GeneConsent FOSS Community

The key to unlock genomic data value

Peter Walgemoed, GeneConsent.eu

LabSpace Coop

Mission Driven
Personal Membership
Data Co-operative

 © GeneConsent.eu 2022

Table of Contents

Table of Contents	2
Executive Summary	3
Market ecosystem	5
GeneConsent Solution and Values	6
Go-to-market strategy	7
GeneConsent as Free Open Software Service platform	10
GeneConsent platform data service space perspectives	11
LabSpace.coop background	14

 © GeneConsent.eu 2022

Executive Summary

GeneCoop is a NGI-ledger project carried out by Carelliance and Waag Society. The project has a history from both Carelliance and Waag working on bio/data commons and genetic data services. The project led to a Minimal Viable Product called GeneConsent. This product is a software service that is literally the key to unlocking genomic data value as a bio-commons.
GeneConsent gives participants in genetic research full control over their data, by providing a *dynamic informed consent* service. This means that the individual does not at one time grant access to all their data, but only to fractions of it, and for specific (research) operations. As a result, the researchers get precisely (and only) the data they need. Next to the increased control in the hands of the research participant, this model gives researchers the confidence that they are operating within the boundaries of the law. This will become important in the European context because of the GDPR and upcoming Data Governance Act.

GeneConsent's stakeholders described the value of the service as follows:

How to use the full power of data to improve health and healthcare of mankind
The road towards data driven preventative healthcare
Founding a trusted data commons for the advancement of science and societal value.

GeneConsent operates in the research and development market of genomic data. This market is divided in:

- A public financed science/research segment
- A private/industry financed segment with specific direct to consumer genetic test sub-segment
- A plural financed segment.
- Potentially +/- 200 mln genomic data sets of individuals require dynamic informed specific consent (All segments combined)

Researchers operate in all three segments. GeneConsent targets the innovator research behaviour sub-segment, because they influence the other researchers. They are working for the top academic medical centers and research labs in the industry.

At this moment the consent holders are given a **generic** consent that is made by lawyers on behalf of the organisation that wants to use their data. It's not clear how their data is processed and for which specific purpose it is used or re-used (Secondary use).

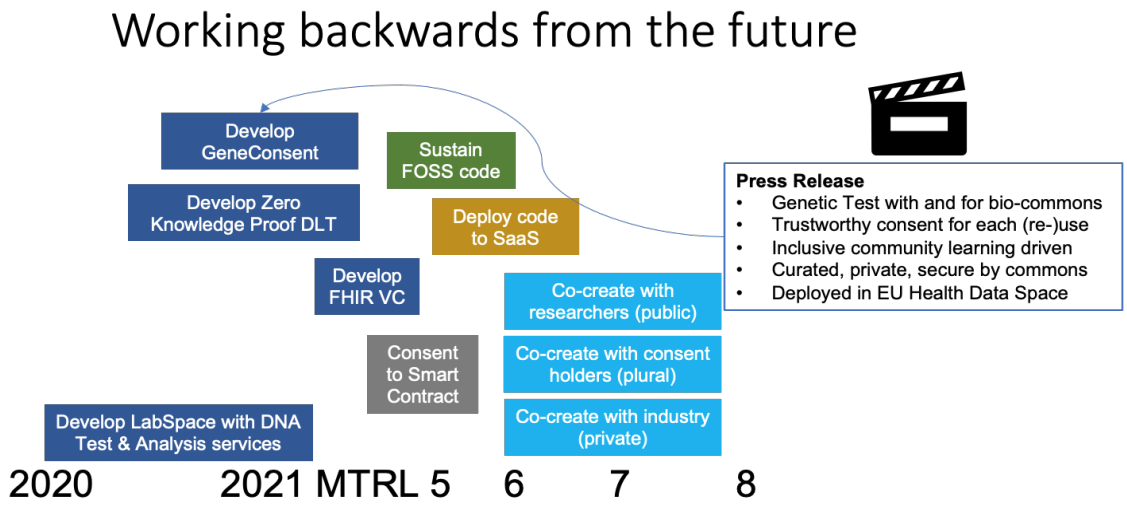
Value proposition GeneConsent

- Participants in genetic research get full control over their data, by providing a *dynamic informed consent* service.
    - The participant will be able to grant or withdraw access to for even fraction's of their data for specific (research) operations at any time.
    - The researcher gets only the data they need.
    - The researcher is confident operating within the boundaries of the law (EU GDPR and upcoming Data Governance Act)
    - The researcher becomes as knowledge contributor part of a network of consented curated data that is shared via a commons catalogue. In this way they gain access to a growing curated data collection that supports their research.
- Arrange GeneConsent with privacy by EU citizen consent holder design as a co-creation process and show data to knowledge value creation per market segment.

 © GeneConsent.eu 2022

- Key of GeneConsent's go to market strategy is to adapt its software service to each segment and especially the way these are financed in the future, because of the need for plural-public-private collaborations.
- **Consent service market fee/valuation check** 1% of project/service transaction value.

The business model projects a 1st revenue model of 600.000, with a market share of 15% and a 1% commission on the DNA market test price for the consent service. The analysis/collection value taken has an exponential profile. The number of analysis options that every research, care/diagnosis or personal heritage/lifestyle does on these data points is enormous. This value is not achieved at this moment, because each silo keeps the test data for itself without the consent holder knowing this. If the test data set is shared under a commons, it can be reused for any research/discovery analysis with consent. This explodes with the combination of analytics/AI tools, which combine individual consents into new collections for their value purpose.
A 2nd revenue stream is created by turnkey project service model with research consortia to tailor the consent in a co-creation process between researchers and consent holders.



With the NGI ledger challenge a MVP with Market Technology Readiness Level[1] (MTRL) 4 "Small prototype" has been achieved. The next steps are to find investors for further development in:

- MTRL 5 Large Scale prototype with 1000 citizen consents and the start of a bio-data commons movement to curate Genetic Data . This is supported by a small-scale stakeholder campaign and extended to
- MTRL 6-7 with a beta version of the SaaS platform and a large scale early adopter campaign that are supported by co-creation processes with consent holders and public and private researchers.
- MTRL 8 where the SaaS is delivered for 600.000 consents that are supported by 10 research consortia.
- MTRL 9. Full commercial SaaS platform. The 600.000 consents lead to an analysis re-use value by the 10 research consortia of € 3.000.000.

[1] MTRL model by SWForum.eu. https://swforum.eu/events/swforumeu-webinar-market-technology-readiness-levels

 © GeneConsent.eu 2022

Market ecosystem

The total market ecosystem and their stakeholders with the human/society at the center is shown in the ECHAlliance figure below.



Digtial Health Society
ECH Alliance
Ecosystem

The DNA test that is used for the public research market is also used by the private market. It could become the basic test for NGO/citizen data co-operatives. The figure below shows its analysis potential.

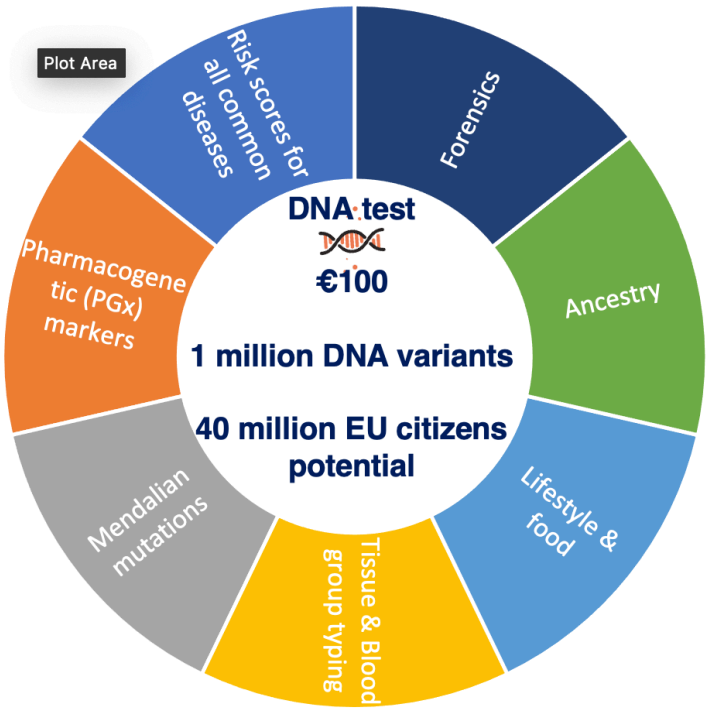


Figure based on a presentation made by Prof. André Uitterlinden, Erasmus MC Rotterdam Complex Genetics and Head of the Genomics Core Facility

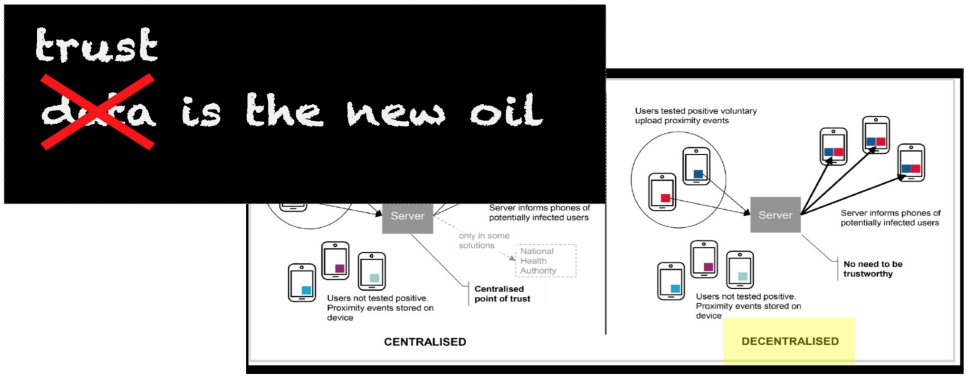
 © GeneConsent.eu 2022

The problem

- We give **our DNA data** to industry (**B2C**) in exchange for their **service** (Consumer genetics & Pharma/big tech).
- We let them create the **value** with it on the **collective/market** level, **without** our **informed consent**.
- We allow that **our DNA data** is locked up and **underused** by **research/academia.**
- We let our governments (**G2C**) sponsor this with **generic consent, without** us knowing **who adds knowledge.**

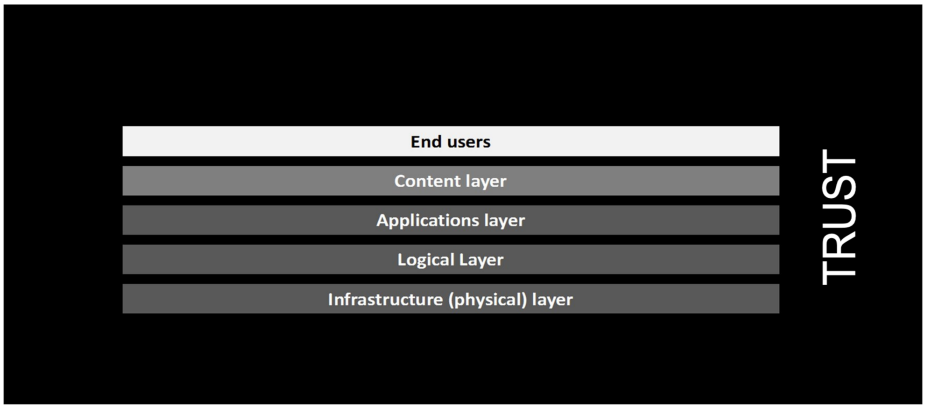


GeneConsent Solution and Values



Source; Prof. Andrea Renada. ***Trust in digital technologies: from contact-tracing apps to policies for Trustworthy AI***

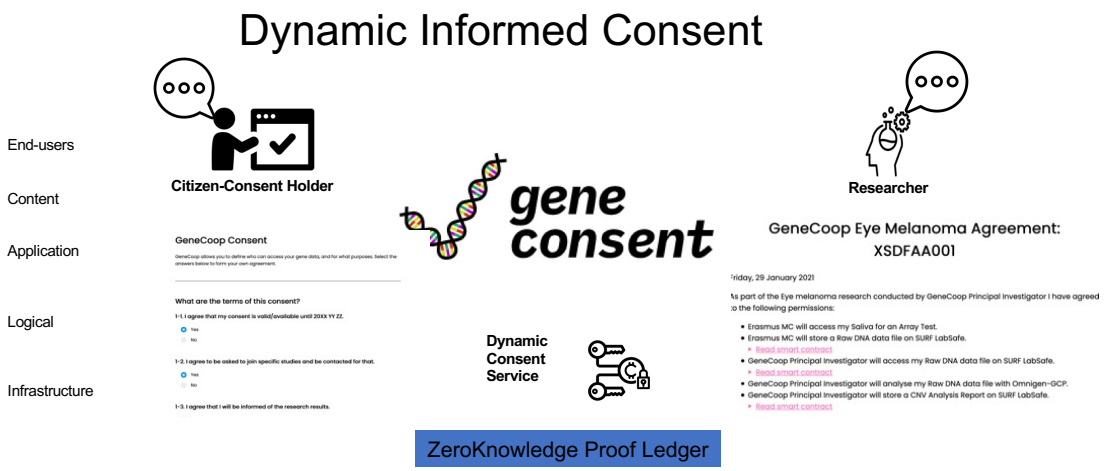
GeneConsent needs to regain the trust of its consent holders. This is not limited to technology, but requires a transparent professional organisation that proves its trustworthiness by taken care of their end user customers. Trust is build by integrating the following service layers.



Source; Prof. Andrea Renada. ***Trust in digital technologies: from contact-tracing apps to policies for Trustworthy AI***

The figure below describes the different service layers and the consent perspective (and language) of the consent holder (citizen) and the researcher's perspective. The consent can be regarded as a smart contract that is executed in a trusted space with the agreed data set and analysis services. The above implementation provides European Researchers (consortia) with a DNA Service LabSpace for its data and knowledge sharing.

 © GeneConsent.eu 2022



Go-to-market strategy

GeneCoop is a NGI-ledger project carried out by Carelliance and Waag Society. The project has a history from both Carelliance and Waag working on bio/data commons and genetic data services. The project led to a Minimal Viable Product called GeneConsent. This product is a software service that is literally the key to unlocking genomic data value as a bio-commons.

GeneConsent gives participants in genetic research full control over their data, by providing a *dynamic informed consent* service. This means that the individual does not at one time grant access to all their data, but only to fractions of it, and for specific (research) operations. As a result, the researchers get precisely (and only) the data they need. Next to the increased control in the hands of the research participant, this model gives researchers the confidence that they are operating within the boundaries of the law. This will become important in the European context because of the GDPR and upcoming Data Governance Act.

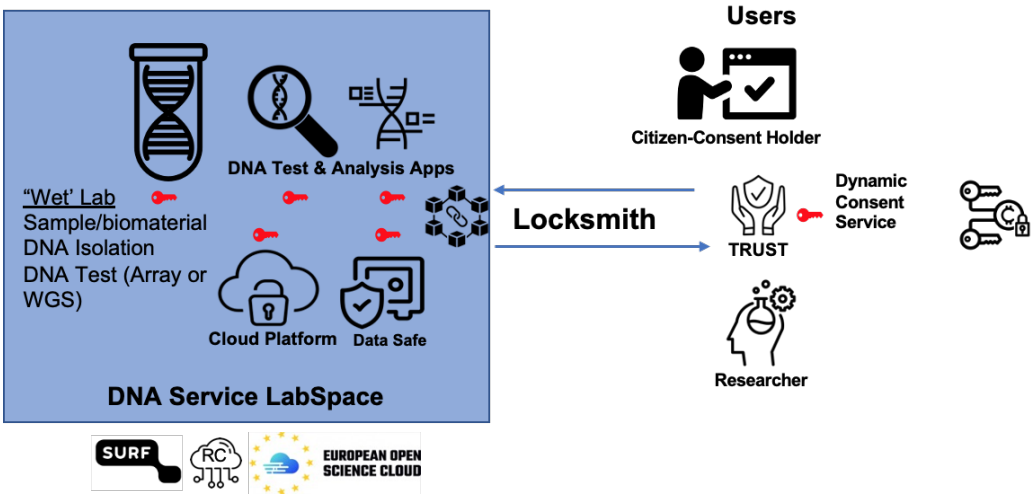
Outlook

One of the important next steps for the venture is to test the consent service with users. After all, it is crucial for consent holders to understand what they give consent for. Otherwise, the consent is hardly 'informed'. This goes hand in hand with the idea that consent should be seen as a process, and not an action. Consent holders are ideally informed about what happens with their data, and should be able to revoke consent at any moment. In this process, a clear user interface is crucial, as well as close contact with a medical expert community who can explain the consequences of giving consent in various situations.

Even when the GeneConsent is fully functional, it will not fully cover the challenges of handling and sharing genetic data in a research context. Next to the consent service, solutions need to be developed for the secure storage of genetic data, as well as for performing research operations in a secure digital environment. These areas are relatively

 © GeneConsent.eu 2022

unexplored and raise many questions. For example: what research operations carried out by the researcher are communicated to the consent holder?



The LabSpace.coop facilitates the development and discovery of the above challenge, by providing a Data Space and technology platform to its members. It's role in terms of NGI eSSIF lab architecture is called "jurisdiction[2]" It's a "data network" in terms of Sitra's fair data economy[3]. The LabSpace.coop platform is built on the SURF/European Open Science Cloud infrastructure and uses the following services:

1. SURF Research Cloud: Software defined Data, Applications, Storage & Compute. SURF is a key node of the European Open Science Cloud and is the trusted gateway to the research and education end-user community.
2. Data services DNA Test Lab, Ingest, Data Archive, Safe and Catalogue to curate data independently from (processing) applications operated by Carelliance.
3. Data processing (analysis) applications with partners in a CmyDNA platform catalogue operated by Carelliance.
4. GeneConsent service operated by the Gene-LabSpace.coop community.

GeneConsent is the NGI ledger software development project. It has a public research funded use case where a Clinical Researcher-Doctor specialized in Eye melanoma (Severe cancer) requests consent to a citizen for a test to analyze their DNA data. The consent is dynamic, because the citizen can withdraw it at any time and it needs a new consent at the moment the analysis/process or research (secondary use) purpose changes. GeneConsent role can be seen as the locksmith that designs, develops, builds and deploys the (ledger/de-central) lock and key technology. This powers the dynamic consent service for the end-users.

The above implementation provides European Researchers (consortia) with a DNA Service LabSpace for its data and knowledge sharing.

---

[2] **Jurisdiction** is a foundational concept for organizing collaborations between **parties**, e.g. in **communities** or **ecosystems**. Basically, a jurisdiction is an **authority** that has mechanisms for defining and maintaining rules, enforcement thereof, and a mechanism for resolving conflicts within its **scope of control**.

[3] Data Network is group of organizations and/or individuals that build data sharing solutions. / A group of companies and other organizations or even individuals that share data in accordance with a rulebook or other contractual arrangement. https://www.sitra.fi/en/blogs/what-is-a-data-network/

 © GeneConsent.eu 2022

A key program/project for GeneConsent is the European 1 million Genome project. The preparation for this initiative driven by the European member states ministers is at this moment taken place. The proposal will combine existing data from (academic) cohort/epidemiological studies with new programs. GeneConsent's academic medical center network has the consortium lead and strong cohort data position. GeneConsent solves two problems the consortium faces.

1. The existing consent was given for genome wide association studies that were not identifiable to individual patients. To allow the new 1 million research/discovery use a new consent is likely needed.
2. The existing DNA data collection is highly biased to a white/European population. It does not represent all ethnicities/races, which are needed for a good reference genome. Especially in big multicultural cities like Amsterdam and Rotterdam the population has a global profile. This last part opens the opportunity to start Gene-(Lab)Space.coop that curates the data on behalf of the consented citizens as a bio-commons. Because citizens are organized in cities/regions/countries a federated model like the EU can be built bottom-up. GeneConsent will explore this with the economic boards of municipalities because they have a long-term interest for their citizens. Value drivers for them:
    a. Inclusive co-creation with citizens to drive informed consent process
    b. Acceleration of population based healthcare data exchange and discovery
    c. NGI decentralized self-sovereign identity and zero knowledge proof technology services will provide faster access to eIDAS services than waiting for national programs that are struggling with the existing legacy identity services infrastructure.
    d. Dashboards for municipal/regional/country/EU policy makers based real time consented citizen data

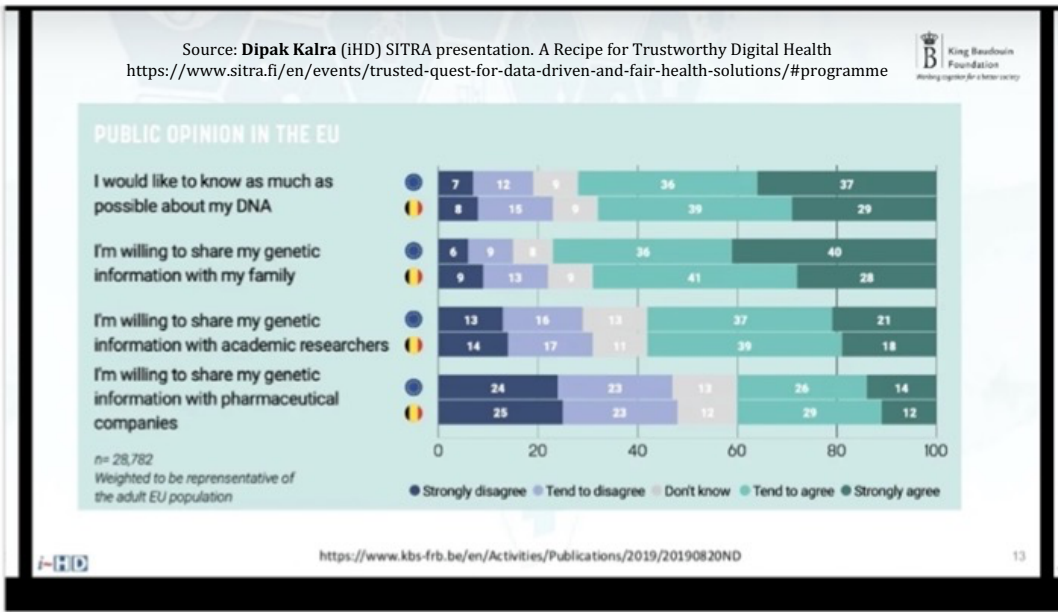
To kick-start the co-creation process the following pilot is investigated.

- Citizens are recruited via a co-creation process led by a municipality partner or philanthropic organization.
- Pilot proposal: 1000 biodiverse Dutch or European citizens.
- Dutch Distribution in three groups.
    1. Rotterdam (Metropole) region: 300 citizens
    2. Amsterdam (Metropole) region: 300 citizens
    3. Rest of the Netherlands (G40 Municipalities): 400 citizens
- Each Gene-LabSpace.coop starts as a project, where the citizen (consent holder) commons become the researcher of their data collection. This model can become a blueprint for (smart) municipalities to start similar co-creation processes and federate them via the European Health Data Space.
- The advantage of setting up a national DNA biobank and use it for all citizens to improve population health is shown by Estonia. They are using the DNA test shown at page 5 and have doubled their participants to 200.000 since 2018. Estonia is a great example for innovative European municipalities with similar sized 1.000.000 populations.

The importance of setting up co-creation processes from a (Health Data) Space society perspective is discussed by this OECD Observatory for Public Sector Innovation thought leader session.

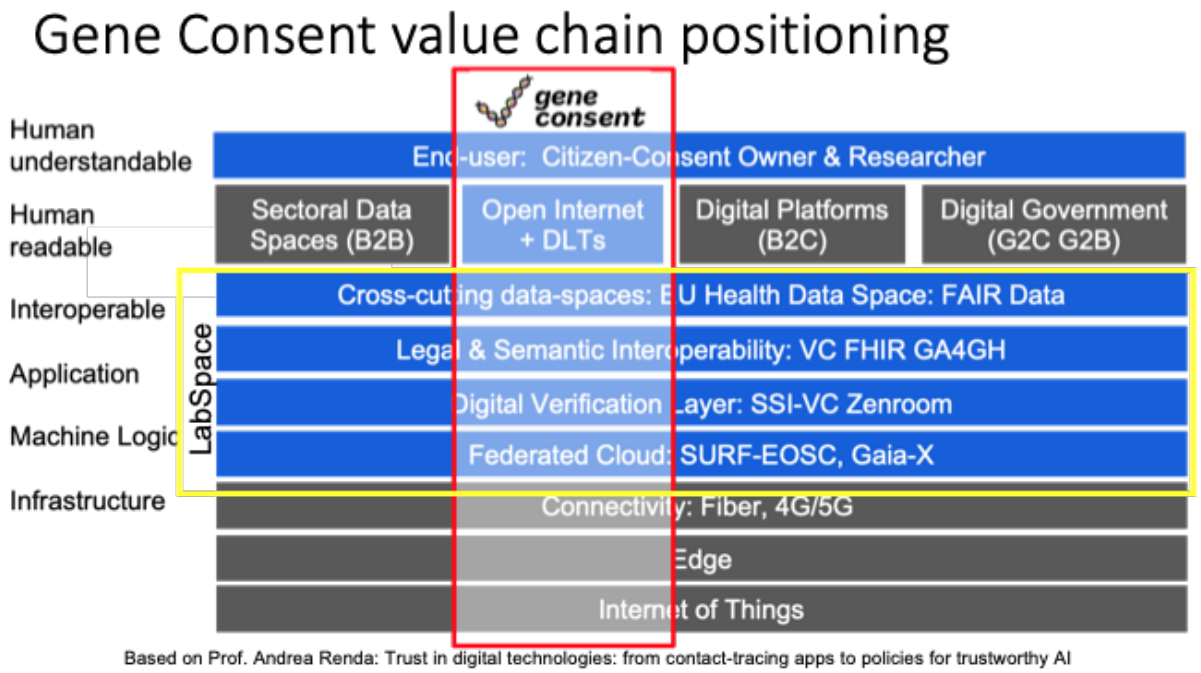
King Boudewijn Foundation did a European citizen study on DNA data sharing. The figure below shows that the majority of citizens are interested to know as much as possible about their DNA. They are also willing to share their data with their family. A bit less to share for (academic) research and much less with (pharmaceutical) companies.

 © GeneConsent.eu 2022



GeneConsent as Free Open Software Service community

The figure below shows the service stack of GeneConsent. The challenge is the orchestration of the FOSS design and development community and its modulair roadmaps. All components need to be integrated into a SaaS platform that can be run in the federated research cloud infrastructure.



Based on Prof. Andrea Renda: Trust in digital technologies: from contact-tracing apps to policies for trustworthy AI

Service layer	GeneConsent FOSS Platform
End user	User experience with content match between consent holder and researcher (Informed consent trusted and understood)

 © GeneConsent.eu 2022

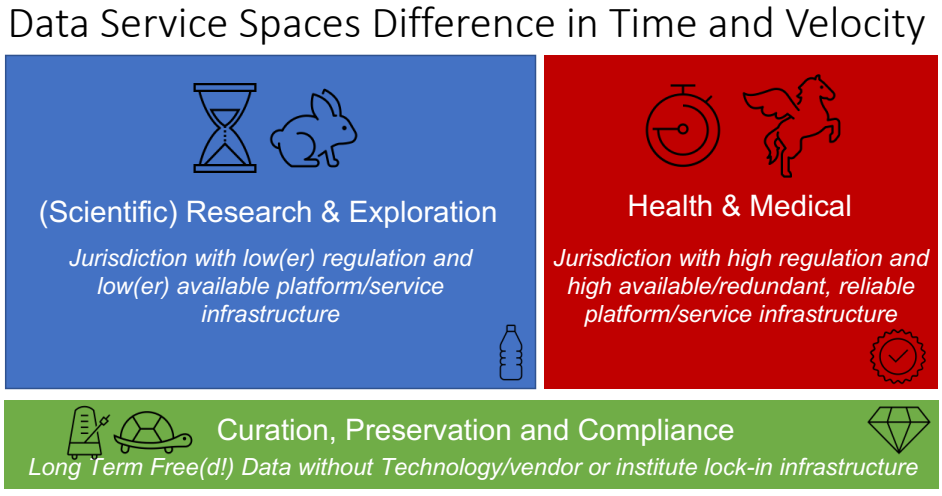
Content layer	Content matched to legal, ethical, and domain specific semantic and interoperability standards (W3C, DLT, FHIR-VC)
Application layer	Web plug-in GeneConsent
Logical layer	Execution of consent with data, application/processing, and user transactions in trusted LabSpace with DLT control. (SSI-Zenroom and SURF/EOSC identity management)
Infrastructure (physical) layer	Software defined research cloud infrastructure SURF/EOSC

GeneConsent community data service space perspectives

At the moment that the end-user is the citizen consent holder, who orders and owns the DNA test (instead of the researcher) like proposed for the 1000 citizen pilot, the Open Internet DLT model turns from G2C and B2C into C2G and C2B. The citizen with their consent determines the re-use/analysis process on their DNA data. As a result, GDPR compliance is arranged with privacy by design with the citizen as controller of their DNA data.

It is important to view the GeneConsent community development from different data service space perspectives. The spaces will co-exist with existing B2B, B2C, G2C and G2B service platforms. The figure below shows that the Research & Exploration space and Health & Medical space have different jurisdiction, regulation, and compliance requirements. As a result the time (HourGlass vs. stopwatch) and velocity (Rabbit vs. super horse) are different. The third curation, preservation and long-term compliance space is completely absent at this moment. It only exists in other sectors like cultural heritage, media and aerospace. The metronome and turtle represent it's time and velocity attributes. The third sustainability dimension is represented by the

- (Recyclable) bottle for science and research, which is a short lived consumable, because of its experimentation focus. However it's publication results with connected (FAIR) data collection needs to be accessible for generations.
- The seal of quality and trust and compliance for health and especially medical.
- The diamond for the sustainable curation, preservation and compliance.



 © GeneConsent.eu 2022

Software components

Component	Functionality	FOSS	Developer community
GeneConsent Front End	Consent Holder Researcher Interface	Yes	GeneConsent Django Framework. Mozilla browser plug-in https://github.com/LedgerProject/GeneCoop
ZenRoom	DLT ZeroKnowledgeProof	Yes	Dyne.org https://github.com/dyne/Zenroom
FHIR VC	Healthcare data verifiable credential	Yes	https://github.com/smart-on-fhir/health-cards/blob/main/docs/credential-modeling.md
FHIR VC	Resource Consent	Yes	https://build.fhir.org/consent.html
SURF/EOSC Research Access Management	Federated identity and access management	Yes	https://wiki.surfnet.nl/display/SRAM
SURF Research Cloud	Federated Cloud for research collaboration	Yes	https://researchcloud.surf.nl/
REUSE	Software service license manager	Yes	https://reuse.software/

Identity and access management Roadmap

1. Federated Researcher Identity provider (SURF Research Access Management/EU-Open Science Cloud)
    a. First test case with 10 consent holders that are known by Researcher Use Edu-ID/ OpenID as identity provider connected to SURF-SRAM.
    b. Second scale out pilot test case for 1000 citizens connected to (EU) citizen relation based data and services access marketplace.
2. SSI for consent holder. Based on Verifiable Credentials.

REUSE

The REUSE service is a key component of the GeneConsent FOSS community. By integrating it with the consent service application/processing transactions it could evolve to a smart contract manager in trusted research cloud spaces.

The Commons Conservancy

All of the above FOSS code needs to be preserved by design. The Commons Conservancy mission and smart governance structure could fit this purpose.

Data curation, privacy and security by design

It's key to separate the data from the applications and decouple identity from services. A FAIR Array Genetics service workflow for ordering a DNA test at a wet lab and curating the data from the lab to the SURF Research Cloud has been developed. The data is archived for the long term independent from the lab and applications that created it. An active data set copy is made available in the Research Cloud. Access to this data set is controlled via GeneConsent. The consented data, metadata and analysis applications processing can only

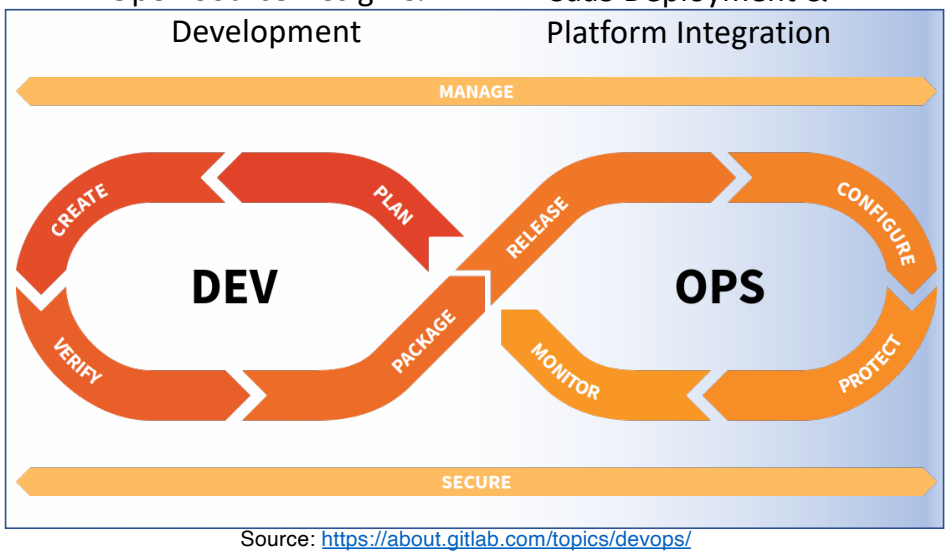
 © GeneConsent.eu 2022

be executed with a signed consent. This solution service process is orchestrated by the LabSpace.coop.

FOSS to SaaS platform(s) development

The development of the open source software components into the GeneConsent SaaS follows the DEV-OPS deployment model below. The ecosystem relation between the FOSS development community and SMEs that provide the SaaS components and integrated GeneConsent SaaS platform solution to customers is key. This free open-source software components development supported by SME ventures for operation and support services is part of the LabSpace.coop orchestration process.

Open Source Design & Development — SaaS Deployment & Platform Integration



Source: https://about.gitlab.com/topics/devops/

FOSS Development Community <> SME SaaS Platform Community supported by big/med tech scale-up/out

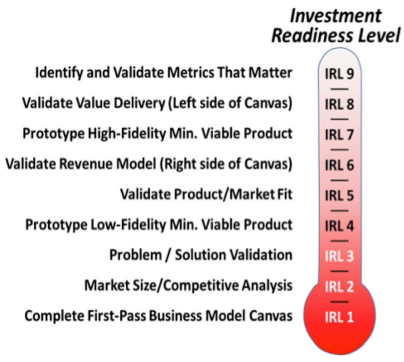
Market Technology and Investment Readiness Levels

The figure below shows the steps to develop software into industrial or publics services for end-users.





- Idea
- Prototype
- Users <> Validation <> Testing & Traction
- Paying Users <> Protection <> Contracts



 © GeneConsent.eu 2022

LabSpace.coop background

LabSpace Coop

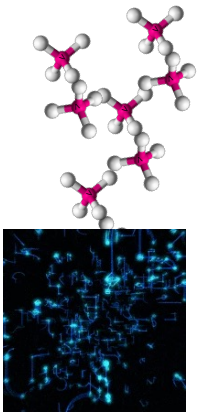
Federated cooperatives curate data as commons to improve health and care of biodiversity

The LabSpace.coop is driven by

Data are “anti-rival” in their nature. They are **products that gain value** while being used (Elinor Ostrom)

Data needs to be **curated by a sustainable organization**. It will **die fast** due to **exponential growth**, dependency on a **complex** application and infrastructure **vendor spaces** and **shortening technology lifecycles**.

Data value creation is a **multidisciplinary team** effort that **synthesizes knowledge with consent, data curation and technology orchestration** in a co-operative federation.



Data Value synthesis

LabSpace Coop

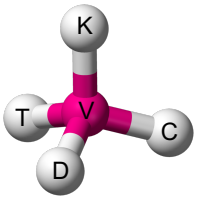


DNA as basis for biodoverse coop life

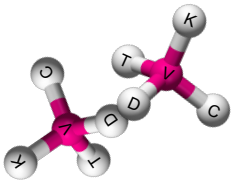


DNA components (A, T, C, G) as Life code synthesis base

Individual, community and Societal **V**alue
= Synthesis of
Consent **H**older
Knowledge **C**ontributor
Data Curator
Technology **O**rchestrator

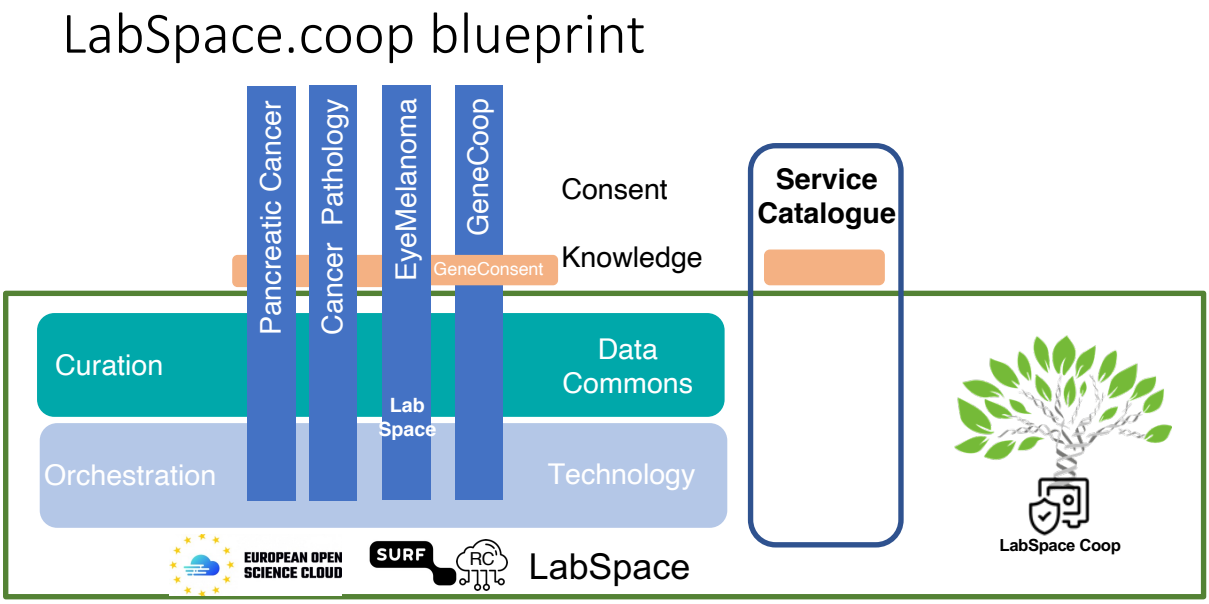


LabSpace.coop DNA components
(CH, KC, DC, TO)



LabSpace A+B Data Commons
= LabSpace.coop UA Federation

 © GeneConsent.eu 2022



At LabSpace.coop we identify four personas. This narrative describes what they mean and why we believe it's important they co-operate to create value for their community and society.

As a person I want to enjoy listening to music (consumer) from an artist. The artist (provider) wants to make it a personal experience and needs to know some trusted data from me to create the most value. We both sign an agreement where I consent the artist (provider) to use my data and contribute his knowledge to me. We are both supported by a technology (platform) to share our data and experience. We both want to be free to change the technology and share our experience with our communities and society to get more insight. For this reason, we both become member of the LabSpace.coop that curates our data and knowledge contribution as a commons, independent from technology/platform provider(s). Value creation is a synthesis process that combines consented curated data with knowledge contributions and orchestrated technology. The resulting (platform) product is what the consent holders and knowledge contributors as consumers have co-designed and implemented with their data and knowledge contribution collection and technology orchestrator provider partners. The technology is a combination of open existing platforms with new FOSS developed SaaS components that are integrated in LabSpace community platforms. This is a continuous experimentation and sustainable deployment process.

In case of health data any individual that shares his data with a researcher or care provider to contribute their knowledge forms the basis of their trusted consent agreement. By orchestrating both like-minded individuals with knowledge contributor networks that are not funneled by silo organizations or monopolists, an open biodiverse community and ecosystem is promoted. The consent is dynamic and specific for the data and knowledge contribution that is shared as well as the trusted technology that is used to process it. The trusted technology is developed by using FOSS DevOps with the connected SaaS venture network for deployment, integration and support. The Data is curated by the community as a commons following the coop's principles of fair biodiversity use.

It is key to orchestrate the ecosystem for the above development. By design this needs to be an open diverse network that is driven by its local community and stakeholders. At the same time the communities are connected globally driven by shared interests and missions.