts.) |
| **Consent Mechanism**  **(Consensus Mechanism)** | Information not available |
| **Payment in currency** | Cryptocurrency - Wibson (symbol “ WIB ” ) |

**AIRBLOC**

Airbloc is a distributed ad data marketplace where users have the opportunity to earn money by offering their personal data to advertisers who intend to leverage that data to run targeted marketing campaigns (Airbloc, 2017). Airbloc's platform is distributed and based on Blockchain technology for the real-time exchange of data in a transparent manner between data owners, providers and consumers. Examples of personal data exchanged on the Airbloc marketplace include: email addresses, phone numbers, names, ages, home addresses, location data, IP addresses, mobile advertising identifiers, a device's installed list of apps, device app usage history, and even behavioral data such as personal interests and preferences etc.

The Airbloc platform develops a fully controllable data supply chain through interconnected user-oriented services, applications and businesses.

* Personal data is collected after the explicit consent of users (data owners) through a Blockchain -based data authorization protocol called DAuth. Data can be audited and traced by its owner on a public Blockchain.
* Airbloc aims to create a token ecosystem that will incentivize stakeholders to maintain and grow the Airbloc data ecosystem. These tokens are called "ABL tokens" and are used to buy and sell data on the platform. Ecosystem participants who provide and process data to Airbloc will receive a revenue share from the data assets. Participants who contribute through data processing, verification or replication will receive additional contribution rewards from the network.
* To ensure that data for Airbloc is treated as an asset/good, access to data should only be provided to legitimate data consumers after purchase. Data privacy and data access of data owners are protected through proprietary Privacy Shield technology that encrypts, anonymizes and secures personal data, even though the data is traded on a public blockchain.

***Airbloc platform ecosystem*** (Airbloc, 2018)*.* The Airbloc ecosystem shown in Figure 12 can be divided into the supply side and the consumption/demand side.

**Diagram

Description automatically generated**

*Figure 12**: Airbloc ecosystem*

The roles of supply side data participants *are* as follows:

* **Data Owner.** Individual users who are the source of Airbloc's personal data. They can monitor and control their data flows and receive ABL tokens when they provide their personal data.
* **Data Provider.** Applications that can collect and sell the data with the consent of the data owners and share the revenue generated from the data with the data owners. Applications can be any kind of services over any device such as mobile phone, personal computer, smart devices, IoT devices, digital machines, etc.
* **Identity Manager.** A data tracking service or personal information management service that stores identity data on behalf of users. Identity managers can receive contribution rewards by storing data owners' identity data and informing them when their data is purchased by data consumers. Identity managers are also able to provide identity matching services.

The roles of the participants making *use* of the data on the demand side are as follows:

* **Data Consumer.** Businesses, research institutions or organizations interested in purchasing data for business intelligence, research or targeted advertising.
* **Data Processors** . Data processors are data analytics services or data science companies that specialize in refining raw data into valuable data. Anonymized, unprocessed data may be provided by a data provider for further processing by a data processor.
* **Data Relays.** A data relay is a node that supports data discovery by searching the data market. It helps data consumers find and purchase their specific subsets of data under certain filters by aggregating the data from multiple data providers.

Finally, a role that does not belong to any of the above categories and helps to verify data is the **Data Validation Node**. Data validation nodes can verify, challenge and replicate the data to receive contribution rewards.

Airbloc uses a system called *the Contribution Graph* to incentivize network participants. The Contribution Graph measures the contribution of network participants by analyzing the works of network participants and the reputation of entities (e.g. Data, Applications, etc.) associated with the participants. Contribution Rewards are then distributed according to contribution levels. Contribution can be expressed as a product **of reputation** and **participation**. Reputation represents the degree of influence of the network and the indirect credibility of each participant's activity. Contribution represents the amount of work a participant has done on the network during a 7-day contribution period. Therefore, to earn the rewards, a participant must work in the network, regardless of how high the reputation is.

**MEECO**

The MEECO platform[[46]](#footnote-47) establishes a personal data marketplace that allows users to add, organize, edit and share their personal data. MEECO provides access, control, representation and consent from the perspective of each user, enabling users (data subjects) to provide their own verified records and controlled consent.

The MEECO ecosystem consists of a MEECO client program (MEECO Client) whose functions are supported by the API-of-Me, as well as the MEECO distributed ledger technology (Meeco, 2018). The MEECO client program is available on iOS and Android devices and provides event chain, consent engine, attribute wallet and data vault functions:

* The *event chain* is an immutable record of a user's events containing the type of event, when it happened, and with whom it happened. Event types stored include: create, update, delete, view, milestone, start, stop, pause, resume, share, and share request. All events and interactions in MEECO are recorded as immutable events, providing a personal chain of behavior. The event chain is a data good/asset created for the individual.
* The value of a user's data is unlocked by sharing. MEECO's *consent engine* stores the conditions each user sets for sharing their data and keeps a record of their consent. Receiving consent is stored in the event chain. The MEECO consent mechanism can impose rules such as the duration and sharing of data in the system.
* A card is created from a template and contains specific data about a user in the *data wallet* . Data can be added in three ways; a) pushed to the wallet via an API, e.g. customer bank data b) connection to data integrations via certified APIs, e.g. social data, health data or IoT c) self-verified data, e.g. add directly to wallet.
* The MEECO *Data Vault* is a flexible API-accessible repository that provides categorization and semantic interpretation of data. The wallet stores encrypted data that is accessible in plain text only on the user's device.

API-of-Me enables control of the chain of data interactions at scale. Based on MEECO distributed ledger technology, service providers can trust the authenticity of data provided through a smart contract. A smart contract can verify the authenticity of data provided by a user using the Zero Knowledge Proof approach, without requiring knowledge of the user's details from the service provider wishing to exploit the data. Data is stored in distributed ledger technology supported by multiple master nodes responsible for processing and verifying transactions and user identity managers.

MEECO establishes a trusted ecosystem for all stakeholders participating in the system and provides incentives for participation and synergy. Stakeholders are individual users (data providers), governments, service providers (financial services, telecommunications and utilities), masternode operators and masternode supporters. MEECO's revenue streams come from user subscription fees, data stream rental from platform usage fees.

***Federated Data Markets***

This section discusses platforms that create federated IoT data markets, i.e. markets that combine data from multiple IoT platforms (members of the federation) and are exchanged within a common market.

**GAIA-X**

The GAIA-X research project[[47]](#footnote-48) provides a European data infrastructure for the development of innovative, reliable and sustainable digital economies, based on standards and open-source software. GAIA-X provides a federated platform that can bring together multiple cloud service providers and data owners in a user-friendly ecosystem that ensures reliable data exchange and the creation of new common data spaces (GAIA-X, 2021). The vision of GAIA-X is the exchange of data and services, implementing policies that ensure the preservation of ownership of them. The core architectural principles of GAIA-X include transparency, interoperability, federation, authenticity, and trust. The following technical features enforce these principles and ensure compliance with the GAIA-X vision (Otto, Eitel, Schleimer, & Lange, 2021):

* Security-by-design
* Privacy-by-design
* Facilitate federation, distribution and decentralization
* User friendliness and simplicity
* Machinability
* Semantic representation

Currently, more than 70 use cases describe scenarios that benefit from the GAIA-X platform, demonstrating the added value of the data infrastructure. Use cases span more than 10 different sectors in 17 countries covering agriculture, health, education and skills, energy, geo-information, industry 4.0, finance, mobility, smart living, public sector . All these use cases highlight the cross-sectoral needs, but also the particularities of each sector.

As shown in Figure 13 (GAIA-X, 2020), the GAIA-X ecosystem is formed by the synergy of an infrastructure ecosystem and a data ecosystem. These two ecosystems are complementary and cannot be viewed separately, yet each ecosystem focuses on different functions. The infrastructure ecosystem deals with the services provided or consumed, while the data ecosystem uses data as the primary business asset. Data and infrastructure ecosystems must be combined to enable the seamless exchange of data and services across a federated cloud architecture and an underlying interconnected network infrastructure. The synergy between these two ecosystems is realized by the federation services of the GAIA-X platform, which include, Identity and Trust, Compliance, Federated Directory and Data Ownership services. Ecosystem actors are classified into the basic roles of ***Provider*** and ***Consumer*** of data services or infrastructure. An entity can have both roles simultaneously. The role of the ***Federator*** , adopted by one or more participants in the GAIA-X platform, appears as the connecting layer that provides all the necessary services to create the federation.

**Diagram

Description automatically generated**

*Figure 13**: GAIA-X ecosystem*

The core elements of the two ecosystems are nodes, services and data. Specifically:

● Nodes refer to infrastructure components and mainly include computing resources. Nodes appear in the infrastructure ecosystem.

● Services are primarily cloud services that run using resources located on nodes, however they can also be services that leverage datasets. These services are provided by Providers and appear both in the data ecosystem and in the infrastructure ecosystem.

● Data goods are provided to Consumers through the services. This data can be searched and consumed by participants in the GAIA-X federation. The discovery of data and services is done through the Federal Directory.

Federation services are described below:

* Data needs proper *Self-Description* to be searched and located by interested Data Consumers. For this purpose, an open and transparent search algorithm is applied that helps Consumers to identify the appropriate offers from the set of registered data. It is important to note that GAIA-X aims to leverage existing standards as well as open technologies and concepts. Combining existing solutions, GAIA-X acts as an orchestrator and integrator. The Federated Catalog is an indexed repository of self-descriptions that enables the discovery and selection of providers and their services.
* The Identity & Trust service aims to facilitate verification and trust in the identities of federation participants. In particular, the federated identity model makes identities functional across domains by connecting various national and international identity providers. To ensure compliance between Service Providers and Consumers, rights and obligations for participants and onboarding and certification procedures have been established, complemented by a binding legal framework regarding service/data usage restrictions and consumer policies.
* The Sovereign Data Exchange service allows a Provider to make sovereign decisions regarding the use and processing of data by Consumers. This service enables the decentralized and controllable recording of data, while providing the possibility to enforce data usage policies through self-descriptions.
* The Compliance service includes mechanisms that ensure a high level of data protection, security of GAIA-X platform participants, transparency, privacy and interoperability. All participants should adhere to the respective policy rules for their inclusion in the platform and for the provision of services/data. Most policy rules are based on European regulatory frameworks and focus mainly on data ownership and control.

The conceptual model of GAIA-X shown in Figure 14 (GAIAX, 2019) describes the concepts required to build the data and infrastructure ecosystems.

Diagram

Description automatically generated

*Figure 14**: GAIA-X Conceptual Model*

***Business features.*** The GAIA-X platform develops a *Structured Market* for GAIA-X compliant data and infrastructure services, along with a broader policy and standards framework (Dietrich & Ferrari, 2021). The GAIA-X platform looks forward to the adoption of its approach by other digital ecosystems in order to facilitate the controlled exchange of data on the GAIA-X platform. Adopting GAIA-X's approach can facilitate expanded collaboration and potentially merging these ecosystems with the GAIA-X platform.

The GAIA-X platform supports the *Reseller business model* , where services can be discovered and selected through the GAIA-X Marketplace, and additional services provided by the platform, such as payment and contract management services, can be integrated into them, making able to collect supplies for the selected service. Commissions may include various pre-paid or non-monetary arrangements. The display of pricing and contract options is described in the GAIA-X Portal specification, which provides a web interface to the Federal Catalogs, but no mechanism is defined for managing or processing this information. The Provider of each service is responsible for defining the pricing policy, consumption monitoring, pricing policy and contract management policy with each Consumer. The GAIA-X platform also aspires to offer some assembler tools to help customers integrate their chosen GAIA-X services with orchestration tools. Service orchestration is described as a desired federation service and could benefit from GAIA-X self-descriptions, but this would require contracting and sourcing outside of GAIA-X.

No information is available on how federation services and other shared functions will be paid for. It also remains unclear how the Federators responsible for the operational implementation of the GAIA-X data ecosystem will be compensated for this activity. However, GAIA-X *members pay an annual membership fee* ranging from €5,000 to €75,000 for for-profit organizations (based on their consolidated revenues) and €2,500 for non-profit organizations. Based on the participation of more than 200 members and an assumed average membership fee of €10,000, the total revenue from platform members is in the order of €2 million.

**FIWARE**

The FIWARE platform[[48]](#footnote-49) is a software platform based on open-source standards. The main mission of the FIWARE platform is to develop an open sustainable ecosystem that encourages the adoption of open standards for the collection, access, management and exchange of data from IoT devices or other sources. The FIWARE platform provides tools and software to develop smart applications that leverage data to make smart decisions in various industry sectors such as smart cities, smart energy, smart agriculture and smart industry. The value proposition of the FIWARE platform is presented in Figure 15 (Tejado, 2018).

The FIWARE platform includes the following parts:

* **FIWARE Catalog** which includes a framework of open source software building blocks that can also be combined with software building blocks from other platforms to build smart solutions faster, easier and cheaper. A simple API allows the integration of the various frameworks, facilitates their interoperability, as well as the portability of smart solutions.
* **FIWARE Lab**,which is a non-commercial environment for experimenting with FIWARE technologies. Users can test the technology as well as experiment with their applications under development in the FIWARE lab, taking advantage of open data published by cities and other organizations. FIWARE Lab is built on top of a geographically distributed network of federated nodes leveraging a wide range of experimental infrastructures.
* **FIWARE Accelerator Programme**, which aims to promote the use of FIWARE technologies with a particular focus on SMEs and start-ups.

Diagram

Description automatically generated

*Figure 15:* *FIWARE value proposition.*

* **FIWARE iHubs Programme** , which aims to support the creation and operations of innovation centers (iHubs) around the world.
* **FIWARE Mundus Programme** , which is designed, in particular, to involve local actors in the field of Information and Communications as well as actors from the sectors that can benefit from the FIWARE platform.

The FIWARE ecosystem includes around 150 cities, 19 iHubs and strategic partnerships with organizations and initiatives such as GSMA, TMF and ETSI. The FIWARE Community includes technology providers and who contribute to building the FIWARE platform and ecosystem. Organizations committing relevant resources to FIWARE Lab, FIWARE Accelerator, FIWARE Mundus or FIWARE iHubs activities are considered members of the FIWARE community.

***Participant benefits, services and value proposition.*** The FIWARE platform offers different products, services and benefits to participants. Specifically:

* Access to a large set of **open-source software building blocks**, intelligent data models and support services (educational and business webinars, software development support, best practices, etc.).
* Join a large **community** based on open, transparent and meritocratic principles, including startups, developers, entrepreneurs, mentors, investors, students, academics, industry and public bodies.
* Participation in the **FIWARE Marketplace**, which is a one-stop-shop that gives access to a wide range of solutions supported by the FIWARE platform, technologies that can be leveraged on FIWARE compatible platforms, and services training/guiding or providing application development advice and support.

The FIWARE platform is currently funded by its members themselves, who pay subscription fees. Fees can change depending on the participation the member seeks, but are also related to other parameters, such as the size, type of organization, etc.

**FIESTA IoT**

FIESTA-IoT was a research project that was active between the years 2015 and 2018. The FIESTA project worked to integrate IoT platforms, experimental infrastructures and applications that would otherwise remain in silos. The FIESTA platform allowed its users to run experiments on multiple federated IoT experimental infrastructures in a seamless manner by creating a large-scale virtual infrastructure. The main goal of the FIESTA project was to open new horizons for the development and deployment of IoT applications and experiments at the European and global level, enabling the interconnection and interoperability of different IoT platforms. The Fiesta-IoT platform infrastructure spans several existing IoT facilities located in Ireland, the United Kingdom, Spain, France, Italy, Greece and Korea.

***Participant benefits, services and value proposition.*** Fiesta-IoT was intended to be a multi-sided platform with different value propositions for consumers and data providers, which are classified into five types of customers, namely: a) IoT infrastructure providers and experimental infrastructure owners, b) developers experimenters and software/solution integrators, c) experimenters and researchers, d) IoT application developers, and e) standards bodies. The value proposition for data consumers is based on accessing a dataset from heterogeneous domains with semantically aligned data, in real time, from real implementations. The value proposition for data providers is access to a community of potential data consumers who may be willing to pay for the data they offer.

In addition to direct and indirect funding, the FIESTA IoT platform anticipated commercial revenue from the participation of commercial enterprises in the data market of the FIESTA platform. The various types of clients of the platform interact through the data market, while the platform has the role of a broker. The platform proposes to introduce four different business models that could be applied to the IoT data market (Fiesta IoT, 2017):

* **Peer-to-peer broker service:** The platform has the role of a broker and focuses on bringing together data providers and consumers (*transaction-centric*). Consumers pay fees directly to data providers, while the platform receives a percentage of each transaction as pay-per-transaction.
* **Advertising service:** The platform again has the role of broker. The difference with the previous model is that data providers pay to advertise/publish their data. Consumers continue to consume data by paying fees directly to data providers.
* **Bait and hook subscription marketplace:** The platform offers a part of its services/data for free (freemium model), for a certain period of time, attracting data consumers to the market (bait-bait). When consumers have now discovered the services of the platform, and the free access period has ended, they are hooked and consume more services for a fee, e.g. for a subscription fee.
* **Peer-to-Peer Data Sharing Service:** This case differs slightly from the first one in that the data exchange is not based on financial transactions, but based on vouchers. Each entity that provides data earns tokens, with which it can consume platform services.

**DENODO**

Today's data management landscape is becoming increasingly complex, as data is spread across many heterogeneous data systems (data warehouses, columnar databases, specialized data stores, cloud applications, etc.) that may reside in multiple locations. This makes it difficult to offer a single view of data across business applications and ensure that policies and governance rules are applied across the entire data supply chain. The *Logical Data Web* is the vision of a unified data delivery platform that removes access to multiple data systems for business consumers, hiding complexity and exposing data in business-friendly formats, while guaranteeing data delivery according to predefined rules semantic and governance. Data *Virtualization* is the key technology to achieve the vision of the Logical Data Web. As one of the key data integration solutions designed for distributed architectures, data virtualization provides a logical level of data access across multiple heterogeneous systems in hybrid, distributed architectures. Today, the data management ecosystem is distributed in nature, so a logical layer of data access, such as data visualization, is most appropriate.

The Denodo platform, Figure 16 (Denodo, 2020), incorporates features that accelerate the delivery of data sets to business applications, in the most appropriate format for each consumer, across multiple, geographically distributed, heterogeneous systems. The Denodo platform offers support for the main data visualization use cases, such as logical analytical architectures, logical data warehouses, and data service APIs. However, it goes beyond traditional data visualization scenarios to better support new types of users and new use cases, such as data science and machine learning (ML). The Denodo platform decouples the heterogeneity of underlying data sources at lower system levels from client applications at higher levels, thereby providing seamless access to all data formats in any location to serve any informational or operational business need. Data virtualization provides consumer applications with a single logical point to access data from sources, avoiding complex, unmanageable point-to-point connections and disconnecting data consumers from data sources. At the data visualization layer it is possible to create multiple logical business views over the same data to suit the requirements of each business user and application, with specific naming conventions and providing the data in the format most appropriate for each From these. The Denodo platform is a middleware technology that makes it possible to create a heterogeneous data marketplace.

Timeline

Description automatically generated with low confidence

*Figure 16**: Denodo architecture*

As a data abstraction middleware technology, the Denodo platform consists of three layers: the Data Access (Connectivity) layer, the Data Integration (Modeling) layer, and the Data Publishing (Delivery) layer, which are discussed below:

* The Data Access layer accesses information from the various repositories and decouples the upper layers of the system from the heterogeneities of the underlying communication protocols and data formats. This layer is responsible for performing the necessary conversions on the data format and normalizing it to match the base view defined by the user. At higher levels all base views appear as relational views of the data, regardless of the underlying data source technology, which may be databases, web services, files, etc.
* The Data Integration layer allows the user to create virtual data models to expose data in a business-friendly format to data consumers, according to a predefined business model to comply with business governance policies defined by the respective organization. The data model can be built using a multi-layered approach, with different layers to facilitate governance and reuse.
* The Data Publishing layer facilitates the creation of a single point of access and interaction with the underlying data sources and abstract data views in a standardized way by exporting the views as web services. A view can be exposed to multiple interfaces of different types at the same time. Unified exposure of all data views at the data service layer leads to faster application development and greater security and reliability.

As a data abstraction layer, the Denodo platform bridges the gap between the real underlying complexity in IT systems (multiple sources, different data models, different granularity in information representation, etc.) and the data consumption needs of business users. Business users expect a clear and simple representation of their key business entities that can satisfy their business information needs. For this purpose, Denodo allows the user to define a business model that includes basic business entities, as shown in Figure 18 (Denodo, 2020).

Graphical user interface, diagram

Description automatically generated

*Figure 18 :* *User defined business model*

The Denodo platform enables business users and applications to access data and metadata through all possible access paradigms:

* **Search:** Business users and researchers will be able to perform google-style searches on all information, external and internal, relevant to the business. This method does not require the user to know the details of the available models (eg their data entities and fields). The results of a search are offered as structured data conforming to the entities defined in the business data model.
* **Browse:** The Denodo platform allows users to traverse data relationships between entities in the business model to discover relevant data. Users can dynamically configure the view to display only the information they are interested in in the results of a search. Based on their role users may have access to additional information relevant to their search. Imagine a web of relationships that users can traverse, navigate, and discover information without having to programmatically develop each view individually.
* **Reference:** The Denodo platform provides direct access to each data resource through universal, linkable identifiers. Following the principles of Denodo's internal RESTful architecture, each individual entity is mapped to a global, unique Uniform Resource Identifier (URI) with which users can access the entity's data from anywhere.
* **Querying and reporting:** The Denodo platform enables the retrieval of accurate information based on the entity schemas of the business data model.
* **Events/Notifications:** The Denodo platform reacts to input data events that generate notifications that are pushed to data consumers.

Denodo is provided as middleware to businesses and is offered in the computing infrastructure environments of major providers (AWS, Azure, etc.). Denodo's ***revenue streams*** come from fees that businesses pay to use the Denodo software, which can be annual or hourly.

# ***Referral Platforms.*** Another similar platform that adopts the method of data visualization to provide Data-as-a-Service (DaaS) is Dremio[[49]](#footnote-50). Specifically, Dremio is an open-source platform for DaaS and is one of the best tools for documenting and tracking data provenance/succession, supports all major data sources and has super-fast analytical algorithms. Finally, a corresponding platform is Teiid[[50]](#footnote-51), which specializes in data visualization in cloud computing environments.

# privacy-aware data processing (FEDERATED LEARNING AND Privacy Enhancing Techs)

# Cross-border data transfer technologies

Lately, fears that foreign governments might (mis-)use their sovereign powers to obtain illegal or disproportionate access to personal data normally protected under GDPR have been proven to be more than mere conjecture by conspiracy theorists. This, in turn, has clearly demonstrated the need to further strengthen the GDPR’s inherent mechanisms aimed at protecting personal data when such data leave the European Economic Area. Only when such protection transcends mere paperwork, can real assurance be offered to data subjects that their rights under the GDPR are truly upheld in case of transfers to third countries. Under impulse of active supervisory authorities and a courageous European Union Court of Justice, the EU has shown its clear intent these past few years to have the GDPR’s principles enforced even at the international level.

Generally speaking, the GDPR provides three different mechanisms for transferring personal data in a structural manner:

1. Adequacy decisions of the European Commission whereby the European Commission assesses the adequacy of the level of protection for personal data in a particular third country.
2. Contractual mechanisms such as the Standard Contractual Clauses[[1]](https://euc-word-edit.officeapps.live.com/we/wordeditorframe.aspx?ui=en%2DUS&rs=en%2DUS&wopisrc=https%3A%2F%2Feurodyn.sharepoint.com%2Fsites%2FTRUSTCHAIN%2F_vti_bin%2Fwopi.ashx%2Ffiles%2F56816b17376148de8ffaf174a6452a0e&wdenableroaming=1&mscc=1&hid=5F6F09A1-B084-7000-FCF4-99E06C260944&wdorigin=ItemsView&wdhostclicktime=1707329790136&jsapi=1&jsapiver=v1&newsession=1&corrid=334e7868-6bf4-4d55-aae9-aa91781b7115&usid=334e7868-6bf4-4d55-aae9-aa91781b7115&sftc=1&cac=1&mtf=1&sfp=1&instantedit=1&wopicomplete=1&wdredirectionreason=Unified_SingleFlush&rct=Normal&ctp=LeastProtected#_ftn1) proposed by the European Commission.
3. Unilateral mechanisms such as the Binding Corporate Rules whereby controllers or processors can unilaterally determine in a sort of internal rulebook how they will transfer personal data in a manner that complies with the GDPR.

While all of these mechanisms have their merits, certainly, the CJEU’s Schrems II judgement has shown that they in themselves may not always be enough. Indeed, an adequacy decision may become outdated in case of a regime change in the third country benefiting from such a decision. At the same time, governments are typically not bound by a contract concluded between two companies, let alone a unilateral declaration made by a single company. Hence, all of these mechanisms require something additionally in order to be truly effective against foreign government intrusion, something more tangible and real.

The European Data Protection Board has drafted Recommendations on measures which could supplement the aforementioned mechanisms.[[2]](https://euc-word-edit.officeapps.live.com/we/wordeditorframe.aspx?ui=en%2DUS&rs=en%2DUS&wopisrc=https%3A%2F%2Feurodyn.sharepoint.com%2Fsites%2FTRUSTCHAIN%2F_vti_bin%2Fwopi.ashx%2Ffiles%2F56816b17376148de8ffaf174a6452a0e&wdenableroaming=1&mscc=1&hid=5F6F09A1-B084-7000-FCF4-99E06C260944&wdorigin=ItemsView&wdhostclicktime=1707329790136&jsapi=1&jsapiver=v1&newsession=1&corrid=334e7868-6bf4-4d55-aae9-aa91781b7115&usid=334e7868-6bf4-4d55-aae9-aa91781b7115&sftc=1&cac=1&mtf=1&sfp=1&instantedit=1&wopicomplete=1&wdredirectionreason=Unified_SingleFlush&rct=Normal&ctp=LeastProtected#_ftn2) These measures are of a technical, organisational or contractual nature. The proposed technical measures show that blockchain solutions can play a real added value, because they include things like strong encryption, transfer of pseudonymized data only and split or multi-party processing. At the same time, blockchain technologies might just as well do the opposite (as is probably true for any technology). Indeed, the technology itself may lead to potentially unlimited cross-border transfers across the globe, particularly in globally distributed blockchain networks.

When implemented well, however, blockchain technology can ensure that during cross-border transfer of personal data such data is immutable when recorded on the blockchain and only accessible to authorized parties having the right credentials to access specific data. This prevents unauthorized third country governments from accessing the personal data recorded on the blockchain.

[[1]](https://euc-word-edit.officeapps.live.com/we/wordeditorframe.aspx?ui=en%2DUS&rs=en%2DUS&wopisrc=https%3A%2F%2Feurodyn.sharepoint.com%2Fsites%2FTRUSTCHAIN%2F_vti_bin%2Fwopi.ashx%2Ffiles%2F56816b17376148de8ffaf174a6452a0e&wdenableroaming=1&mscc=1&hid=5F6F09A1-B084-7000-FCF4-99E06C260944&wdorigin=ItemsView&wdhostclicktime=1707329790136&jsapi=1&jsapiver=v1&newsession=1&corrid=334e7868-6bf4-4d55-aae9-aa91781b7115&usid=334e7868-6bf4-4d55-aae9-aa91781b7115&sftc=1&cac=1&mtf=1&sfp=1&instantedit=1&wopicomplete=1&wdredirectionreason=Unified_SingleFlush&rct=Normal&ctp=LeastProtected#_ftnref1) Commission Implementing Decision (EU) 2021/914 of 4 June 2021 on standard contractual clauses for the transfer of personal data to third countries, *OJ L* 199, 7 June 2021, p. 31–61.

[[2]](https://euc-word-edit.officeapps.live.com/we/wordeditorframe.aspx?ui=en%2DUS&rs=en%2DUS&wopisrc=https%3A%2F%2Feurodyn.sharepoint.com%2Fsites%2FTRUSTCHAIN%2F_vti_bin%2Fwopi.ashx%2Ffiles%2F56816b17376148de8ffaf174a6452a0e&wdenableroaming=1&mscc=1&hid=5F6F09A1-B084-7000-FCF4-99E06C260944&wdorigin=ItemsView&wdhostclicktime=1707329790136&jsapi=1&jsapiver=v1&newsession=1&corrid=334e7868-6bf4-4d55-aae9-aa91781b7115&usid=334e7868-6bf4-4d55-aae9-aa91781b7115&sftc=1&cac=1&mtf=1&sfp=1&instantedit=1&wopicomplete=1&wdredirectionreason=Unified_SingleFlush&rct=Normal&ctp=LeastProtected#_ftnref2) EDPB, Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data, 18 June 2021, [Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data | European Data Protection Board (europa.eu)](https://edpb.europa.eu/our-work-tools/our-documents/recommendations/recommendations-012020-measures-supplement-transfer_en).

# 

# Data identification, data provenance, data tracking mechanisms

Data management in technology has evolved significantly with new technologies, impacting aspects like data identification, integrity, and traceability. The advancement in areas such as blockchain, AI, and IoT has revolutionized how data is identified, stored, and tracked.

The advent of AI and machine learning has enhanced the ability to identify and categorize data more accurately and efficiently. These technologies can process vast amounts of data, recognizing patterns and anomalies that might be missed by human analysis.

Blockchain technology, combined with IoT for data traceability, revolutionizes the security and transparency of data management. Its decentralized, tamper-proof nature, paired with IoT's extensive data collection capabilities, forms a powerful system for data tracking and verification. This integration is especially valuable in sectors like supply chain management and healthcare, where data authenticity, history, and real-time monitoring are critical. Leveraging blockchain in IoT environments enhances data provenance, reliability, and security, making it a key solution for complex systems requiring accurate and transparent data tracking.

The importance of **data identification** in technological contexts stems from the need to accurately recognize, categorize, and manage data in increasingly complex digital environments. Efficient data identification methods are vital for ensuring data integrity, security, and usability. These methods involve various technologies and approaches, such as metadata tagging, data fingerprinting, and AI-driven data classification systems. In sectors ranging from healthcare to finance, effective data identification is crucial for operational efficiency, compliance with regulations, and informed decision-making. Understanding and implementing advanced data identification techniques is key in leveraging the full potential of digital data assets.

The latest tools and technologies in data categorization and recognition include advanced machine learning algorithms, natural language processing (NLP), and AI-driven classification systems. These technologies enable more accurate and efficient data sorting, even in large datasets. Machine learning algorithms, for instance, can identify patterns and categorize data based on learned criteria. NLP is particularly useful in processing and categorizing unstructured textual data, while AI-driven systems offer adaptive and sophisticated ways to handle diverse data types, enhancing accuracy and efficiency in data management.

Tracking the origin and history of data, known as **data provenance**, is crucial in modern data management. It involves understanding where data comes from, how it has been processed or altered over time, and who has interacted with it. This knowledge is essential for ensuring data integrity, verifying authenticity, and maintaining compliance with regulatory standards. In fields like healthcare, finance, and scientific research, data provenance is vital for making accurate decisions, ensuring transparency, and building trust in data-driven systems. It also plays a key role in security and privacy, helping to trace the source of data breaches or unauthorized access.

Blockchain technology plays a pivotal role in ensuring data provenance, particularly in complex systems like supply chains or healthcare. Its decentralized and immutable ledger provides a transparent and tamper-proof record of data transactions, making it ideal for tracking the origin, movement, and changes of data. In supply chains, blockchain can trace the journey of a product from manufacturer to consumer, ensuring authenticity and compliance. In healthcare, it helps maintain the integrity of patient data and medical records, crucial for treatment and research. Other technologies like digital signatures and cryptographic hashing also contribute to securing data provenance.

Various mechanisms and technologies for **tracking data** include IoT and AI. IoT devices collect vast amounts of real-time data from their environment, providing invaluable insights for various applications. AI, on the other hand, plays a crucial role in analyzing this data, identifying patterns, and making predictive decisions. These technologies together enable efficient data tracking in areas such as environmental monitoring, smart homes, and industrial automation, leading to more informed decision-making and enhanced operational efficiency.

IoT and AI technologies facilitate real-time monitoring and data management across various sectors. In healthcare, they enable constant health monitoring and patient data analysis, enhancing care delivery. In manufacturing, IoT devices track production processes, while AI analyzes this data for efficiency improvements. In agriculture, these technologies monitor crop and soil conditions, aiding in precision farming. The real-time insights provided by IoT and AI are transforming sectors by enabling proactive decision-making and optimizing operations.

* [NetObjex](https://www.netobjex.com/) has introduced a smart parking solution using the blockchain IoT. This platform helps you find out a vacant place in the parking space and automates the process of making payments using crypto wallets.
* [Golden State Foods (GSF)](https://goldenstatefoods.com/) is a renowned manufacturer of food products working in collaboration with IBM to improve its business processes by making use of blockchain IoT. GSF has created a transparent ledger system that is accessible to stakeholders in real-time.
* [Telstra](https://www.telstra.com.au/) a telecommunication and media company provide blockchain IoT enabled smart home solutions that allow our home security systems to be easily managed using a remote control.
* Hyundai is a Korean company that is using blockchain IoT for their startup brand HDAC (Hyundai Digital Asset Currency), which is building its own private blockchain designed specifically for IoT.
* Filament a business startup using a chip designed to enable industrial IoT devices to work with multiple blockchain technologies. The chip provides a secure platform for decentralized interaction.

The innovative applications of blockchain IoT in various industries hold significant potential for transformation. They are reshaping sectors by improving operational efficiency, enhancing transparency, and offering better user experiences. For example, smart parking solutions streamline the parking process, food industry transparency improves supply chain management, and smart home solutions increase convenience and security. The future potential of these technologies lies in their ability to further integrate into daily operations, offering more automated, secure, and efficient systems. This integration is expected to continue evolving, driving innovation and new opportunities across industries.

Implementing blockchain IoT solutions faces several challenges and limitations. Scalability is a major concern, as blockchain networks must handle large volumes of IoT data efficiently. Security is another critical issue, especially in protecting sensitive data against cyber threats. Additionally, integrating blockchain IoT with existing systems poses technical challenges, requiring significant investment and expertise. These factors must be addressed to fully realize the potential of blockchain IoT technologies in various industries.

# Data certification/verification methods should be developed to verify the trustworthiness of the data.

Issuers

# Challenges and Limitations

# Future Trends and Innovation Potential

# NEXT STEPS in trustchain

# OC3: Economics & Democracy

# OC4: Multi chains support for NGI protocols

# OC5: Green scalable and sustainable DLTs

* What is currently in proposals

# REFERENCES

*“2019 Food & Health Survey” by Food Insight (2019)* [*https://foodinsight.org/wp-content/uploads/2019/05/IFIC-Foundation-2019-Food-and-Health-Report-FINAL.pdf*](https://foodinsight.org/wp-content/uploads/2019/05/IFIC-Foundation-2019-Food-and-Health-Report-FINAL.pdf)

*“9 Ways Blockchain IoT Union Help Elevate Your Business Value” by Valuecoders.com (2019, April)* [*https://www.valuecoders.com/blog/technology-and-apps/9-ways-blockchain-iot-union-help-elevate-your-business-value/*](https://www.valuecoders.com/blog/technology-and-apps/9-ways-blockchain-iot-union-help-elevate-your-business-value/)

*“A modular solution built on blockchain, benefiting all network participants with a safer, smarter and more sustainable food ecosystem” by IBM (2019)* [*https://www.ibm.com/products/supply-chain-intelligence-suite/food-trust*](https://www.ibm.com/products/supply-chain-intelligence-suite/food-trust)

*“A Privacy-Preserved and User Self-Governance Blockchain-Based Framework to Combat COVID-19 Depression in Social Media”* by Ma Zirui & Gu Bin (2023, April 5) <https://ieeexplore.ieee.org/document/10092736>

Accredible <https://www.accredible.com/>

Accenture <https://accenture.com/>

Aimultiple.com. What is Data Governance? Use cases, Best Practices & Tools (2023, January 1) <https://research.aimultiple.com/data-governance/#:~:text=Data%20governance%20is%20important%20because,Image%20source%3A%20Global%20IDs>

Artifacts <https://artifacts.ai/file/>

Audius <https://audius.co/>

BigSolar Co-Op <https://bigsolar.coop/>

BitPesa <https://www.bitpesa.io/>

“Blockchain-based Data Provenance for the Internet of Things” by Marten Sigwart, Michael Borkowski, Marco Peise, Stefan Schulte, Stefan Tai (2019, May 15). <https://arxiv.org/pdf/1905.06852v1.pdf>

*BMC Medical Ethics*. Data-driven research and healthcare: public trust, data governance and the NHS (2023, July 14) <https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-023-00922-z>

Builtin.com. “How Healthcare Cybersecurity Is Affected by the Coronavirus Pandemic” by Stephen Gossett (2020, May 5) <https://builtin.com/cybersecurity/hospital-healthcare-cyberattacks>

Blockcerts <https://www.blockcerts.org/>

BlockFi <https://blockfi.com/>

BurstIQ <https://burstiq.com/>

Chainanalysis <https://www.chainalysis.com/>

*“Characteristics of Blockchain and Its Relationship With Trust”,* by Vizaad Ali, Azah Anir Norman and Saaidal Razalli Bin Azzuhri (2023, February 17) <https://ieeexplore.ieee.org/document/10041154>

Circle <https://www.circle.com/en/>

Clarivate <https://clarivate.com/>

*Concentric.ai*. 9 best practices for data governance in a healthcare setting (2023, May 8) <https://concentric.ai/data-governance-in-healthcare-a-technical-overview/#:~:text=,and%20compliance%20with%20industry%20regulations>

“Consumers want traceability, transparency in their products” by Amway North America (2029, January 9) <https://amwayconnections.com/healthy-living/consumers-want-traceability-transparency-in-their-products/>

Coursera <https://www.coursera.org/>

Datagalaxy.com. *“8 Data Governance Use Cases You Should Know”* (2023, September 20) <https://www.datagalaxy.com/en/blog/8-data-governance-use-cases-you-should-know/#:~:text=1,how%20data%20should%20be%20handled>

*Dataversity.net*. The Data Trifecta: Privacy, Security, and Governance from Reactivity to Resilience (2023, April 5) <https://www.dataversity.net/the-data-trifecta-privacy-security-and-governance-from-reactivity-to-resilience/#:~:text=Data%20privacy%20focuses%20on%20protecting,groups%20%E2%80%93%20privacy%2C%20security>

Decentraland <https://decentraland.org/>

DHL <https://www.dhl.com/>

“DHL and Accenture Unlock the Power of Blockchain in Logistics“ by Accenture (2018, March 18) <https://newsroom.accenture.com/news/2018/dhl-and-accenture-unlock-the-power-of-blockchain-in-logistics>

“DHL and accenture unlock the power of blockchain in logistics” by DHL (2018, December 3) <https://www.dhl.com/global-en/home/press/press-archive/2018/dhl-and-accenture-unlock-the-power-of-blockchain-in-logistics.html>

“DHL and other supply logistics companies partner with Accenture blockchain” by Freightwaves staff (2018, March 15) <https://www.freightwaves.com/news/blockchain/dhl-and-other-supply-logistics-accenture-blockchain>

Embleema <https://www.embleema.com/>

Energiewende <https://www.agora-energiewende.org/about-us/the-german-energiewende>

EnergyHub <https://www.energyhub.com/>

Enjin <https://enjin.io/>

“Enjin Coin: All About Fake Swords, Fake Real Estate, And NFTs Near Future” by Kenneth Rapoza (2021, August 15) <https://www.forbes.com/sites/kenrapoza/2021/08/15/enjin-coin-all-about-fake-swords-fake-real-estate-and-nfts-near-future/?sh=68c31e04a7c9>

*“Ethical concerns about social media privacy policies: do users have the ability to comprehend their consent actions?”* by Annmarie Hanlon & Karen Jones (2022, August 30) <https://doi.org/10.1080/0965254X.2023.2232817>

*Estonia e-resicency* [*https://www.e-resident.gov.ee/*](https://www.e-resident.gov.ee/)

*European Centre for Disease Prevention and Control* [*https://www.ecdc.europa.eu/en*](https://www.ecdc.europa.eu/en)

*“Everything you need to know about Estonia’s e-Residency" by Expatica (2020)* [*https://www.expatica.com/global/living/self-employment/estonia-e-residency-16558/*](https://www.expatica.com/global/living/self-employment/estonia-e-residency-16558/)

*“Food manufacturing on blockchain” (2020)* [*https://www.ibm.com/blockchain/resources/food-trust/manufacturing/*](https://www.ibm.com/blockchain/resources/food-trust/manufacturing/)

*Gaine.com*. Evolving Data Privacy Regulations in Healthcare (2023, November 7) <https://gaine.com/blog/mdm/evolving-data-privacy-regulations-in-healthcare/#:~:text=,Century%20Cures%20Act%2C%20and%20HITECH>

GuardTime Health <https://guardtime.com/health>

HIPAAJournal.com “June 2022 Healthcare Data Breach Report” by Steve Alder (2020, July 20) <https://www.hipaajournal.com/june-2022-healthcare-data-breach-report/#:~:text=Over%20the%20past%2012%20months,of%2057.67%20breaches%20a%20month>

“How Sidewalk Labs is helping make cities more sustainable in 2022” by Prem Ramaswami (2022, August 23) <https://blog.google/outreach-initiatives/sustainability/how-sidewalk-labs-is-helping-make-cities-more-sustainable-in-2022/>

“How walmart brought unprecedented transparency to the food supply chain with hyperledger fabric” by Hyperledger (2018) <https://www.hyperledger.org/case-studies/walmart-case-study>

IBM <https://www.ibm.com/>

IBM Food Trust <https://www.ibm.com/products/supply-chain-intelligence-suite/food-trust>

“India Has Been Collecting Eye Scans and Fingerprint Records from Every Citizen. Here’s What to Know” by Billy Perrigo (2018, September 28) <https://time.com/5409604/india-aadhaar-supreme-court/>

“India’s Top Court Limits Sweep of Biometric ID Program” by Vindu Goel (2018, September 27) <https://www.nytimes.com/2018/09/26/technology/india-id-aadhaar-supreme-court.html>

Interational Renewable Energy Agency (IRENA) <https://www.irena.org/>

<https://www.ibm.com/products/supply-chain-intelligence-suite/food-trust>

Maersk <https://www.maersk.com/>

*MidLedger.com* <https://www.mediledger.com/>

*Microsoft.com*. Governance for the retail industry (2022, January 12) <https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/industry/retail/govern>

MIT's "Sustainable Energy Initiative" <https://energy.mit.edu/>

“LO3 Energy Deploys Pando Software to Maximize Renewable Asset Use in Australia” by Justin Williams (2022, March 24) <https://www.globenewswire.com/news-release/2022/03/24/2409576/0/en/LO3-Energy-Deploys-Pando-Software-to-Maximize-Renewable-Asset-Use-in-Australia.html>

Open Government Partnership <https://www.opengovpartnership.org/>

*Patientory.com* <https://patientory.com/>

Provenance <https://www.provenance.org/>

“Provenance supporting traceability for the 2020 International Woolmark Prize finalists” by Provenance (2019, October 30) <https://www.provenance.org/news-insights/provenance-supporting-traceability-for-the-2020-international-woolmark-prize-finalists>

Ripple <https://ripple.com/>

*Savedelete.com*. How to Use Data Governance to Drive Data-Driven Decision Making: A Case Study for Data Scientists (2023, November 7) <https://savedelete.com/news/technology/how-to-use-data-governance-to-drive-data-driven-decision-making-a-case-study-for-data-scientists/464880/#:~:text=Updated%3A%20October%2028%2C%202023%20How,managed%20data%20pipelines>

*Sganalytics.com*. Governance in EdTech: How Institutions Can Ensuring Data Privacy and Ethical Practices (2023, September 5) <https://us.sganalytics.com/blog/governance-in-EdTech-ensuring-data-privacy-and-ethical-practices/#:~:text=,concerning%20data%20privacy%20and%20security>

SideWalkLabs <https://www.sidewalklabs.com/>

Siemens Energy Automation <https://www.siemens.com/global/en/products/energy/energy-automation-and-smart-grid.html>

“Siemens Invests in LO3 Energy, Making Blockchain a Piece of Its Microgrid Strategy” by Katie Fehrenbacher (2017, December 26) <https://www.greentechmedia.com/articles/read/siemens-invests-in-lo3-energy>

Siemens Smart Grid Technologies <https://www.siemens.com/global/en/products/energy/energy-automation-and-smart-grid.html>

Singapore's Smart Nation Initiative <https://www.smartnation.gov.sg/>

“Singapore smart nation initiatives and possible opportunities” by SCS (2021) <https://www.scs.org.sg/articles/smart-nation-singapore>

“Slow-and-Steady Waymo Is Winning the Self-Driving Race” by Dave Lee (2024, January 9) <https://www.advisorperspectives.com/articles/2024/01/09/slow-steady-waymo-winning-self-driving-race>

Student1 <https://www.student1.org/>

*“Technology and regulation 2020”* by Leenes Ronald Leenes & Aaron Martin (2021, July 4) <https://www.jstor.org/stable/community.34023115>

Tesla <https://www.tesla.com/>

“Tesla Achieves Impressive Sustainability Goals in 2022 Impact Report” by Kevin Armstrong (2023, April 25) <https://www.notateslaapp.com/news/1369/tesla-achieves-impressive-sustainability-goals-in-2022-impact-report>

“Tesla recalls more than 1.6 million cars in China over problems with Autopilot, locks“ by Ashley Capoot (2024, January 5) <https://www.cnbc.com/2024/01/05/tesla-recalls-over-1point6-million-cars-in-china-over-autopilot-locks.html>

“The main stories of Germany's Energiewende” by Ruby Russel and Julian Wettengel (2019, February 6) <https://www.cleanenergywire.org/factsheets/main-stories-germanys-energiewende>

Unique Identification Authority of India <https://uidai.gov.in/en/>

Veem <https://www.veem.com/>

Walmart <https://www.walmart.com/>

Waymo <https://waymo.com/>

“Waymo's Self-Driving Cars: How They Work” by Jeremy Laukkonen (2020, February 13) <https://www.lifewire.com/waymo-self-driving-cars-4171314>

“Waymo’s Robotaxis Are Hitting The Highway, A First For Self-Driving Cars” by Alan Ohnsman (2024, January 8) <https://www.forbes.com/sites/alanohnsman/2024/01/08/waymos-robotaxis-are-hitting-the-highway-a-first-for-self-driving-cars/?sh=11fe5aa923a3>

“What is Audius? The Decentralized Music Sharing and Streaming Service” by Ekin Genç (2021, August 23) <https://decrypt.co/resources/what-is-audius-the-decentralized-music-sharing-and-streaming-service>

“What Is Decentraland?” By Michelle lodge (2023, December 31) <https://www.investopedia.com/what-is-decentraland-6827259>

Woolmark <https://www.woolmarkprize.com/>

"The Algorithmic Foundations of Differential Privacy. Foundations and Trends in Theoretical Computer Science" by Cynthia Dwork and Aaron Roth (2014) http://dx.doi.org/10.1561/0400000042

"A Comprehensive Survey on Local Differential Privacy toward Data Statistics and Analysis. Sensors" by Teng Wang (2020) https://doi.org/10.3390/s20247030

"k-Anonymity: From Theory to Applications. TRANSACTIONS ON DATA PRIVACY" by De Capitani di Vimercati (2023) https://hdl.handle.net/2434/954133

"A Survey on Homomorphic Encryption Schemes: Theory and Implementation. ACM Comput. Surv." by Abbas Acar (2019) https://doi.org/10.1145/3214303

"Secure multi-party computation: Theory, practice, and applications. Information Sciences" by Chuan Zhao (2019) https://doi.org/10.1016/j.ins.2018.10.024

"Uc non-interactive, proactive, threshold ecdsa. Cryptology ePrint Archive" by Ran Canetti (2020) https://eprint.iacr.org/2020/492

"A Survey of Blockchain From the Perspectives of Applications, Challenges, and Opportunities. IEEE Access" by Ahmed Afif Monrat (2019) https://doi.org/10.1109/ACCESS.2019.2936094

1. <https://www.iota.org/> [↑](#footnote-ref-2)
2. <https://www.caruso-dataplace.com/> [↑](#footnote-ref-3)
3. <https://oceanprotocol.com/> [↑](#footnote-ref-4)
4. <https://inrix.com/> [↑](#footnote-ref-5)
5. <https://aws.amazon.com/data-exchange/> [↑](#footnote-ref-6)
6. <https://www.dawex.com/en/> [↑](#footnote-ref-7)
7. <https://datarade.ai/> [↑](#footnote-ref-8)
8. <https://otonomo.io/> [↑](#footnote-ref-9)
9. <https://wibson.io/en> [↑](#footnote-ref-10)
10. <https://cloud.google.com/marketplace> [↑](#footnote-ref-11)
11. <https://azure.microsoft.com/en-us/services/data-catalog/> [↑](#footnote-ref-12)
12. <https://www.advaneo.de/en/> [↑](#footnote-ref-13)
13. <https://dih.telekom.net/en/> [↑](#footnote-ref-14)
14. <https://www.airbloc.org/> [↑](#footnote-ref-15)
15. <https://www.meeco.net/> [↑](#footnote-ref-16)
16. <https://datum.org/> [↑](#footnote-ref-17)
17. <https://www.amo.foundation/> [↑](#footnote-ref-18)
18. <https://burstiq.com/> [↑](#footnote-ref-19)
19. <https://dataeum.io/> [↑](#footnote-ref-20)
20. https://www.caruso-dataplace.com/ [↑](#footnote-ref-21)
21. data harmonization: the gathering of data of various file format specifications and naming conventions (naming conventions) from different sources and their subsequent conversion into one coherent data set in a format comparable to the analysis services of a data market and therefore useful for buying and selling. [↑](#footnote-ref-22)
22. data sourcing: is the ability to extract and integrate data from multiple sources. [↑](#footnote-ref-23)
23. (https://www.dawex.com/en/solutions/generate-new-revenue-streams/) [↑](#footnote-ref-24)
24. <https://oceanprotocol.com/> [↑](#footnote-ref-25)
25. [https :// www . Advaneo . de / de / data - marketplace](https://www.advaneo.de/de/data-marketplace) [↑](#footnote-ref-26)
26. Data-centric: It is the platform that provides tools for analyzing, processing, and visualizing data by creating metadata. [↑](#footnote-ref-27)
27. <https://dih.telekom.net/en/> [↑](#footnote-ref-28)
28. <https://www.cloudera.com/> [↑](#footnote-ref-29)
29. OPC Unified Architecture: is a machine-to-machine communication protocol for industrial automation developed by the OPC Foundation <https://opcfoundation.org/about/opc-technologies/opc-ua/>. [↑](#footnote-ref-30)
30. <https://origintrail.io/> [↑](#footnote-ref-31)
31. <https://otonomo.io/platform/> [↑](#footnote-ref-32)
32. <https://www.caruso-dataplace.com> [↑](#footnote-ref-33)
33. Multi-signature address: is a Bitcoin-like digital signature associated with more than one private key. It requires more than one private key to sign i.e. authorize the transaction in order to be added to the blockchain (it is usually m =2 keys required out of a total of n =3 keys). It finds application in cryptocurrencies, contract execution and private sales of digital data goods. [↑](#footnote-ref-34)
34. Azure Eventhub. Retrieved January 3, 2022, from: <https://docs.microsoft.com/en-us/azure/event-hubs> [↑](#footnote-ref-35)
35. Azure StreamAnalytics. Retrieved January 3, 2022, from <https://azure.microsoft.com/en-us/services/stream-analytics> [↑](#footnote-ref-36)
36. <https://www.databroker.global/> [↑](#footnote-ref-37)
37. <https://www.databroker.global/help/selling-data> [↑](#footnote-ref-38)
38. <https://datapace.io> [↑](#footnote-ref-39)
39. It is a piece of code that describes functions and events for a transaction. [↑](#footnote-ref-40)
40. The network contains faulty processes that do not stop but send messages with any content, regardless of the instructions of the distributed algorithm. The state of a faulty process can change without a message. [↑](#footnote-ref-41)
41. During this specific process, two different nodes wish to both place a block in the chain at the same time, so it splits into two creating two sub-chains (forks). Then all the nodes of the network that did not participate in the problem are invited to choose one of the two subchains, which will be considered valid to continue the process. [↑](#footnote-ref-42)
42. Cosmos Network: considered the Internet of blockchains, a decentralized network of independent blockchains that offers interoperability and exchange of tokens and information (e.g. for scalability, governance issues) and is designed 1) to help developers build blockchain networks for use cases and 2) to operate itself as a very low-fee and instant transaction verification marketplace. Cosmos Hub: is the first of the interconnected Cosmos Network blockchains with its own cryptocurrency/token, ATOM (https://cosmos.network/). [↑](#footnote-ref-43)
43. <https://www.zenodys.com/> [↑](#footnote-ref-44)
44. It is an agent with high data processing capabilities that runs on any operating system, device, gateway , smart phone, cloud, chip , distributed computing systems. [↑](#footnote-ref-45)
45. https://wibson.io [↑](#footnote-ref-46)
46. [https :// www . meeco . me /](https://www.meeco.me/) [↑](#footnote-ref-47)
47. [https :// www . data - infrastructure . eu / GAIAX / Navigation / EN / Home / home . html](https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html) [↑](#footnote-ref-48)
48. [https :// www . firmware . org /](https://www.fiware.org/) [↑](#footnote-ref-49)
49. <https://www.dremio.com/> [↑](#footnote-ref-50)
50. <https://teiid.io/> [↑](#footnote-ref-51)