REPORT

GOVERNANCE AND BUSINESS MODELS FOR SUSTAINABLE HEALTH DATA COLLABORATIVES

Belgium, July 2025

The Data Tank

Impact Licensing Initiative

IMPACT LICENSING INITIATIVE

The Data Tank

Table of contents

Table of contents **2**

Executive summary **3**

1. Introduction **5**

2. Methodology **6**

3. Regulatory context in Europe **7**

- **3.1 The EU Data Governance Act and the Data Act** **7**
- **3.2 Towards a European Health Data Space (EHDS)** **9**
  - **The Belgian university hospitals common position: a P + 4P analysis** **12**

4. Public perceptions and meaningful engagement: towards building social and impact licences **16**

- **4.1 The social licence approach** **16**
- **4.2 The role of impact licensing** **17**

5. An overall framework for situating emerging data collaboratives in the playing field **19**

6. Key dimensions in exploring a business model **20**

- **6.1 Main challenges** **24**

7. Case studies **26**

- **7.1 Data cooperatives** **27**
  - **MIDATA** **27**
  - **Salus Coop** **27**
  - **The Holland Health Data Cooperative (HHDC)** **28**
- **7.2 Health records banks** **28**
  - **UK Biobank** **29**
  - **OpenHumans** **30**
  - **Patients Like Me** **30**
  - **Unforgettable** **30**
- **7.3 Data trusts** **31**

8. Pathways for a business model for data collaboratives **33**

- 8.1 Impact licensing as strategic instrument 35
- 8.2 Scenarios for the financial model 35
- 8.3 Next steps for data collaboratives 36

9. Conclusion and recommendations **39**

References **41**

EXECUTIVE SUMMARY

Executive summary

This report develops an overview of mechanisms and components of business and data governance models for health data collaboratives. In this report, we use the terms ‘data collaborative’ or ‘ecosystem for data reuse’ indistinguishably to refer to collaborations between different stakeholders across multiple sectors to exchange data in a way that overcomes silos to create public value (Susha et al., 2017).

This report synthesises a rapid literature review of academic, policy, and industry documents, including case studies, to examine governance and business models for health data reuse. We examine in the literature different dimensions involved in sustaining an ecosystem for data reuse. The report sets these out in the current regulatory context. It also considers the role that mechanisms like a social license and impact licensing play in the sustainable governance of the different business models as essential complements to the regulatory context. It analyses case studies that can be mapped onto these models and offers pathways for a process to decide on a business model.

We conclude with the following recommendations:

- Business and governance models for data collaboratives depend on the added value that all stakeholders -both on the supply and the demand-side of data- assign to these data.

  Recommendation 1: We recommend that all stakeholders in a data collaborative explicitly articulate the value that data reuse offers to them and therefore what they seek out of the exchange and how they can contribute to the public interest.

- While a data collaborative can serve different purposes, it is important for the governance and business model that there is agreement between stakeholders and clarity on the main purpose of the entity.

  Recommendation 2: We recommend that all stakeholders in a data collaborative also agree on the main value proposition that their partnership is designed and built to offer.

- A business model goes hand-in-hand with the values and principles that underpin the governance of the collaborative and its approach to decision-making.

  Recommendation 3: We recommend that all founding stakeholders of a data collaborative are in agreement of a charter or Code of Practice that sets out the values and principles underpinning its aims and governance.

- The regulatory context sets some guidelines but, ultimately, the reuse of data for secondary uses (that is uses other than those for which the data was created in the first place) need additional layers of individual and collective consent that address the barriers of individual-only, binary and static consent processes.

  Recommendation 4: We recommend that any data collaborative puts in place mechanisms for participation of the relevant publics and stakeholders to acquire legitimacy and collective *consent for the reuse of data for different purposes other than the ones for which the essential individual informed consent is obtained.*

Recommendation 5: We recommend that a data collaborative considers the role of contracts and agreements with well-defined objectives for the use of data (ie. impact licenses) for potential different users of the data.

- The financial model needs to be aligned with the value proposition of the data collaborative as well as the ecosystem.

Recommendation 6: We recommend that data collaboratives articulate the type of model (transactional, relational, systematic) that is most aligned with its value proposition, and maps out the different sources of income (public funding, private funding, membership revenue, commercial revenue, etc.) and its likelihood in the context of the ecosystem of the data collaborative.

01

INTRODUCTION

1. Introduction

Sustaining ecosystems for data reuse is a complex challenge. Many ecosystems for data reuse, including data intermediaries and data collaboratives, struggle to sustain themselves with an approach that goes beyond public funding (Ponti et al., 2024; Krewer and Warso, 2024). In addition, the semantics of ‘a business model’ may be deceiving in a context of a data reuse ecosystem that aims to improve research and wellbeing. Talking about value creation and value creation models might be more appropriate as these semantics point towards an approach which is accommodating for contributions from all stakeholders.

This report presents the reader with the following flow:

- The regulatory context for health data reuse in Europe
- Insights from a rapid review of literature on key components for business and governance models
- Pathways towards adopting business models for data collaboratives
- Conclusions and recommendations

We hope this report equips emerging data reuse ecosystems with a lens to make sense of value creation and how this could translate into a more sustainable data collaborative. Data reuse ecosystems bring together different stakeholders with an interest or stake in data, be it the supply or demand side of it- and with different incentives and resources. Contexts also vary and there are no one-size-fits all approaches to making them work, especially when they involve health data, which is particularly sensitive. Therefore, this report should be considered as a guide and a working document aimed at emerging or young data collaboratives or intermediaries to honing in on agreement within their stakeholders for operationalising their foundational principles into a platform and value creation model that works in practice.

02

METHODOLOGY

2. Methodology

To inform this report, ILI and TDT conducted a rapid review of academic and grey literature related to governance and business models on ecosystems for data reuse. The review included 40 sources consisting of peer-reviewed academic literature, policy documents, and industry papers to provide insights into existing business models and emerging trends in healthcare data governance, including case studies. While we started focusing on sources related to health data, we expanded this to include lessons learnt from other types of data intermediaries, such as those used to reuse data for climate and environmental change. The references are listed at the end of this document.

We synthesised the literature to identify recurrent themes related to:

(i) Key components of a business model for ecosystems for data reuse
(ii) Key components of governance models for ecosystems for data reuse and their relationship with business models
(iii) Defining the regulatory landscape and the implications for ecosystems for data reuse.
(iv) Case studies of ecosystems for data reuse that reflect different business and governance models

The ILI and TDT then held a series of working sessions to design pathways for business models based on these insights. To establish these pathways we considered:

- A P + 4P framework for the governance of ecosystems for data reuse (adapted from Marcucci et al., 2023; Fritzenkötter et al., 2022) which we applied to map out the key dimensions of business models.
- The need for a three-tier approach to consent for the reuse of data: (i) essential individual consent, (ii) meaningful collective consent and participation in decisions on data reuse (a ‘social license’ for data reuse), and, finally, (iii) an impact license, which refers to the practice of establishing contracts and agreements having well-defined objectives for the use of data.

03

REGULATORY CONTEXT IN EUROPE

3. Regulatory context in Europe

Even if the digital economy in Europe is making inroads at an unprecedented pace, a key component of its success, the reuse of data, is still facing significant challenges. In this regard, recent research has highlighted that 'the reuse of open data in Europe is still in its infancy and requires further research and implementation efforts to fully express its potential for public and private value creation' (Ferro & Pizzamiglio 2023). Data silos, in particular, have also been identified as a barrier to achieve a fully fledged digital economy (Azcoitia & Laoutaris 2022).

In 2020, the European Commission made the case for an European Data Strategy (European Commission, 2020), with the aim to create "governance structures for handling data and to increase its pools of quality data available for use and reuse". This strategy encompasses a set of acts and regulations, such as the Data Governance Act, the Data Act or domain-specific spaces, such as the European Health Data space (EHDS). In April 2025, and as part of its new AI Continent Action Plan, the European Commission announced a new Data Union Strategy to make more data available for AI development and innovation (European Commission, 2025).

3.1 The EU Data Governance Act and the Data Act

The European Data Strategy, launched in 2020 envisions a single EU market for data, where data can flow securely across sectors and borders. The Data Governance Act (DGA) and the Data Act (DA) are two of its foundational pillars, each addressing distinct but complementary aspects about data sharing and data governance in the digital economy.

First, the DGA offers a new framework for data intermediaries to facilitate data sharing between data holders and users, introduces the concept of 'data altruism' for sharing data for public good, and sets up mechanisms for the reuse of sensitive public sector data. In sum, it provides the regulatory infrastructure for data sharing, notably the trust mechanisms, the organisational frameworks, and the governance structures for responsible and effective data sharing.

The DGA regulates three main types of data organisations or entities through different chapters, and establishes different requirements for each. Relevant to this report are the requirements for data intermediaries. From a structural perspective, the DGA establishes that data intermediaries (Chapter III) must be legally separate entities focused on data intermediation, cannot use the data for purposes other than placing it at the disposal of data users, and must remain neutral in their data handling. Therefore, they have to operate under a clear legal structure, one that is separate from other commercial activities that can be undertaken. As Micheli et al. add:

> If a data intermediary aims to conduct data analytics and derive insights and data-driven services from such data, this activity should take place through a separate legal entity. Business models that abide by the DGA requirements might present challenges in terms of

> economic sustainability and outcomes. For instance, if certain data intermediaries need to maintain high membership fees to sustain themselves, this could relegate them to a niche service, accessible only to well-resourced individuals or larger companies.' (Micheli et al., 2023:63).

According to the European Commission (2024), these rules and conditions are 'a deliberate means to create trust of the users of such service'. Also, 'the commercial terms (including pricing) for the provision of intermediation services should not be dependent on whether a potential data holder or data user is using other services.' (idem, 10/17).

The other type of organisation that the DGA regulates is the 'data altruist organisation' (Chapter IV). Under the DGA, these are entities that enable and facilitate voluntary data sharing for the public good and other general interest objectives (to be determined by national laws). Data altruist organisations operate on a non-profit basis, collecting and processing data that individuals and companies voluntarily donate without seeking compensation. To enhance transparency and trust, the DGA establishes a formal recognition system where these organisations can register with a central EU-level register of those recognised entities. This formal structure is supported by specific regulatory requirements, obligations, and oversight mechanisms designed to ensure these organisations operate transparently, securely, and ethically. The framework aims to create a trusted ecosystem for data donation that can benefit society through research, innovation, and public services while maintaining high standards of data handling and protection.

Data altruism organisations under the DGA will need to find a business model that balances operational sustainability with their public benefit mission. While they cannot profit from data sharing itself, they can charge fees to cover operational costs, receive grants and public funding, and accept donations to maintain their infrastructure and activities. Their value proposition centres on facilitating data sharing for societal benefit, supporting research, innovation, and public service improvement. Arguably, this model will require significant investment in technical infrastructure, expertise, and security while maintaining relationships with various stakeholders including data donors, users, funding partners, and regulatory authorities. Success will depend on building trust and demonstrating clear societal impact, while maintaining operational efficiency.

Second, the Data Act represents another cornerstone of the EU's data strategy, establishing comprehensive rules for fair access to and use of data generated in the EU across all economic sectors. The DA aims to unlock the value of data generated by connected devices (IoT), clarifying who can create value from such data, and establishing conditions for data sharing. The Act introduces provisions that enable users to gain access to data generated by their devices, facilitate switching between cloud services, and establish safeguards against unlawful data transfer to non-EU states. It complements the DGA by focusing on the practical aspects of data sharing, that is, specifying how data should be made available, under what conditions, and with what protections.

As opposed to the DGA, the DA does not create new organisations or institutional structures. Instead, the DA focuses on establishing rules, rights, and obligations for existing actors in the data economy. It defines how different entities –manufacturers of connected products, service providers, data holders, data recipients, users, public sector bodies, data processing services (e.g. cloud providers)–should behave and interact regarding data access and sharing. When it comes to healthcare and medical data, the DA explicitly excludes personal data (which remains under GDPR governance), and recognises that healthcare data often falls under specific sectoral regulations. However, it does impact certain aspects of healthcare data, particularly regarding connected medical devices, where manufacturers must make non-personal device data accessible to users and healthcare providers. This may include technical data about device performance, usage patterns, and efficiency data, while maintaining strong protections for sensitive health information.

The DA also establishes frameworks for handling non-personal health data, such as machine-generated data from medical devices, anonymised operational data from healthcare facilities, and environmental monitoring data in healthcare settings. Importantly, it includes provisions for public sector access to data in cases of public health emergencies, healthcare system planning, and emergency response (Chapter V). This approach aims to enable beneficial data sharing that could improve healthcare delivery and research while working alongside, rather than replacing, existing healthcare data regulations, ensuring that sensitive health data remains protected.

3.2 Towards a European Health Data Space (EHDS)

The European Health Data Space (EHDS) is the first of these data spaces proposed by the European Commission and approved in April 2024 by the European Parliament. The EHDS is a landmark agreement that should enable EU citizens to access their health data and share it with their healthcare professionals across EU countries, if needed.

The text adopted lists in (Article 34(1)) the purposes for which electronic health data can be processed: (a) public interest activities in public and occupational health; (b) policy making and regulatory activities for public sector bodies; (c) statistics (national, multi-national and Union level official statistics); (d) education or teaching activities in health or care sectors, (e) scientific research related to health or care sectors that aims to benefit end-users (that is, development and innovation activities for products or services, and training, testing and evaluating of algorithms, including in medical devices, in-vitro diagnostic medical devices, AI systems and digital health applications; and (h) improving delivery of care, treatment optimization and providing healthcare, based on the electronic health data of other natural persons (EC 2024). The EHDS, in sum, is the first common data space to emerge from the EU Strategy for Data, with the aim to ‘foster a genuine single market for electronic health record systems’ while facilitating both primary and secondary uses of health data (EHDS 2024).

Under this framework, there is a requirement for Member States to designate health data access bodies (HDABs) that will enable and oversee the secondary use of health data. Each Member State must establish at least one HDAB as a centralised permission-granting authority and coordinator for health data access requests. The HDABs are tasked with evaluating data access applications, ensuring compliance with data protection regulations, and providing secure processing environments for data analysis.

At this point in time, the EHDS presents different technical, practical, and legal challenges as identified by multiple industry, professional, or civil society stakeholders. While it is out of the scope of this report to review them all, it is important to highlight some legal concerns. Thus, it has been noted that the EHDS does not consistently clarify its interplay with related legal frameworks (e.g. GDPR, Data Governance Act, Data Act, and Medical Devices Regulation) (EUCOPE 2024) or that it could even undermine GDPR transparency requirements by introducing waivers related to provision of individual-level information to data subjects (Art. 7), and disfavoring consent as a legal basis for sharing electronic health data (Art. 33.5) (Marelli et al, 2023). From a governance perspective, there are questions about how the functions of HDBAs will align with other oversight bodies such as Ethics Review Boards (ERBs) and Data Protection Authorities (DPAs), or about effective and responsible data stewardship (idem).

At the time of writing, the EHDS2 Pilot project is building and testing a first version of the European Health Data Space (EHDS) by creating a cross-border network infrastructure and developing guidelines for data and metadata standards, data quality, data security and data transfers. The legislative timelines are estimative but we can expect that the EHDS will be finally adopted and published in the Official Journal by the end of 2024 or early 2025, opening up the transition and application periods. In any case, aligning a data collaborative platform infrastructure, service, data and metadata standards, and regulatory requirements will be key to advance its objectives and selected business model.

3.3 Belgium as a case study: The interplay between the GDPR and national law

In line with the GDPR and the Belgian Law on the Protection of Natural Persons with regard to the Processing of Personal Data, data related to health are considered as “sensitive personal data” or a “special category of personal data”. As such, its processing requires a valid legal basis, such as informed consent, medical diagnosis by a healthcare professional, or public health reasons (Articles 6 and 9 of the GDPR). The GDPR and the national law, therefore, remain the overarching regulatory framework when it comes to access and processing of personal health data.

As Van Obberghen et al. (2024) have specifically detailed for the Belgium case: ‘Employees of the medical devices industry or health app providers may be in direct contact with patients and process their personal data. Under the GDPR, one may only process personal health-related data when one of the grounds of art. 9.2 applies. Personal data may be processed for purposes of preventive or occupational medicine, medical diagnosis or the provision of health or social care treatment, but this may only be done under the responsibility of a professional subject to the obligation of professional secrecy (arts 9.2(h) and 9.3 of the GDPR). Accordingly, health app providers cannot benefit from this provision and will have to rely on any of the other exceptions in art. 9 (e.g. freely given, specific and informed consent (art. 9.2(a)), where processing is necessary for reasons of public interest in the area of public health (art. 9.2(i)) or where processing is necessary for scientific research purposes (art. 9.2(j))).

As the same authors continue:

> The GDPR maintains a purpose limitation principle, meaning that personal data that is collected for a certain purpose cannot be used for a new and incompatible purpose (art. 5.1(b) of the GDPR). It is thus important to establish all purposes for which the personal data will be used at the time of collection. This is particularly relevant in the context of clinical trials. All too often, personal data collected in the course of a clinical trial (first use) may become of interest for the use in other research, independent of this clinical trial (further use). The purpose limitation principle prohibits further processing of personal data incompatible with the initial purpose; however, further processing in accordance with art. 89(1) of the GDPR for scientific research purposes shall not be considered incompatible with the initial purpose. Nonetheless, if the legal basis for the further processing of personal data (secondary use) is consent under art. 6.1(a) of the GDPR, this may pose certain problems. Consent must be freely given, specific, informed and unambiguous. However, often at the beginning of the clinical trial (first use) when consent of the data subject is sought, it is not yet entirely clear for which further research purposes the personal data may also be used (further use). Fortunately, recital 33 of the GDPR allows for some flexibility in this regard and notes that data subjects should be permitted to give their consent for the further use of their personal data for scientific research on a more general level. Ensuring that data subjects give their consent at the time of collection for all purposes for which one intends to use the personal data is good practice and avoids the situation where one would have to go back to the data subject to ask for consent for additional purposes. (Van Obberghen et al. 2024).

The Belgian model is therefore similar to the French model. If public interest is the legal basis, there would be no need for consent from the data subjects, but their prior information would still be required by the GDPR. This applies to public and private entities, who can also rely on public interest as a legal basis (e.g. proving their research fulfils the public interest). Requests still need to be reviewed by the corresponding ethical review board, which will assess how the request for health data has a public interest and decide if the arguments fulfil that public interest.

In parallel to this legislative framework, an important development in Belgium is the new Health Data Agency (HDA), launched in 2023. The HSA mission is to ‘facilitate access to quality health (care) data and data related to health (care) in a simplified and more uniform, reliable, transparent and secure manner, through the development of a framework in which the re-use (secondary use) of quality health (care) data and data related to health (care) is optimally facilitated’ (HDA 2023). The goal of the HDA is to establish a transparent and efficient governance model for the reuse of health data (Article 5, 4°) which is operationalised in three ways: (i) serving as the primary point of contact for all requests to reuse health data (Article 5, 1°); (ii) facilitating access to health data by handling and processing requests (Article 5, 2°); and (iii) documenting and streamlining the process of requesting health data reuse. The latter includes providing guidance on making health data standardised and FAIR (Findable, Accessible, Interoperable, and Reusable) (Article 5, 6°) and creating a metadata catalogue to provide information about available health datasets (Article 5, 14°).

As Saelaert et al. (2023: 3) note, some of the tasks mentioned above (e.g. providing a national data catalogue, expanding the availability of health data, promoting common data standards in health, or harmonising data request procedures) align with the responsibilities that the EHDS assigns to the HDABs mentioned above. In contrast, there are two different aspects where the role of the Belgian HDA differs from HDABs. First, the HDA is not a Trusted Third Party (TTP) as defined in the EHDS, and therefore it is not involved in processing electronic health data (including from other data holders) for purposes of collecting, preparing, linking or disclosing these data for secondary use (idem). It is not holding the encryption keys of pseudonymised data either, which is also the responsibility of a TTP. This is a design feature intended not to overlap with already authorised TTPs in Belgium such as the eHealth platform, the Crossroads Bank for Social Security and Statbel (idem). Second, as opposed to HDABs that will be responsible for deciding on data access applications, as well as for issuing data permits to access electronic health data for secondary use, the HDA assigns the role of issuing data permits to other federal or regional institution(s) (idem).

One of these authorised TTPs is the eHealth platform launched in 2008, which includes organisations such as the Brussels Health Network, the Federal Agency for Medicines and Health Products, Medex, the Data protection authority, Agentschap Zorg & Gezondheid, the Federal Center of Expertise for Health Care (KCE), Walloon Health Network (RSW), the Agency for a Quality Life, Zorgplatform collaborator, Vlaams Ziekenhuisnetwerk KU Leuven, the National Intermutualist College (CIN), Mutual funds, The League of Users of Health Services (LUSS), or ZNA – care portal, among others (TEHDAS 2022a). As the TEHDAS report indicates, any of those organisations ‘can act as controller, processor or neither depending on the specifics of the data processing’. The caveat is the absence of “a ‘fixed’ definition of who can act as data processor or controller because this depends on specific applications” (TEHDAS 2022a: 32). This has legal consequences in terms of who will be the parties in collaboration agreements. Access health data in Belgium is granted through collaboration agreements between the data holder and the organisation requesting access. In this context, a data collaborative based in Belgium will need to consider its interaction with the HDA or HDAB as a Trusted Third Party

This is an evolving area where the Belgian eHealth Action Plan 2022–2024 aims to make a significant step forward with a flagship initiative, the Belgian Integrated Health Record (BIHR). Starting in 2025, BIHR will allow individuals and healthcare professionals to access personal health information and enable a complete view of a patient's health history. It also aims at streamlining reuse of data for the public interest, while facilitating opportunities to develop new products and services based on the exploitation of health data (INAMI 2024).

To exemplify the situation In Belgium with a national case study, let us consider the “Common position establishing a framework for secondary use of real-world data (routinely) collected in hospitals”, the framework developed by the seven Belgian university hospitals (Belgian University Hospitals 2024).

The Belgian university hospitals common position: a P + 4P analysis

The P + 4P analysis, adapted from Marcucci et al. (2023), is a straightforward method to establish the focus and scope of the case study and ensure that the analysis is directed towards the specific issue or phenomenon we want to cover, even if broadly. Thus, the problem definition guides the subsequent 4P framework, targeting the selection of relevant information. For the Belgian University Hospitals' 'Common position establishing a framework for secondary use of real-world data (routinely) collected in hospitals' we have identified the key Ps as follows:

Problem: The increasing volume and diversity of requests for secondary use of real world health data (RWD) collected in university hospitals presents a challenge in balancing the competing interests of research, innovation, and patient privacy and data protection. The growing complexity of data reuse exacerbates the challenge, particularly in the context of AI technologies and, more specifically, machine learning applications.

Purpose: the common position's purpose is to facilitate the secondary use of real world health data (RWD). It aims to make RWD accessible for both non-profit and commercial initiatives. The framework also seeks to maintain data protection and privacy standards when health-related data is processed for secondary purposes, allowing a better understanding of legal and ethical considerations.

Principles: although not specifically stated, the framework suggests that it is based on principles such as excellence in providing care based on evidence-based medicine; responsibility as data stewards of electronic patient health records, or transparency in designing the requirements and processes for effective sharing of RWD.

Processes: the processes suggested in the common position are based on the 'six conditions' for the secondary use of routinely collected RWD. These are:

1. Registration: All requests and decisions must be registered.
2. Privacy and Compatibility Assessment: Evaluation of lawfulness, relevance, and scientific underpinning of requests.
3. Right to Information: Transparency through general and personalised information.
4. Legal Basis and Exemption: Compliance with GDPR Articles 6 and 9.
5. Right to Opt-out: Patients can opt-out unless general interest prevails.
6. Security: Pseudonymisation as standard practice.

Practices: embedded within the 'six conditions' above, the common position suggests good practices for healthcare professionals when making assessments of RWD requests. As a synthesis of these (and extracted from) the framework:

1. Registration
    - Maintain a comprehensive register of all requests for secondary use and subsequent decisions
    - Use this register as a reference database for fulfilling data subject requests
2. Privacy and Compatibility Assessment:

  - Conduct thorough assessments of lawfulness, relevance, and scientific underpinning of requests
  - Involve a multidisciplinary body with necessary expertise to evaluate medical, scientific, legal, ethical, and technical aspects
  - Include patient representatives in the evaluation process
  - Perform Data Protection Impact Assessments (DPIA) when required

3. Transparency and Right to Information:

  - Provide layered, easily accessible information about secondary use of data
  - Offer general information through privacy policies and brochures
  - Give patients personalised information about specific projects using their data
  - Consider digital solutions for direct patient access to information
  - Adapt information provision for minors

4. Legal Basis and Exemptions:

  - Assess and specify the legal basis for processing under Art. 6 GDPR
  - Identify appropriate exemptions for processing health data under Art. 9 GDPR
  - Conduct compatibility tests for secondary processing
  - Use informed consent judiciously, only when other legal bases are not applicable

5. Right to Opt-out:

  - Provide patients with the ability to opt-out of secondary processing

  - Implement technical and organisational measures to detect and respect opt-outs
  - Clearly communicate the opt-out option to patients
  - Allow patients to review and modify their opt-out decision at any time
  - Perform a balancing exercise between individual and general interests when evaluating opt-outs

6. Security Measures:

  - Implement pseudonymisation as a standard practice
  - Apply additional security measures for secondary processing of RWD
  - Use automated pseudonymisation tools or federated learning models when possible
  - Establish transparent internal procedures for access rights and conditions
  - Require data transfer agreements with external parties, even between independent controllers

By establishing a clear purpose, guidelines and practices for secondary data use, this common position also addresses the need for patient empowerment and ensures compliance with GDPR regulations. Yet, as the authors also note, ‘the necessary contractual agreements such as data processing or transfer agreements have to complement the framework and the accompanying internal hospital-specific procedures. Such agreements should, for example, foresee in commitments [sic] to purpose limitation, the return or destruction of data, the sharing of results...’ (idem).

In this context, we argue that enabling participatory processes towards achieving social and impact licences would facilitate the purpose of facilitating the reuse of RWD preserving the balance between advancing research and innovation in a patient-centred perspective.

04

PUBLIC PERCEPTIONS AND MEANINGFUL ENGAGEMENT: TOWARDS BUILDING SOCIAL AND IMPACT LICENCES

4. Public perceptions and meaningful engagement: towards building social and impact licences

A 2022 report of the joint action Towards the European Health Data Space (TEHDAS 2022b) summarises public perceptions of how health data could be used in the future. Key highlights from respondents were, first, that secondary use of health data should be in the public interest (e.g. research and policymaking to improve public health) with more hesitancy about their data being used for commercial purposes. Second, that there is a need for transparency, ethical standards, and a solid legal framework for people to trust data sharing authorities and processes, including being consulted in those processes (idem). Likewise, as reported in Saelaert et al. (2023), a survey from the King Baudouin Foundation in Belgium (KBF 2022) revealed that ‘three out of four persons would share data for scientific research, whereas almost 50% would share with governmental institutions to support health related policy making. A bit more than half of the people would share data with pharmaceutical companies’ which shows that ‘although many people might be willing to share their health data for secondary use, this goodwill does not equal a blank cheque and people judge different data users in different ways’ (idem, 4).

Other public perceptions research has found that people have nuanced views on how data and data-driven AI systems should be used and that they want to have a meaningful say in them (Colom et al., 2023). More broadly, the implications of using health data both in terms of common good and in terms of risks such as privacy, bias and discrimination require meaningful engagement processes. As pointed out elsewhere, it is becoming increasingly urgent to link democratic institutions and practices with data, technology and its users (e.g. Casanovas et al., 2017; Colom & Poblet, 2025).

In line with the need for collective engagement and meaningful consent for public interest data, below we expand the concepts of social license and impact license respectively.

4.1 The social licence approach

There are multiple ways to define ‘social licence’. To Milne et al., ‘a social licence describes how the expectations of society regarding some activities may go beyond compliance with the requirements of formal regulation’ (Milne et al, 2021: 323).

The reuse of data can have great benefits for people and society when done responsibly, but, when not, it can cause serious harms, such as amplifying bias, inequalities or violating human rights. Moreover, reusing data entails the need for consent and permission to ensure these data can be used for purposes beyond those originally intended. Binary and individual approaches to consent are essential but not enough to ensure there is collective permission for using data for a range of uses. A social license then refers to securing and maintaining ‘public support for uses (and reuses) of people’s data’ (Aitken et al., 2020).

By a social license we mean not a legal, technical, procedural label, like a certification, but a set of governance processes that ensure relevant and diverse publics and stakeholders are included in decision-making processes for agreeing on what type of data reuse is appropriate and legitimate in a given context. Unlike legal licenses, which are static and fixed for a defined use and time period, a social license is a tacit permission granted by communities or stakeholders, which evolves as data uses and their ecosystems also evolve (Aitken et al., 2020; Moffat et al., 2016).

A social license can be obtained, for example, through embedding relevant publics and stakeholders in the decision-making organs of an organisation –such as including a representative citizens' panel as an organ alongside or inside an organisation's board–; or through commissioning participation processes that then inform an organisation's code of practice, rules or decisions on specific reuses of data that can add collective legitimacy in addition to individual consent.

The risk of using a concept that can be interpreted in multiple ways is that it can be used opportunistically through, for example, tokenistic consultation processes rather than meaningful participation that grants permission. An accepted definition referred to in the literature comes from Moffat et al. which defines a social license as 'the ongoing acceptance or approval of an operation by those local community stakeholders who are affected by it and those stakeholders who can affect its profitability' (Moffat et al 2016: 480).

Trust is an underlying component of a social license (Verhulst and Saxena, 2022) and this means that the process by which an organisation or data collaborative operates needs to be trustworthy, for which accountability, transparency and explainability are key. When a social license has been meaningfully granted, data holders, data subjects and those likely to be impacted by the reuse of data will 'trust that their data will be used as they have agreed, and accept that enough value will be created' (Data Futures Partnership, 2017 in Aitken et al., 2020).

4.2 The role of impact licensing

Impact licensing is the practice of establishing contracts and agreements having well-defined objectives for the use of data. In the context of data collaboratives, impact licensing can be defined as a strategic tool for maximising the societal valorisation potential of data. It entails the voluntary granting of a time-limited permission by the data collaborative platform to use data intentionally, measurably, and in an enhanced manner to a predefined market for a specific societal purpose. This purpose may involve addressing personalised citizen needs, unmet needs specific for certain citizen groups and unmet societal needs at the population level. Or it can be geared towards a more open ended research process. More specifically it encompasses the following qualities:

- Intentionality: The pursuit of societal impact must be deliberate and not restricted to a mere potential. This criterion contrasts with certain "washing" marketing practices that, e.g., label patents according to ESG criteria, but that only identify the capability or potential of the protected inventions, not their real achievements. The goal of impact licensing must, to the contrary, be explicit and dedicated to delivering a clear and affirmative societal benefit, not a mere capability. This necessitates a pre-event declaration and an equitable, non-discriminatory licensing approach that actively seeks economic and societal objectives.

- Measurability: The impact should be quantitatively and qualitatively assessed, ideally by independent third parties, to enhance their trustworthiness and credibility.

- Additionality: Impact licensing activities must prioritise undercapitalised regions and evaluate whether the impact achieved would have materialised without the licensing effort. They do not undermine other valorisation efforts in the traditional and more profitable markets.

Impact licensing is based on the premise that data needs protection via contractual mechanisms to achieve societal development. Social impact requires investments that can only be attracted if the principles that underpin the way data or a technology is used can be protected. Impact licensing, therefore, aims to create value for the individual citizen, the organisations involved and society at large without losing sight of the economic aspect and the goal of generating a return on investment so that the process can be sustainable.

An impact licensing agreement contains at least five enforceable clauses:

- The impact domains and measurable key indicators where data will have material effect upon (ie. indicators determined on the basis of the UN Sustainable Development Goals framework). The definition of the impact scope defines the main societal market boundaries of the impact licensing agreement and makes it possible to specify methods and processes whose impact will be monitored and reported.
- The eligibility conditions (such as geographies, target groups, use settings, and/or tipping points) to bring developments to the defined societal markets.
- Exploitation clause: a clause that returns the impact license to for example the platform of the data collaborative in case the licensee is not or using (or is under exploiting) the IP in certain domains, applications and geographies
- The boundary conditions in data usage that have to be adhered to in the societal value generation (individual and collective consent, security, quality monitoring, reputational risks,..)
- The economic safeguards: the restrictions to protect the potential of these technologies for economic growth ex-ante and ex-post the duration of the licensing, the use of technologies beyond the societal market, or time boundaries, among other safeguards.

We argue that impact licensing can help data collaboratives position themselves strategically to shepherd the reuse of data sustainably towards social value-creation and to meet societal needs.

05

AN OVERALL FRAMEWORK FOR SITUATING EMERGING DATA COLLABORATIVES IN THE PLAYING FIELD

5. An overall framework for situating emerging data collaboratives in the playing field

As stated earlier, data collaboratives are partnerships with different stakeholders and operate in complex ecosystems. To make sense of this complexity, we highlight three core relationships of a data collaborative or a data reuse ecosystem: (i) with the individual, (ii) with the organisational ecosystem, and (iii) with society as a whole. We propose three key mechanisms to structure these three types of relationships: informed consent, impact licensing, and social licensing respectively.

Informed consent is the process by which an individual voluntarily agrees to participate in research (and this including consenting on the collection, use, or disclosure of their data), after being 'provided with all information likely to be relevant to their decision whether or not to participate, and where such information has in fact been understood' (Boulton and Parker, 2007). The informed consent shapes the relation with citizens providing them with rights to decide on the use and sharing of their data. This decision will depend on the scope of use and needs to be clarified upfront. Informed consent is a mechanism to share and control data coming from the individual data vaults. It is the source of data influx for a data reuse ecosystem.

As described in the previous section, **impact licensing** is a strategic tool for maximising the societal valorisation potential of technologies and data without losing sight of the economic aspect and the goal of generating a return on investment. It entails the voluntary granting of a time-limited permission by a technology holder to introduce a technology and/or data intentionally, measurably, and in an enhanced manner to a predefined market for a specific societal purpose. Impact licenses therefore can help data collaboratives to safeguard the use, commercial and other, of shared data towards well-defined objectives.

As also described earlier, **social licensing** refers to the collective informed and deliberated acceptance or approval that an organisation or partnership receives from the community and stakeholders to operate in a specific area. It is a process of participation in which a tacit license to operate is negotiated at the collective level. As such this process also arbitrates on which business models are aligned with the foundations of a data collaborative. It goes beyond legal permits and regulations, focusing instead on the social and ethical dimensions of projects and activities and can inform codes of practice or the frameworks within which formal governance organs, like a board,make decisions. Table 1 provides more information on the content of the mechanisms.

The combination of informed consent, impact licensing and social licensing can allow a data collaborative to put its foundational caring technology principles into practice. By using these mechanisms in a smart adaptive way, the meta-principles of caring technology are embodied by everyone involved: promote humane technology and citizen-centred data management, support societal anchoring, stimulate participatory governance, monitor quality and foster systemic coherence. Moreover, the shaping of the combination of informed consent, social licensing and impact licensing allows for paving the way towards new models of value creation in which social and economic value go hand in hand while respecting the intrinsic value of agency and responsibility of individual citizens, collectives and society.

Table 1. Summary of three mechanisms to shape the playing field of data collaboratives.

	Informed consent	**Impact licensing**	**Social licensing**
Relationship	Members of the public	Organisations	Society
Purpose	Enable citizens to retain their individual agency in deciding over the use of their data	Safeguarding that the (commercial) use of data is directed towards well-defined objectives	Arbitrating on the principles and values that underpin the data collaborative's business model
Attributes	Disclosure Understanding Voluntariness Capacity Revocability	Definition of impact Scope Eligibility conditions Exploitation clause Technological usage Parameters Economic safeguards	Inclusive and shared decision-making Collective learning ethos Understanding the whole Community engagement Participatory process Sensemaking Co-creation of vision
Regulatory and legal context	GDPR EHDS	Enforceability mechanisms Competition law	Participatory governance, value-systems, principles and codes of practice

Source: Own creation

06

KEY DIMENSIONS IN EXPLORING A BUSINESS MODEL

6. Key dimensions in exploring a business model

This section summarises the key components of a business model for an ecosystem for data reuse. By business model we refer to the set of principles, processes and practices that ensure the service can deliver value to its users while ensuring its own financial sustainability. Different sources in our literature review outline different lists, categories and frameworks for definite business models. At the same time, there are in essence very consistent key components of a business model across the literature reviewed. We categorise them into the following three overall categories:

1. Value proposition: The value refers to the purpose and set of principles, processes and practices that underpins a service so that it offers value or meets the needs of its users. Others have defined it as 'the customer's need(s) met by the organisation and what the organisation does in order to satisfy such needs.' (Micheli et al., 2023). As stated by Micheli et al, the value proposition is an essential element but is not enough to ensure the financial sustainability of the service. The value proposition needs to be 'clearly articulated' because the decisions on the business model will follow from this vision (OECD, 2017). Not being clear on the value proposition might result in other components dictating what the service can offer (such as the type of technology used influencing what value is possible to add to the data and for whom, or the need for financial sustainability enforcing specific value and beneficiaries over others). As argued by Lauf et al., 2023: 'Naturally, a data trustee can serve multiple purposes, but only one purpose constitutes its main purpose. The purposes of a data trustee are linked to the Customer Value provided' (p.8). Specifically, it is worth considering the following components of a value proposition:

1.1. Purpose of the service. This relates to the problem that the service or ecosystem for data resume is trying to solve. For example, it could be enhancing privacy, making intermediation trustworthy, or ensuring that the social value of data is shared for social impact, such as advancing health and healthcare.

1.2. Value added to the data and how it differs for different potential data holders or subjects (supply side). This refers to the different types of value the service can offer (supply) for different data holders or subjects. For example, it might be adding privacy-enhancing value to some users but also adding knowledge value to other users through the provision of analysis and insights. It is essential to be clear on what are the boundaries for data reuse and how the ecosystem for data reuse will be governed so that the purpose of the service and the value added to the data for different data holders are in harmony.

1.3. Value given to the data (or incentives) for different potential users (demand-side). This refers to the principles and expectations of value from the potential users of the service. Users might be aligned in prioritising the value of reusing data to advance health research but other potential users might value it as a commercial asset and source of income.

For example, in a 'gatekeeper model', the data intermediary 'serves as a trusted third party that negotiates terms of access to previously closed data by users selected through a call for proposals.'

The value, in this case, comes from the legitimacy of the process.’ (Susha et al., 2020). In an ‘information-as-a-service’ model, the intermediary provides data visualisations to targeted users to improve their decision-making. In this case, the value ‘comes from ease-of-use and quality of decision support’ (Susha et al., 2020).

It is important to find a balance between supply and demand. Defining the supply and services offered based solely on the demand of a range of actors can compromise the purpose and principles for which the service was initiated in the first place. This is why the governance of the service is essential, as the other components listed below. For example, in the case of the health data cooperative MIDATA (see Case Studies section), the service aims to foster digital self-determination and this is also a key incentive for members of the general public to join the cooperative as members (Gille and Vayena, 2021). In addition, for a service that offers data intermediation it is important, as Micheli and others point out, that the reduction of transaction costs is an added value: ‘if it was easy and costless for a data holder and a data user to meet, agree on the terms of data sharing, and carry on with the exchange (for example, because they know and trust each other and they both use the same standardised technical protocols and data sharing licence), they would not need a third party for data sharing to happen.’ (Micheli et al., 2023:35). Micheli et al refer to the range of transaction costs, such as search costs, bargaining costs, monitoring and enforcement costs or information costs. This is also highlighted by others as an important component of the competitive advantage of a data collaborative: ‘Sources of competitive advantage may vary depending on the type of barriers encountered during the processes of data acquisition and analysis’ (Ferro and Pizzamiglio, 2023).

1.4. Level of agency over data. This entails defining how much agency do data holders or subjects have over the data before it is reused. This will affect how decisions are made and therefore the type of services and revenue that can be included in the business model. It will also determine -alongside the legal requirements- the need for processes for meaningful engagement of data sharers, such as the need for consent or processes to enable a social license (participation processes to determine collective consent) for data reuse.

1.5. Governance. This relates to the processes and principles that will govern the service. It includes how decisions will be made, who will be involved on what decisions and at what stage, or what systems of accountability there will be in place.

1.6. Level of data openness. The extent to which data is open will determine the services and type of revenue that can be generated to sustain the service. However, it is the other elements of the value proposition -such as the purpose of the service and the value given to the data by data sharers-, as well as elements of the ecosystem -like the type of data- that will determine this level of openness in the first place.

2. Ecosystem: Defining a business model also entails being clear on the infrastructure and stakeholders that will interact with this data. There are a range of components in the ecosystem:

- the type of data being stewarded and shared
- the type of stakeholders involved
- the technology itself and what it enables or limits

- the regulatory context
- the skills and resources available
- the funding and revenue landscape.

For example, different types of data entail different risks, need to follow different processes based on regulatory requirements, demand higher protection and privacy from those holding it, or can add different value to different stakeholders. Similarly an ecosystem that enables easy access to seed funding can help the service start and buy it time for the business model to develop, adapt, and help the service be sustainable.

Others, like Susha et al. (2020) have referred to this category of components as two-folded: the 'value architecture' (or the 'architecture for the technological and organizational infrastructure used in the provisioning of products and services') and the 'value network' (the 'collaboration and coordination with other organizations') (2020: 307).

3. Financial proposition: The financial proposition is essential to a business model as it sets out the sources of revenue and income that can sustain the service. It needs to be aligned with the components of the ecosystem and be part of the value proposition rather than compromising it. Key components of the financial proposition are:

- the pricing mechanism
- the use and redistribution of revenue

As described by Susha et al. (2020) these components refer to the 'value finance', operationalised through the '[p]ricing and revenue breakdown associated with sustaining and improving the creation of value' (p. 307).

Table 2. A summary of the components of a business model

Business model framework dimensions	Business model components	Data governance framework dimensions (P+ 4 P's)
VALUE PROPOSITION	Purpose/aim of the platform/service (ie. data protection, enhanced privacy, accessibility, information/analysis as added value, social impact,...)	PROBLEM (that the service addresses)
VALUE PROPOSITION	Value added to data - this relates to incentive for the different actors (supply focused, demand focused, privacy focused, social impact focused, revenue focused)	PURPOSE
VALUE PROPOSITION	Value given to the data by different potential users (economic, social, environmental...)	PRINCIPLES

VALUE PROPOSITION	Level of agency over data, including processes like consent, social licensing	PRINCIPLES AND PROCESSES
VALUE PROPOSITION	Governance	PROCESSES AND PRINCIPLES
VALUE PROPOSITION	Extent of data openness	PRACTICES
ECOSYSTEM - DATA	Type of data	PRACTICES
ECOSYSTEM - POLICY AND REGULATION	Regulatory incentives and rules	PRACTICES
ECOSYSTEM - ACTORS	Type of exchange and data usage (from which supply actors to which demand actors - evidence on how data is used)	PRACTICES
ECOSYSTEM - TECHNOLOGY	Type of platform and technology affordances (Value added to the data through platform/service	PRACTICES
ECOSYSTEM - FINANCIAL	Sources of funding	PRACTICES
ECOSYSTEM - SKILLS/RESOURCES	Skills and resources	PRACTICES
FINANCIAL PROPOSITION	Pricing mechanism	PRACTICES
FINANCIAL PROPOSITION	Use and redistribution of revenue	PRACTICES

6.1 Main challenges

While these are components consistently mentioned in the literature, aligning them to create business models for the sustainability of ecosystems for data reuse is challenging. A recent report by the European Commission’s Joint Research Centre (JRC) Science for Policy concluded that it ‘is not clear which business models can guarantee the economic sustainability of the different types of data intermediaries.’ (Micheli et al., 2023) and others have also emphasised that this is a sector still ‘in its infancy’ and ‘and requires further research and implementation efforts to fully express its potential for public and private value creation.’ (Ferro and Pizzamiglio, 2023).

Some of the challenges mentioned in the reviewed literature are related to:

- **Limited funding**

Data collaboratives in the social and health spaces often find it difficult to find and/or sustain funding (Lauer et al., 2024). This not only limits its sustainability, and therefore impact, but also can affect its value proposition, as offering certain services or building partnerships with the private sector might be more difficult without enough funding to cover the costs, salaries and time that this entails. Some have pointed out that public funding can limit agency and independence (Lauer et al., 2024) or come with challenges in relation to adapting to ‘procurement rules and the risk of preferential treatment

07

CASE STUDIES

or picking winners’ (Krewer and Warso, 2024). While cooperatives can be sustainable if they include large numbers of members, they also depend on initial funding that can be sustained for long enough to transition to other forms of income (Hafen, 2019). At the same time, obtaining subsidies and public funding also requires time and people to ensure enough potential funders know, understand and are willing to support the data collaborative (Lauer et al., 2024).

- **Failing to see data as relational and collective**

Data collaboratives or data reuse ecosystems exist because of the need to address a complex ecosystem and address power imbalances. Yet, the literature points to a shortcoming for some initiatives in failing to fully acknowledge and account for the relationality of data. For example, Wong et al., 2024 argue that ‘data are always co-created by a source and a collector’ and that ‘data are inherently relational and collective’ and that this has implications for its governance as well as the value that can be generated. The authors argue that a collective governance approach can both provide more complete information and reduce transaction costs as well as rebalance interests and how power is shared (Wong et al., 2024).

- **Decision-making in different types of participatory governance**

As indicated above, collective governance is a key aspect in many data-sharing ecosystems but the literature also points out, especially in relation to cooperatives, that shared decision-making can be complex (Hafen, 2019). Building on a governance structured with clear decision-making processes that balance the need for individual and collective agency as well as transparent and accountable governance is therefore key and done in conjunction with the business model.

Other challenges mentioned in the literature, although related to the ones just outlined, refer to: limited skills and resources, unclear identity, no value generation, and difficulty in data access and transferability. These reinforce the need for a clear value proposition that is aligned with the ecosystem and the financial proposition.

7. Case studies

In this section we present some case studies from entities that act as intermediaries or data banks of health data and which present different business and governance models. Specifically, we cover data cooperatives, health records banks, and data trusts. There are, however, other models that can be followed to establish platforms for data reuse. The table below, sourced from Micheli et al. (2023) maps them out, indicating their specific characteristics:

Figure 1. An overview of different types of data intermediaries adapted from Micheli et al. 2023

Type	Main goals	Value proposition/business models considerations	Approach
PIMS	Provide tools to individuals to take control over their personal data	Finding alternative revenue streams to user payments (ie. data-driven services).	Individual
Data cooperatives	Establish a bottom-up governance structure and produce benefits for the members of the community	Providing incentives for individuals to devote time to the different tasks they carry out	Collective
Data trusts	Establish responsible data management through independent decision processes in the interest of data subject holders	Focusing on increasing uptake in order to generate sustainable revenue streams	Individual
Data unions	Establish collective bargaining on rights to personal data generated through platforms	Providing incentives for individuals to devote time to the different tasks they carry out	Collective
Data marketplaces	Match data supply and demand	Reducing search costs and facilitating data exchange initiation between supply and demand	Business oriented
Data sharing pools	Leverage data synergies between stakeholders with complementary datasets	Improving data products and services, creating new business opportunities, and tapping unexploited data value	Business oriented

Source: Micheli M, Farrell E, Carballa-Schmichowski B, et al. (2023)

However, not all of these models are suitable for the management of health data. For example, it has been argued that ‘Data marketplaces’ focus solely on the monetisation of personal data and we could not find evidence that individuals selling their data (such as data from their health devices like wearables) have meaningful input and information to decide whether data should be used or can serve the public good (Kariotis et al., 2020).

7.1 Data cooperatives

Data cooperatives are organisational structures whose members pay a fee to join, can store their data on its digital infrastructure and become part-owners with a say in the governance (Kariotis et al 2020, Micheli et al. 2023). As Micheli et al. note, the model is largely inspired by the broad cooperative movement over the last two centuries both in the UK and the European continent (Micheli et al. 2023). We retrieve below the most salient examples of data cooperatives as found in the literature review.

MIDATA

The MIDATA cooperative was founded in 2015 by researchers at the Swiss Federal Institute of Technology Zurich and Bern University of Applied Sciences. The focus is on health data use in the context of medical research projects, with special consideration to ethical and legal aspects such as data protection, data ownership, and informed consent (Lauer et al, 2024). The cooperative 'operates a data platform, acts as a trustee for data collection and guarantees the sovereignty of citizens over the use of their data' (MIDATA 2024). MIDATA promotes an open innovation ecosystem by allowing startups, technology providers, and research institutions integrate their apps with the platform. Data account holders on MIDATA (individual citizens who upload their data into the platform) can participate in medical research and clinical studies by granting controlled access to their personal data (idem). They may also become members of the cooperative, giving them a say in its governance. According to MIDATA's Articles of Association, cooperative members have to be Switzerland residents, have to pay a one-time fee (CFH 40), and do not get paid dividends as net profits are reinvested into the services offered on the platform (MIDATA 2019).

MIDATA also offers 'data stewardship' for research projects and partnerships by supporting 'all processes regarding planning and implementation of research projects or launching of health and lifestyle apps' (MIDATA 2024). According to the website, MIDATA selects projects that meet the 'legal, ethical and technological requirements for a fair and transparent data use', assure the 'digital self-determination of the citizens' must be assured, and are designed 'to bring a real benefit for the participants and society' (idem). Specific projects currently hosted on the platform are Sero (suicide prevention), OpenProms (Patient Reported Outcome Measures), or Approches (support for family caregivers). The platform software is published as "Open MIDATA Server", under GNU General Public License v3.0.

Salus Coop

Founded in 2017, Salus Coop is a Catalan-based, citizen-led data cooperative with the mission of connecting data donors with health researchers (Salus Coop 2024). The cooperative has developed a mobile application, the Salus.Coop App, for donors to register and upload their data. In the backend, registrations are recorded as transactions on the blockchain and donors' data are pseudonymised and stored both in IPFS (InterPlanetary File System) and AWS (the Amazon cloud). The latter allows researchers to access the data when licensing conditions are met. In this regard, Salus Coop conducted a public consultation among its members on their preferences for data sharing and the outcome were two different licences: the Common Good License (for non-profit research projects linked to health) and the Personal Return License (for research projects that allow economic return, maintenance of the cooperative, or the start of new projects by the cooperative itself). In both cases, licensees must adhere to the following conditions: (1) use of data for health research; (2) use by non-profit entities, (3) open sharing of the results; (4) anonymization at the highest possible level, and (5) use of the data until the consent of the donor is withdrawn (Creus 2024: 33). The Common Good License has been integrated into the Ethical Commission Review of Barcelona’s Hospital Clínic (idem: 33).

Even if information on the business model is not available on the website, founding member Javier Creus has noted that ‘Salus.coop’s business model was supposed to generate income from fees charged to research projects in exchange of recruiting data donors that would use the Salus app to share their data, or by sharing its platform Software as a Service with other communities’ (idem: 33). That model may have been hampered by misalignments between Salus Coop agenda and ‘the actual incentives for the use of data’ (idem: 33).

The Holland Health Data Cooperative (HHDC)

Launched in 2017 with support from the Municipality of Rotterdam, My Data Our Health Foundation, TNO, IJsselwijs Foundation, and Medical Delta, the Holland Health Data Cooperative (HHDC) focuses on giving citizens control over their health data while promoting collective benefits and reinvesting in health-focused initiatives. The funding model relies on the Roadmap Next Economy, a collaborative initiative of South Holland municipalities and provinces, but the HHDC website also lists various projects in partnership with other organisations (e.g. Okapion and Share2Care) receiving different funding sources (e.g. SIDNfonds and Diabetesfonds) and stakeholders. The cooperative has an Ethical Data Use Assessment Committee to evaluate both projects and data use requests.

HHDC helps its members decide who can use their health data and for what purpose. Members can set a filter to only receive requests that match their preferences, based on criteria such as who is asking, why they want the data, what kind of data they need, whether the member needs to do anything, what the member might get in return, and how long they would have to participate. This makes it easier to find requests that match the member's interests and/or filter out unwanted requests. The cooperative’s platform keeps track of granted permissions and makes these visible on the member's personal page. A member's personal web page shows the outstanding data requests with the criteria used and shows any outstanding requests. The cooperative also considers a process of delegating consent (and being able to revoke it) complying with the legal framework, but the implementation of this must be further elaborated (HHDC 2021).

7.2 Health records banks

Health record banks can be broadly defined as centralised repositories to store and manage health-related data for individuals. The Health Record Banking Alliance (HRBA), founded in 2006 in the UK offers key defining elements of a health record bank, compared to those offered by Personal Health Record Systems (PHRs), where consumers enter their own medical records. Google Health and Microsoft HealthVault (now discontinued) are early examples of PHRs.

An example of a health record bank is the French Health Data Hub (HDH) that was established by law in 2019 to replace the previous *Institut National de Donnés de Santé* (INDS). Unlike its predecessor, the HDH offers a robust technological infrastructure for processing and managing health data. While it doesn't grant permission for secondary data use, it serves as a centralised platform for researchers to securely access and analyse non-identifiable health data.

Table 3. Differences between Health Record Banks (HRBs) and Personal Health Record (PHR)

Feature	Health Record Bank (HRB)	Personal Health Record (PHR)
Data Source	Direct from healthcare providers	Self-reported by patients
Data Collection	Automated	Manual
Data Completeness	Comprehensive	Limited
Data Reliability	High	Lower (due to self-reporting)
Data Integration	Integrates various data types	Limited integration
Mobile App Integration	Strong	Limited
Consumer Empowerment	High	Lower

Source: Own elaboration from HBRA source

The other examples that we have found in the literature, nevertheless, are platforms combining aspects from HRBs and PHRs. For example, data is self-reported by the public but both data and mobile integration are high, and so is the level of engagement and empowerment of their users. The three examples below can be considered as hybrids between the two.

UK Biobank

The UK Biobank defines itself as ‘a large-scale biomedical database and research resource, containing in-depth, de-identified genetic and health information from half a million UK participants’ (UK Biobank 2024). It started in 2006 as a long-term longitudinal research project, funded by the Wellcome Trust, the Medical Research Council, the UK Department of Health, the Scottish Government, and the Northwest Regional Development Agency. UK Biobank collects de-identified genetic and health data from half a million UK participants, and enables access to these data to approved researchers from academia, government, philanthropy and commercial organisations. In 2021, a large Covid-19 study recruited more than 200,000 of Biobank participants and their families to contribute antibody test results, to understand the persistence of antibodies over time (idem). At UK Biobank, The Access Committee of the organisation is responsible for making decisions about granting access to the data, while the Ethics Advisory Board provides advice on current and/or future ethical issues regarding the use and stewardship of these data.

In November 2023, nevertheless, an investigation found that the UK Biobank had opened its datasets to insurance companies, despite having pledged not to do so (Das 2023). The UK Biobank alleged that insurance research was done by reusing anonymised data in the ‘public interest’ and with consent from participants to reuse their data for commercial purposes. However, recent research has contended these claims and proposed a ‘significantly expanded understanding of the publics who must be involved in biobank governance and data-sharing decisions involving insurers' (Metcalf and Sadowski 2024).

OpenHumans

OpenHumans was launched in 2015 as a project of the Open Humans Foundation with grants from the Robert Wood Johnson Foundation, Knight Foundation, and Shuttleworth Foundation. At the time of writing, the community has more than 10,000 members and 40 tools and activities are running on the site (Open Humans, n.d.). The open source project is 'dedicated to creating public genome, health, and trait data'. Open Humans primarily relies on its users to upload their own health data. This can be done through various methods: (i) users can manually upload data files from their computers or external devices; (ii) API integrations with third-party health data sources, allowing users to automatically sync their data, and (iii) participation in research projects, which may have specific data collection requirements and may provide tools or instructions for users to upload data.

Open Humans has a multi-tiered governance system. Individuals have full control over their own data, while the community oversees the platform. Research projects need approval before they can collect data from members, and members vote for the board that runs Open Humans. As members share data with projects and give permission for data reuse, thay are also facilitating the creation of data commons for specific data types and patient patient communities (Lauer et al, 2024).

Open Humans primarily operates on a non-profit business model (501(c)(3) nonprofit). It doesn't generate direct revenue through selling user data or services, but it has several revenue streams to support its operations, including grants and donations, partnerships with research institutions, universities, and corporations, and optional membership fees.

Patients Like Me

Patients Like Me was founded in 2005 with the mission to empower patients to connect, share their experiences, and contribute to medical research. At present, the platform has 850,000 registered members from a global community, with over 2,800 different medical conditions, sharing personal stories and information about their health, symptoms, and treatments (Empowering Patients Through Community, 2024).

Patients Like Me offers a variety of features to support its community. Users can engage in online discussions through community forums, track their symptoms over time using the symptom tracker, and search for relevant clinical trials. Additionally, the platform collaborates with researchers to conduct studies utilising patient-generated data collected through the platform. Patients can choose to participate in research studies that utilise the platform's data and they need to grant their consent before that.

Unforgettable

Unforgettable is the platform run by the company Unforgettable Research Services, based in Victoria (Australia). The purpose of the platform is to allow its users to upload their medical information, which can then be shared with research partners for research purposes. Health information from users covers a comprehensive range of data (from GPS location and accelerometer data from mobile devices to images, audio and video, etc.). Data from users may come through different sources: the unforgettable.me mobile app, IFTTT (e.g. email, Google Drive or Dropbox) or is entered directly into the website through surveys or embedded experimental software (Unforgettable Me, 2024). Users can also contribute with their data to research projects, on a voluntary basis and after providing consent. In some cases, users participating in a research project may be offered some payment for their contribution to the project.

Registration to the platform is free but added services require a monthly fee. The pricing scheme is not available on the Unforgettable website but the platform offers different subscription plans for their services (full access to the data and personalised analytics).

7.3 Data trusts

Data trusts have been broadly defined as ‘a repeatable framework of terms and mechanisms that is mandated for use (or subject to scrutiny, or certification) in particular contexts to provide oversight of data access’ (ODI 2018), as ‘a legal structure that provides independent stewardship of data’ (Hardinges et al., 2019), ‘an independent structure for stewardship of data that can enable flexible and inclusive data governance and respect multiple interests’ (Milne et al 2021: 324), or as ‘a fiduciary required to keep the interests of the users at the forefront when making decisions about the processing of their data and ensuring their privacy’ (Lauf et al 2023).

Despite the many definitions and analysis of the data trust framework in the literature, specific examples for the health domain that are operational are scarce. This could be due the legal figure of ‘trust’ being absent in non-common law jurisdictions, where other entities (such as cooperatives) may fulfil a similar role. Kariotis et al (2020) consider OpenHumans (examined in the section covering biobanks) as an example of data trust in the health space. In the UK, the Data Trusts Initiative has started two pilot projects in this area. Ones is the Born in Scotland Data Trust, which creates a framework for responsible management of health, administrative, and social data from pregnant women and their children. This project aims to address health inequalities in Scotland while empowering participants to have a say in how their data is used. The second pilot is GP Data Trust, exploring how to give patients who have opted out of NHS data sharing more control over their health information. By building a data trust, GP Data Trust aims to enable these patients to participate in vital research while ensuring their data is managed responsibly.

Blockchains could offer the technical infrastructure for data trusts to flourish in the health sector. Recently, some communities have launched Decentralised Autonomous Organisations (DAOs) with initiatives that, while not formally incorporated as trusts, exhibit similar features, notably in terms of data stewardship and flexible data governance. One example is CureDAO, which can be described as a decentralised data trust, as it empowers individuals to own and control their health data while enabling its use for research and innovation. CureDAO leverages blockchain technology to create a secure platform where users can share their data, participate in research, and receive personalised health insights. The healthcare DAO landscape is highly evolving (Behluli 2025), and new initiatives may emerge in the future. As blockchain technology and decentralised governance continue to mature, we can expect to see more innovative and impactful DAOs in the sector.

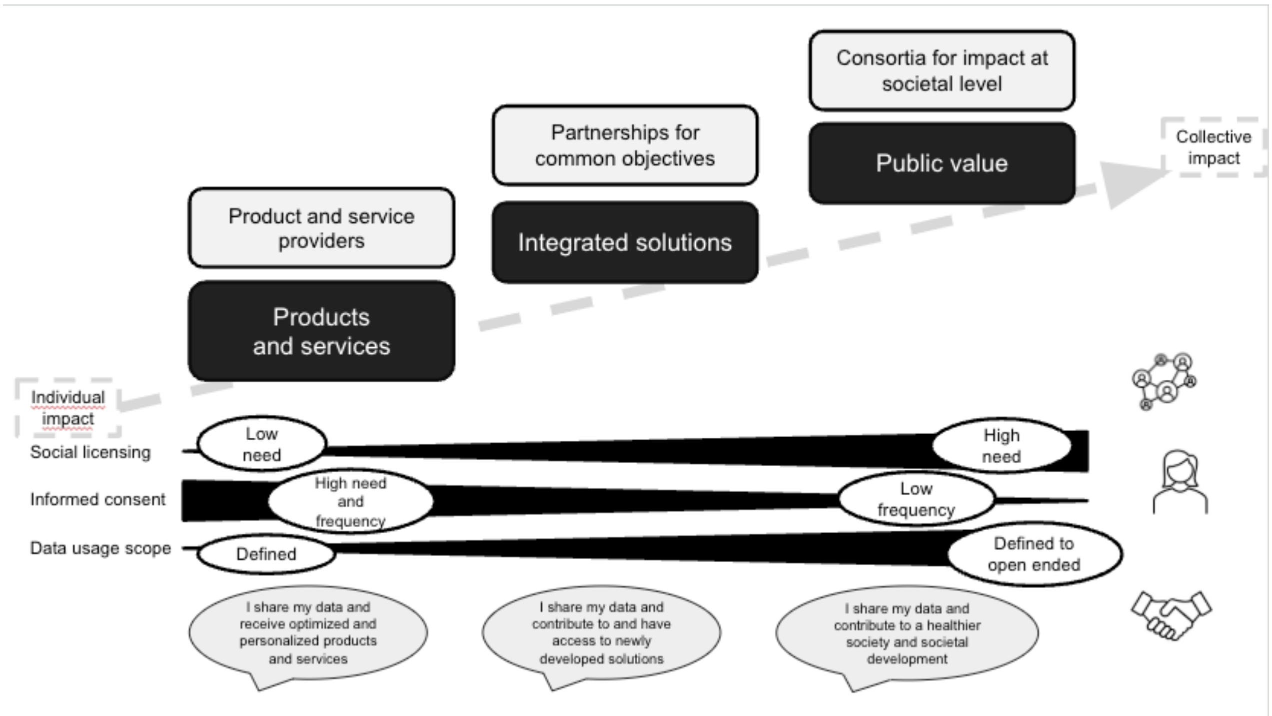
08

PATHWAYS FOR A BUSINESS MODEL FOR DATA COLLABORATIVES

8. Pathways for a business model for data collaboratives

As evidenced from the literature and case studies, data collaboratives or ecosystems for data reuse often depend on public funding. While we have already outlined the key components to consider when deciding on a business model, in this section we aim to structure and make sense of the vast variety of business models that exist beyond data ecosystems, from simple transactional ones to more complex ones operating in a systemic way. Figure 3 below presents an overview of three categories of business models–transactional, relational and systemic business models– and how they relate with informed consent, impact licensing, and social licensing. In each business model, the data collaborativeacts as a platform to facilitate value creation. Yet, its role has a distinct orientation depending on the mode adopted.

Figure 2. Three business models and their relationship to informed consent, impact licensing and social licensing.



Source: Own creation

Transactional business models

Under this category, products and service providers develop personalised products and services through the use of data shared by individual citizens. An example of this could be a personalised service to optimise the relationship between nutrition and intake of medication in Parkinson's disease to enhance quality of life. There is a linear relationship between providers and citizens that would be chaperoned by a data collaborative platform as a quality adviser. The role of the data collaborative would be limited to imposing minimal requirements to ensure quality and safety, for example by means of a quality label. The decision to use certain products and services is with the individual citizens. This implies that for each specific use of personal data, the individual citizen needs to provide consent. No requirements regarding objectives or purpose are imposed by the data collaborative platform. The relative importance of informed consent is therefore high. In principle, the need for deliberation and social licensing would be low if the transaction follows both regulation and informed consent and if no collective harms or consequences are associated with this individual-level transaction. Even then, however, it should be noted that data is collective and relational in nature and that its reuse 'can have consequences for groups and individuals from different communities that individual consent does not account for' (Colom and Poblet, 2025). Ultimately, because we enter the realm of personalised health models, the data usage scope has to be well-defined and the potential risks for or implications of this data being reused beyond the individual should need to be carefully considered.

Relational business models

Relational business models support the development of integrated solutions for groups of citizens. They require clear purpose and objectives, as well as collaboration between organisations. While a transactional approach focuses on a personalised service (e.g. the Parkinson's disease example above), an integrated solution would go beyond this to combine advice regarding nutrition, movement, and cognitive exercise, blending it with home caregiving. All personalised services are combined into an integrated solution. Relational models typically target specific end-user groups through establishing business relations with organisations and managers, rather than with end-users. In this model, the data collaborative platform would impose minimal requirements regarding quality and safety and, in addition, would act as a custodian and gatekeeper, ensuring that shared data are used for a purpose and objectives that have been defined upfront. The usage and scope of the data increase and the informed consent takes place as a more overarching level to include a wider perimeter of data usage. Collective deliberation and arbitration processes become relevant as the legitimacy to operate will be supported by a social license process, deciding which objectives and collaborations live up to the ethical standards rooted in the foundational principles of the given data collaborative or data reuse ecosystem.

Systemic business models

Systemic business models consider both value creation and outcome at a societal level. This broad and holistic view includes the role and input of all stakeholders involved in the creation of value, outcome and impact. These models capture and enhance the added value of impact, such as a better health outcome at the population level. The role of the data collaborative is a trusted third party and data aggregator. The informed consent takes place at meta-level. The data usage scope is very broad and also implies an important research dimension. Citizens share their data to contribute to a healthier society and societal development through research. In this case, the social license to operate becomes fundamental. It is an open model welcoming transactional and relational layers as well. The purpose and objectives need to be well-defined and impact is aimed for at a collective level.

The role of the data collaborative across these categories of business models

The three categories of business models are characterised by an increasing degree in complexity and collaboration. It is also important to note that these categories are not strictly separated, but that they rather relate as potential layers within each.

The role of the data reuse ecosystem boils down to shaping and facilitating the ecosystem with informed consent, impact licensing and meaningful collective decision-making to ensure there is a social license to support the type of data reuse. The role of the data collaborative would change together with the relative weight of the different types of consent and licensing, and the types of governance each entails, including the need to communicate and reinforce mechanisms that ensure quality and safety. In addition, the data collaborative has an important role in ensuring that the data are used for well-defined objectives and purposes, in alignment with the value proposition discussed earlier. It also needs the means to stop data usage if there is misuse. The data collaborative then, could be positioned as a backbone of the ecosystem around data and health by fulfilling distinct roles in the transactional, relational and systemic way of working.

8.1 Impact licensing as strategic instrument

Across the different business models, impact licensing can be used as a strategic instrument to ensure quality and safety and that data usage is aligned with articulated objectives. These outcomes can be tied to improved health outcomes, cost efficiencies, better experience for citizens, or improved quality of life. In addition, an impact license also specifies the terms and conditions for data usage for research purposes and how intellectual property is dealt with. This is a crucial element for generating trust and transparency in collaborations. The impact license also includes enforceability clauses. More information on impact licensing as a tool can be found in the bulletin on social value creation of the IP helpdesk of the European Commission (European Commission. Directorate General for Internal Market, Industry, Entrepreneurship and SMEs., 2020).

8.2 Scenarios for the financial model

While there is a large variety of income streams for different types of business models (freemium, add-on, cross-subsidisation, shared service model, etc) (Fielt, 2013), we focus on providing a framework to make sense of the financial model that has to be linked with the business model. From a high-level perspective, four scenarios can be differentiated: (i) Purpose without Profit, (ii) The Split, (iii) The Combination and (iv) The Open Integrator.

The **Purpose without Profit** is common for non-profit organisations entirely dependent on public funding and/or philanthropic resources for its operations and societal value creation. Very often such organisations are faced with decreasing income streams and are forced to move towards **The Split** scenario in which they set-up commercial or fundraising activities to continue their activities aimed at societal value creation. In the other two scenarios, there is a much more integrated approach in which purposeful activities creating societal value do generate a financial return. The difference between **The Combination** and **The Open Integrator** is the broadened scope of the activities and the value creation at population level in an open model in the latter scenario.

As mentioned in section 6, there needs to be a coupling between the economic and societal value creation, and the ecosystem in which the data collaborative operates and is built on, to ensure that all activities are aligned with its foundational principles. When blending Figure 5 with the examples and scenarios of income streams, we could argue that a layered approach with its gravity point in the relational business model could make sense for many data collaboratives. This is presented in Table 5. This is not a conclusion, nor a suggestion, but rather an example to start the iterative process within a givenconsortium for honing into a 'good enough' starting point.

Table 4. Examples in a layered approach anchored on the relational business model for a data collaborative.

	Personalised products and services for better health (transactional)	**Working together in a club model for better health and quality of life (relational)**	**Collective impact and better health outcomes at population level (systemic)**
Actors involved	Companies and organisations that use data to help a person to find the best suited product or give this person the best possible service based on who they are. This may include health service providers, but may also go beyond that (ie. mobility services).	Partnerships of multiple stakeholders, including companies, non-profit organisations and perhaps certain public actors. They work together to develop and improve integrated solutions using data. These actors already know how to work data driven to develop new products and services.	All kinds of actors. Geared towards creating value at population level. Public and private health and health research service providers become more important.
Client facing identity	Brand of the company using the data and identity of the above mentioned companies and organisations. Together with a data collaborative quality label.	Actors and brands that aim to develop integrated solutions for specific target groups.	Public and private health and health research providers.
Income generating scheme	For every use of the data that the data collaborative holds there is a return or a compensation in a pay per service or a pay per volume model. Or it can be a licensing fee for the use of that sort of personalised services.	A club model with subscription. This can be discriminatory for commercial and non-profit organisations. Licensing fees that also include payments per development and/or sales that comes out of the shared data. The licensing also covers commercial uses.	A health impact scheme with a public provider. Carried forward by a loan, impact bond and/or certain subsidies. Repaid based on criteria and indicators relevant for impact at population level.
Role of the data collaborative	Quality adviser and safety keeper through the combination of informed consent and impact licensing agreements. Social licensing is less important here.	Gatekeeper and custodian to ensure that the shared data are used for specific objectives and purposes. Social licensing becomes more important as an arbiter for approval of the type of activities.	Trusted third party and data aggregator bringing together all actors for collective learning and research, striving for better health outcomes, cost efficiencies and better experiences.

Source: Own creation

8.3 Next steps for data collaboratives

Having laid out the range of possibilities, we recommend that emerging data collaboratives deliberate on and reaches a decision based on the following questions, which follow the components of a business and governance model for a data reuse ecosystem, as we outlined in section 6.

Value proposition

Question 1. What is the main problem that the data collaborativeis trying to address? And what are other important but secondary problems?

Question 2. What value is the data collaborative adding to the data? Does this value added differ depending on the different users of the platform?

Question 3. What are the reasons why each potential user of the platform will want to use it? What are the incentives for them to use it?

Question 4. How much agency will data holders have over the data -extent of openness, type of informed consent needed and how often-?

Question 5. Will the data collaborative serve as a space to store sensitive or personal data and, if so, how will it decide legitimately and systematically on its uses?

Question 6. How will the data collaborative obtain the collective consent and legitimacy (ie. social license) for potential sensitive uses of the data?

Question 7. What contracts and agreements with well-defined objectives for the use of data (ie. impact licenses) will the data collaborative need to have for different users?

Question 8. How will the data collaborative be governed so that the platform is accountable, transparent and trustworthy?

Ecosystem

Question 9. What is the relationship between the data collaborative and other data intermediaries and data spaces in Belgium and in the EU?

Question 10. Is the data collaborative guaranteeing the quality of the data and if so how?

Question 11. What does the infrastructure behind the data collaborative platform enable or restrict (technology affordances)?

Question 12. Are the affordances/characteristics of the infrastructure aligned with the purpose and added value (value proposition) to data that the data collaborative is based on?

Financial proposition

Question 13. Which category of business model is more relevant to the data collaborative’s value proposition: the transactional, the relational or the systemic?

Question 14. What sources of funding do currently exist that are sustainable for a long period of time and what are the terms and possibilities of this time window?

Question 15. Are there any uses of the data or services based on the data that can be responsibly monetised without compromising informed consent, the social license and legitimacy of the data collaborative, and the regulatory context? If so, which ones?

Question 16. What category of financial model (purpose without profit, split, combination, open integrator) are more relevant to the data collaborative’s value proposition and ecosystem?

09

CONCLUSION AND RECOMMENDATIONS

9. Conclusion and recommendations

In this report we have provided an overview of mechanisms and components of business and data governance models for more sustainable data collaboratives. Starting with an examination of the current European and Belgian regulatory context—including the Data Governance Act, Data Act, and the emerging European Health Data Space—we explore the interplay between different business models and governance mechanisms: informed consent, social licensing, and impact licensing. Through the analysis of existing literature and case studies of data cooperatives, health record banks, and data trusts, we also identify key components of sustainable business models and outline three potential approaches: transactional, relational, and systemic. Finally, we also make practical recommendations for the development of emerging data collaboratives, emphasising the importance of a clear value proposition, alignment with stakeholders' interests, and appropriate governance structures. These different elements should help to guide data reuse ecosystems in selecting and implementing an appropriate business model that balances economic viability with their key values and goals.

- Business and governance models for data collaboratives depend on the added value that all stakeholders -both on the supply and the demand-side of data- assign to these data.

  Recommendation 1: We recommend that all stakeholders in a data collaborative explicitly articulate the value that data reuse offers to them and therefore what they seek out of the exchange and how they can contribute to the public interest.

- While a data collaborative can serve different purposes, it is important for the identity, governance and business model that there is agreement between stakeholders and clarity on the main purpose of the entity.

  Recommendation 2: We recommend that all stakeholders in a data collaborative also agree on the main value proposition that their partnership is designed and built to offer.

- A business model goes hand-in-hand with the values and principles that underpin the governance of the collaborative and its approach to decision-making.

  Recommendation 3: We recommend that all founding stakeholders of a data collaborative are in agreement of a charter or Code of Practice that sets out the values and principles underpinning its aims and governance.

- The regulatory context sets some guidelines but, ultimately, the reuse of data for secondary uses (that is uses other than those for which the data was created in the first place) need additional layers of individual and collective consent that address the barriers of individual-only, binary and static consent processes.

Recommendation 4: We recommend that data collaboratives put in place mechanisms for participation of the relevant publics and stakeholders to acquire legitimacy and collective consent for the reuse of data for different purposes other than the ones for which the essential individual informed consent is obtained.

Recommendation 5: We recommend that data collaboratives consider the role of contracts and agreements with well-defined objectives for the use of data (ie. impact licenses) for potential different users of the data.

- The financial model needs to be aligned with the value proposition of a ata collaborative as well as the ecosystem.

Recommendation 6: We recommend that data collaboratives articulate the type of model (transactional, relational, systematic) that is most aligned with its value proposition, and maps out the different sources of income (public funding, private funding, membership revenue, commercial revenue, etc.) and its likelihood in the context of their ecosystem.

REFERENCES

References

Aitken M, Toreini E, Carmichael P, et al. (2020) Establishing a social licence for financial technology: Reflections on the role of the private sector in pursuing ethical data practices. *Big Data & Society* 7(1): 205395172090889.

Azcoitia SA and Laoutaris N (2022) A Survey of Data Marketplaces and their business Models. *Association for Computing Machinery*. https://doi.org/10.1145/3572751.3572755.

Behluli P, (2025) *The Future of Research Funding and Data: Insights from the DeSci.Berlin 2025 Conference.* Medium. Available at: https://medium.com/@TheDataTank/the-future-of-research-funding-and-data-insights-from-the-desci-berlin-2025-conference-7eff1b8b941d

Belgian University Hospitals (2024) Common position establishing a framework for secondary use of data routinely collected in hospitals, https://www.univ-hospitals.be/common-position-establishing-a-framework-for-secondary-use-of-real-world-data-routinely-collected-in-hospitals/

Boulton M and Parker M (2007) Informed consent in a changing environment. *Social Science & Medicine* 65(11): 2187–2198.

Casanovas P, Mendelson D and Poblet M (2017) A Linked Democracy Approach for Regulating Public Health Data. *Health and Technology* 7(4): 519–537.

Colom A, Modhvadia R and Reeve O (2023) *What do the public think about AI? Understanding public attitudes and involving the public in decision-making about artificial intelligence*. Evidence review. Ada lovelace Institute. Available at: https://www.adalovelaceinstitute.org/wp-content/uploads/2023/11/ALI-2023_What-do-public-think-about-AI_Evidence-review.pdf.

Colom A, Poblet M, (2025) A polytopic approach to democratising decision-making on health data reuse in the EU. *Data & Policy.* [Forthcoming issue]

Creus J (2024) Salus.coop,data cooperative for health research, in Data Cooperatives Report (People-Centered Internet eds.), UN Science Summit Digital Governance Series, https://peoplecentered.net/data-cooperatives-report-2/

Das S (2023) Private UK health data donated for medical research shared with insurance companies, *The Guardian*, 12/11/2023 https://www.theguardian.com/technology/2023/nov/12/private-uk-health-data-donated-medical-research-shared-insurance-companies

Dr. Daniël De Coninck Fund (2024) Eight guiding principles for caring technology. Available at: https://www.fondsdanieldeconinck.be/en/initiative/caring-technology/.

Empowering Patients Through Community (2024) Patients Like Me. Available at: https://www.patientslikeme.com/about

EUCOPE (2024) The European Health Data Space: concerns and opportunities following ongoing trilogues. Available at: https://www.eucope.org/the-european-health-data-space-challenges-and-opportunities-with-respect-to-the-different-proposals-currently-subject-to-the-trilogue-negotiations/

European Commission. Directorate General for Internal Market, Industry, Entrepreneurship and SMEs. (2020) Report on the feasibility and analysis of “Impact Licensing Initiative” (ILI) for technology access during a health crisis. LU: Publications Office. Available at: https://data.europa.eu/doi/10.2873/552686

European Commission (2020). European Data Strategy. Available at: https://digital-strategy.ec.europa.eu/en/policies/strategy-data

European Commission (2024a) Implementing the Data Governance Act – guidance document. *Shaping Europe’s digital future*. Available at: https://digital-strategy.ec.europa.eu/en/library/new-practical-guide-data-governance-act

European Commission (2024b). Proposal for a Regulation Of The European Parliament And Of The Council on the European Health Data Space. Available at:https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0197

European Commission (2024c) European Health Data Space. Available at: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

European Commission (2025) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Available at: https://digital-strategy.ec.europa.eu/en/library/ai-continent-action-plan

Ferro E and Pizzamiglio A (2023) *New Business Models for Data-Driven Services.* European Union: Publications Office of the European Union. Available at: https://data.europa.eu/doi/10.2830/034880.

Fielt E (2013) Journal of Business Models. *Journal of Business Models* 1(1). Journal of Business Models: 85–105.

Gille F and Vayena E (2021) How Private Individuals Maintain Privacy and Govern Their Own Health Data Cooperative: MIDATA in Switzerland. [object Object]. Available at: http://hdl.handle.net/20.500.11850/476759 (accessed 10 September 2024).

Hafen E (2019) Personal data cooperatives–a new data governance framework for data donations and precision health. In: Kurtzinna J and Floridi L (eds) *The Ethics of Medical Data Donation*. e. Oxford, UK: SpringerOpen, pp. 141–149.

Hardinges J, Wells P, Blandford A, Tennison J, & Scott A (2019) *Data trusts: lessons from three pilots.* Open Data Institute (ed.). Available at: https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit?tab=t.0#heading=h.3fngvdcfo2cs

HDA (2023) Annual Report. HDA Available at:https://www.hda.belgium.be/en/about_us#section-annual-report

HHDC (2021) Position Paper on Health Data Cooperatives. Gezond Akkoord. Available at: https://mdog.nl/wp-content/uploads/2021/11/1-GA-Position-paper-Gezond-AKKOORD.pdf

INAMI (2024) BIHR (Belgian Integrated Health Record): Un système d’information intégré pour des soins de qualité. INAMI. Available at: https://www.inami.fgov.be/fr/themes/esante/bihr-belgian-integrated-health-record-un-systeme-d-information-integre-pour-des-soins-de-qualite

Kariotis T, Ball MP, Tzovaras BG, et al. (2020) Emerging health data platforms: From individual control to collective data governance. Available at: https://doi.org/10.1017/dap.2020.14.

KBF (2022) Zorg voor je data. Available at: https://media.kbs-frb.be/nl/media/8912/2022_PUB_3832_ Zorg%20voor%20je%20data

Krewer J and Warso Z (2024) *Digital Commons as Providers of Public Digital Infrastructures*. October. Open Future.

Lauer R, Merkel S, Bosompem J, et al. (2024) (Data-) Cooperatives in health and social care: a scoping review. *Journal of Public Health*: 1–19.

Lauf F, Scheider S, Friese J, Kilz S, Radic M, Burmann A (2023) Exploring Design Characteristics of Data Trustees in Healthcare-Taxonomy and Archetype. Available at: https://core.ac.uk/download/pdf/567667434.pdf

Marelli L, Stevens M, Sharon T, Van Hoyweghen I, Boeckhout M, Colussi I, Degelsegger-Márquez A, El-Sayed S, Hoeyer K, van Kessel R, Zając DK (2023) The European health data space: Too big to succeed?. *Health policy*, 135, p.104861. Available at: https://www.sciencedirect.com/science/article/pii/S016885102300146X

Marcucci S, Alarcón NG, Verhulst SG, et al. (2023) Informing the Global Data Future: Benchmarking Data Governance Frameworks. *Data & Policy* 5: e30.

Metcalf K and Sadowski (2024) Asserting the public interest in health data: On the ethics of data governance for biobanks and insurers. *Big Data & Society*, *11*(4). https://doi.org/10.1177/20539517241290215

Micheli M, Farrell E, Carballa-Schmichowski B, et al. (2023) *Mapping the landscape of data intermediaries Emerging models for more inclusive data governance*. JRC Science for Policy Report. Luxembourg. Available at: https://op.europa.eu/en/publication-detail/-/publication/379cf850-4553-11ee-9854-01aa75ed71a1/language-en.

MIDATA (2019) Articles of Association. Available at: https://www.midata.coop/wp-content/uploads/2019/08/MIDATA_Statuten_20190626_EN.pdf

MIDATA (2024) About us. Available at: https://www.salus.coop/qui-som/

Milne R, Sorbie A and Dixon-Woods M (2022) What can data trusts for health research learn from participatory governance in biobanks?. *Journal of Medical Ethics*, 48(5), pp.323-328.

Moffat K, Lacey J, Zhang A, et al. (2016) The social licence to operate: A critical review. Forestry: An International Journal of Forest Research 89(5): 477–488.

Moyersoen J and Vandermeulen B (2023) *Societal Value creation*. Bulletin 5, December. European IP Helpdesk. Available at: https://op.europa.eu/en/publication-detail/-/publication/a51c1121-9fb9-11ee-b164-01aa75ed71a1/language-en.

Open Data Institute (2018) *What is a Data Trust?* Available at: https://theodi.org/article/what-is-a-data-trust

OECD (2017) *Business models for sustainable research data repositories*. 47, OECD Science, Technology and Industry Policy Papers, 6 December. Available at:

https://www.oecd-ilibrary.org/science-and-technology/business-models-for-sustainable-research-data-repositories_302b12bb-en (accessed 4 September 2024).

Open Humans (n.d.) About. Available at: https://www.openhumans.org/about/.

Salus Coop (2024) Qui som. Available at: https://www.salus.coop/qui-som/

Susha I, Janssen M and Verhulst S (2017) “Data collaboratives as “bazaars”? A review of coordination problems and mechanisms to match demand for data with supply”, transforming government. People, *Process and Policy*, 11(1), 157–172. https://doi.org/ 10.1108/TG-01-2017-0007.

Susha I, Flipsen M, Agahari W, et al. (2020) Towards Generic Business Models of Intermediaries in Data Collaboratives: From Gatekeeping to Data Control. In: Viale Pereira G, Janssen M, Lee H, et al. (eds) *Electronic Government*. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 304–315. Available at: https://link.springer.com/10.1007/978-3-030-57599-1_23 (accessed 22 August 2024).

TEHDAS (2022a) Report on secondary use of health data through European case studies. Available at: https://tehdas.eu/app/uploads/2022/08/tehdas-report-on-secondary-use-of-health-data-through-european-case-studies-.pdf

TEHDAS (2022b) TEHDAS Consultation: People’s support of health data use. Available at: https://tehdas.eu/tehdas1/results/tehdas-consultation-people-support-health-data-use-with-solid-safeguards/

Unforgettable Me (2024) Getting Started. Available at: https://unforgettable.me/for-users/

UK Biobank (2024) About us. Available at: https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us

Van Obberghen et al. (2024) Digital health laws and regulations, https://iclg.com/practice-areas/digital-health-laws-and-regulations/belgium

Verhulst S and Saxena S (2022) *The Need for New Methods to Establish the Social License for Data Reuse*. Data & Policy Blog. Available at: https://medium.com/data-policy/the-need-for-new-methods-to-establish-the-social-license-for-data-reuse-e7c67bdc4aff.

Wong W, Duncan J and Lake DA (2024) Why data about people are so hard to govern. *Regulation and Governance.* Available at: https://doi.org/10.1111/rego.12591.

Authorship

This report has been written by [in alphabetical order] Anna Colom (The Data Tank), Randy Mellaerts (Impact Licensing Initiative) and Marta Poblet (The Data Tank). It has also benefited from review and input by Johan Moyersoen and Bruno Vandermeulen (Impact Licensing Initiative).

Recommended citation

The Data Tank and Impact Licensing Initiative (2025) *Governance and business models for sustainable health data collaboratives.* The Data Tank. https://www.datatank.org/news-publications

Licensing

Impact Licensing Initiative

The Data Tank

Belgium, July 2025