KRAKEN

BROKERAGE AND MARKET PLATFORM FOR PERSONAL DATA

D6.10 Second Advisory Board

www.krakenh2020.eu

This project has received funding from the European Union's Horizon 2020 (H2020) research and innovation programme under the Grant Agreement no 871473

D6.10 Second Advisory Board

Grant agreement	871473
Work Package Leader	INFOCERT
Author(s)	Matteo Marinelli (INFOCERT) / Pasquale Chiaro (INFOCERT)
Contributors	Romualdo Carbone (INFOCERT), Francesca Podagrosi (INFOCERT), Karl Koch (TUG), Angel Palomares (ATOS), Alberto Miranda (Atos), Davide Zaccagnini (LYNK), Donato Pellegrino (TEX), Davide Porro (INFOCERT), Stefan More (TUG), Wim Vandevelde (KUL)
Reviewer(s)	Sara Diez & Juan Carlos Pérez (ATOS), Valerio Cini (AIT)
Version	Final
Due Date	30/11/2021
Submission Date	29/11/2021
Dissemination Level	Public

Release History

Version	Date	Description	Released by
v0.1	19/11/2021	Initial version	Matteo Marinelli
v0.2	26/11/2021	Version for reviewers (Atos and AIT)	Matteo Marinelli
v0.3	26/11/2021	Final version	Matteo Marinelli
v1.0	29/11/2021	Submitted version	ATOS

Table of Contents

List of Tables

List of Figures

List of Acronyms

Acronym	Description
LSTS	Science, Technology and Society Studies
EU	European Union
GDPR	General Data Protection Regulation
ESSIF	European Self Sovereign Identity Framework
IoT	Internet of Things
API	Application Programming Interface
Mx	Month 1, Month 2…
PAB	Project Advisory Board
EBSI	European Blockchain Services Infrastructure
MPC	Multi-Party Computation
Tx.x	Task x.x
WP	Work Package
SMPC	Secure Multiparty Computation
SSI	Self-Sovereign Identity

Executive Summary

As part of the project, KRAKEN consortium has involved external parties to share with them the approach, methodology and results of the entire project. In this second Project Advisory Board meeting have been involved experts in different topics: from blockchain perspective to self-sovereign identity; from legal and ethical issues to business vision.

The KRAKEN second Project Advisory Board meeting aimed to show the advisors the work done in the last months of the project and to collect feedback to better address the next steps in the different work packages and to discover possible areas of improvement. The meeting took place on November 4th, 2021 and was held online., during which the advisors had the chance to get in touch with all work package leaders and contributors.

The first part of the meeting was an introduction with an overview of the project and a brief explanation of KRAKEN's structure and updates on the project status. Then we presented our business approach, in terms of market analysis and business model; we continued with a technical presentation on SSI, and Crypto aspects, and some live demonstrations. In the third part of the session the discussion focused on the ethical and legal aspects, the two project pilots (Health and Education) and a brief presentation related to the future of the project and sustainability.

The general feedback on the progress of the project was very positive and really appreciated by the advisors. Advisors agreed on consortium' choices, both on business and technical questions, and made some relevant suggestions to let the consortium better address the next choices and steps to take.

Both pilots have been deemed appropriate and contemporary. In both health and education there are many opportunities to create value by exploiting data, the crucial element that must always be considered is to have reliable data for data consumers. In health market one market segment to consider is the service to multi-center research projects. Another relevant consideration suggested by the advisors is to understand "Who benefits (most)?" as it helps to direct payment flows. Flexibility in processing these payment flows is a critical success factor. In light of the current developments in the "crypto market" it seems that crypto payments could be the future for achieving this flexibility although cryptocurrencies volatility could be an issue. Another aspect to consider is that cryptocurrencies could be too innovative. A relevant suggestion to highlight is that there are not only monetary compensations to consider in the reward system, but also transparent support for research or charity projects.

The technical approach has been considered innovative and appropriate. SSI technology seems the logical choice to integrate into the solution set. However, the visibility of blockchain data could be an issue for commercially sensible data: even if sensible data are not stored in the blockchain, the volume of transactions and the date of such transactions can be sensible and valuable information for competitors.

Finally, for the next steps of the project, the consortium should consider the new eIDAS draft, eIDAS 2.0, and the evolution of GDPR assessment by data protection European institutions.

1 Introduction

1.1 Purpose of the document

The present document has the purpose to report the Project Advisory Board feedback regarding KRAKEN project's progress and achievements. This is the second of three deliverables: D6.9 submitted in November 2020, D6.10 and D6.11 due in November 2022.

1.2 Structure of the document

The document is structured as follows: the advisors and their areas of expertise are presented in chapter 2, and the meeting agenda in section 3.1. In section 3.2 the different presentations made by the consortium and the advisors' feedback after each presentation is detailed.

Finally, in chapter 4, the main recommendations of the advisors are summarized in a table with references to the different meeting sessions.

2 The Role of Project Advisory Board

As mentioned within grant agreement, the Project Advisory Board (PAB) is an external body to the project, devoted to providing feedback and suggestions to the consortium. To achieve this goal, three official meetings with the Advisory Board have been scheduled as part of T6.5 – Project Advisory Board activities. The reports of these meetings will be submitted as deliverables and the PAB recommendations considered as a relevant input for the achievement of project goals and the technical management of the project.

2.1 Members of PAB

Andrea Migliavacca (male), degree in Business Administration (1988), 26 - year experience in ICT projects. Since 2009 Senior consultant at Lombardia Informatica (Research, Innovation and Financed Projects Ar-ea). Andrea was team leader in LISPA for Palante and Salus Projects and he currently is the CEO of Think4Future.

Carlos Pastor (male) recently joined Inetum as Blockchain Strategy Director, ESSIF convenior EBP-EBSI and collaborating with Alastria as Digital Identity Commission Leader. More than 25 years working experience in national and multinational companies like Telefónica, or Sun Microsystems linked to then emerging technologies like Intelligent Buildings, Electronic Banking, e-Commerce, Internet Gaming, Social Networks, Voice over IP, SWIFT Communication, Federated Identity, Public Key Infrastructure, Electronic Signatures (advanced including biometric voice & voice recognition signature), Self-Sovereign Identity and Blockchain.

Melek Önen (female) is an assistant professor in the Digital Security Department at EURECOM. Her current research interests are the design of security and privacy protocols for cloud computing, Big Data and IoT. She was involved in many European and national French research projects. Melek Önen holds a PhD in Computer Science from ENST (2005).

J. Peter Burgess (male) is a philosopher and political scientist. He is Professor and Director of the Chair in Geopolitics of Risk at the Ecole Normale Supérieure, Paris; Professor at the Centre for Advanced Security Theory (CAST) at the University of Copenhagen; and Research Professor at the Centre for Law, Science, Technology and Society Studies (LSTS) of the Vrije Universiteit Brussel. His research and writing have focused mainly on the theory and ethics of security and insecurity, and more recently on questions of fundamental rights in relation to digitization, data protection and privacy. He is at present Chairman of the Ethics Advisory Group of the European Data Protection Supervisor and co-authored its recent report Toward a Digital Ethics.

Mr Harald Zwingelberg (male) is head of the “Privacy Technology Projects” division at Unabhängiges Landeszentrum für Datenschutz (ULD), the office of the Data Protection Authority of Schleswig-Holstein. On behalf of ULD he participated in a series of EU-funded and national research projects with relation to data protection, privacy and identity management. His focus resides with legal aspects of data protection.

Dr. André Kudra (male) has more than 13 years of information security consulting experience. In his career he held various key positions in major information security projects of global enterprise organizations. He studied business administration at the European Business School (EBS) in Oestrich-

Winkel, Germany, and computer science at the James Madison University (JMU) in Harrisonburg, Virginia, USA. Since 2013 André is CIO of esatus AG, a consulting company specialized in information security matters, with its headquarter near Frankfurt in the Rhine-Main area and offices in Hamburg and Munich. André is a strong advocate of Self-Sovereign Identity and a Sovrin Technical Governance Board member.

3 Second Project Advisory Board Meeting

3.1 Meeting organization, agenda and participants

Due to Covid-19 constraint, the second Project Advisory Board meeting has been organized as online meeting. The meeting took place on Thursday, November 4th, from 9.00 am to 01.45 pm and the platform was Teams, by Microsoft. The meeting has been recorded with the consent of all participant.

The agenda had the aim to show to the advisors the project activities as a "Company Presentation" addressing the business point of view first, and then going in depth into technical, legal and ethical aspects.

The agenda (Table 1) covered all the Work Packages of the project and to make the meeting as interactive as possible all the sessions were followed by a feedback discussion with advisors.

4 th November 2021			
Time	**Description**	**Responsible**	**Duration**
9:00-9:05	**Conference opening – partners & advisors join the conference**		5'
9:05–9:15	Welcome, presentation of the agenda and meeting objectives. - Roundtable to present advisors and partners	**INFOCERT**	10'
09:15-09:30	**Project Overview: update on current status** • Project Overview & organization • 6 objectives • Project updates vs 1st Advisory Board	**ATOS**	15'
09:30-09.40	**Market Analysis Highlights**	**INFOCERT**	10'
09.40-10.25	**Business Model (Education & Health):** • Business cases for healthcare (hospital networks, Consumer data) • Business cases for education • Value proposition, market segments and data requirements • Payment systems (FIAT vs Crypto payments) • Data valorization and pricing • Computation as a service	**TEX** **TUG**	45'
10.25-10.40	**Round table for Feedback Session on Market Analysis & Business Model**	**TEX**	15'
10.40-10.55	**Break**		15'
10:55-11.15	**Technical Approach. Platform Overview:** • SSI Components • Crypto aspects	**ATOS** **TUG**	20'
11.15-11.35	**DEMOS**	**TEX LYN** **ATOS** **INFOCERT**	20'
11.35-11.50	**Round table for Feedback Session on Technical Approach + DEMOS**	**INFOCERT**	15'
11.50-12.05	**Break**		15'

4th November 2021			
Time	**Description**	**Responsible**	**Duration**
12.05-12.20	**Ethics Advisory Board Meeting**	**KUL**	15′
12.20-12.35	**Round table for Feedback Session on Ethical and Legal**	**KUL**	15′
12.35-12.45	**Pilot focus (WP5) Education: Initial marketplace use cases**	**TUG GRAZ**	10′
12.45-13.05	**Pilot focus (WP5) Healthcare: Initial marketplace use cases**	**LYN**	20′
13.05-13.15	**Future of the Project and sustainability**	**ATOS**	10′
13.15-13.30	**Final feedback from Advisors**	**INFOCERT**	15′
13.30	**End of the meeting**		

Table 1: PAB meeting agenda

Five out of the six project advisors participated in the meeting. One of them couldn't attend. The meeting presentations and a questionnaire were distributed to all of them.

Work Package leaders and representatives of all project partners attended the meeting.

3.2 Presentations and live feedback

3.2.1 Market analysis highlights [1]

Identities need to be portable, verifiable but also private and secure. Digital identity is the solution to reduce the level of bureaucracy and increases the speed of processes within organizations by allowing for a greater interoperability between institutions. Digital Identity satisfy the need of a portable and verified identity within a standard format, but at the same time is essential that digital identity can't be stored on a centralized server, otherwise it will become a honeypot for hackers.

Marketplaces solved the Identity issues, ensuring:

- a connected ecosystem where data is shared and aggregated, breaking the paradigm of data collected in silos
- a secure data transmission creating a safe infrastructure by using cryptography, anonymizing the data
- giving a clear responsibility to Data providers.

Thanks to SSI based on Blockchain technology with a decentralized approach the problem of individual control, security, and portability of Digital Identity can be solved.

Health & Education data ecosystem are facing different important issues, such as:

- Data are widespread and stored in silos with a lack of interoperability
- Lack of safety and security of data
- Lack of Data Ownership and control
- Paper-based certifications and documents

To answer these issues for both industries, it is necessary to create an ecosystem where different actors can easily exchange data, breaking down data silos, enhancing the security and privacy

[1] Presentation about market analysis showed the results of D6.2 "Initial market analysis"

management and giving a clear ownership of data. After analyzing 18 companies from all over the world, with a B2B and B2C approach and different type of approach, the main value propositions emerged are the following:

- Empowering the Data Providers as data controllers ensuring a full control over their data
- Provide tools to share and analyze and collect data in a simple, fast, interoperable, and certified way
- Monetize data through a clear reward system between data consumer and data provider.

There are 2 main monetization approaches:

- “Pay per use”: Service providers reward Data Provider once he gives consent to share data and Marketplace generally takes a transaction fee
- “Subscription fee”: Service providers pays a fee to marketplace for their services

Generally, B2B monetization approach is based on a platform fee, cut of sales and subscription for added value services (such as analytics). Marketplace are governed by Data exchange agreements that allows participants to share and access to information according to a reciprocal agreement.

Live feedback from advisors:

No additional feedback.

3.2.2 Business Model Education & Health[2]

Education

Work in the first half of the project and last advisory board meeting was focused on the direct exchange and sharing of education data. In the following months we will focus on the business case and value propositions for two other data sharing modalities:

- Privacy-preserving analytics using SMPC for simple education data
- Potential new Use Case: Privacy-preserving computation of population statistics

The Education Data Market supports the secure sharing of academic data of students, such as graduation certificates, certificates for courses, and the enrollment status for individual terms. The KRAKEN system thereby enabling data consumers to perform analytics on that data in a privacy-preserving way.

Besides the other relevant user groups (students, and human resources agencies/HR departments), the following period will focus on the user group of Statistical Agencies (private and public).

In KRAKEN, statistical agencies can acquire academic data offered by students and use them in computations of statistics and other analysis in a privacy preserving way.

Additionally, those agencies are enabled to combine data sets *from multiple sources* (e.g., universities and governmental statistic agency) and compute statistics on that data while preserving student's privacy, which demonstrates an additional use case for the platform.

Live feedback from advisors:

Remark about the potential complexity of the “new” use case, which could be a usability risk. Dismissed after realization that the involved parties (large, specialized entities on both ends) have the expertise necessary to deal with the processes.

[2] Presentation regarding business model showed the first results of D6.4 “Initial Exploitation Plan”

Health

This section of the Advisory Board meeting covered the business model in terms of the business cases and value propositions associated with the health pilot whilst also examining the problem of data pricing for the KRAKEN marketplace's users and the possible business and revenue models for the KRAKEN marketplace.

The focus of the presentation was on two different data sharing modalities in the marketplace: privacy-preserving analytics using Secure Multi Party Computation for hospital networks, and consumer data streams using Data Unions.

The general business case presented behind the first was that we know that hospitals are sitting on highly desirable and highly-valuable data assets but that the hospital data ecosystem is extremely fragmented and siloed. The two main reasons for this are that hospital IT infrastructures are not well integrated, and hospitals are distrustful or fearful of sharing their data because of the legal and ethical liabilities.

The KRAKEN value proposition for hospital networks presented was that KRAKEN provides hospitals with a way to capture the entire value of their data whilst processing their sensitive data in a secure and private way at scale by allowing them to perform multi-party data analytics in private environments in the marketplace.

It was discussed that the combination of SMPC and blockchain prevents the exposure of patient information by leaving the shared data behind individual hospitals' firewalls which shields them from any ethical and legal risks or liabilities.

Two possible use case or application patterns for the SMPC were presented: 1) A third party organization like a pharmaceutical company, uses the marketplace to run analytics on data from one or more hospitals, without these hospitals providing full access to their data. 2) And multiple hospitals or hospital consortiums pool their data together, making it available for them to perform joint analytics without having access to each other's underlying data.

For the second modality of consumer data streams using Data Unions, the business case presented was that Healthcare organizations are increasingly looking to get their hands on consumer and lifestyle data and that by 2022 more than 1bn people will be connected to wearable devices.

If healthcare organizations can access consumer data, they can use it to forecast health risks and identify health outcomes, for population health management, studying and developing personalized treatments, post-market drug surveillance amongst many other use cases.

But the current situation is that much of this consumer data is siloed and controlled by data monopolies in the form of large organizations, which limits innovation in healthcare and the ability to derive insights. Consumers also lack control over the data they create, and they do not share in the value created.

The Value Proposition presented was that through KRAKEN and Data Unions, citizens and patients can gain greater control over the data they produce, share in the value generated from their data, and enhance their competitiveness in today's data markets through increasing their collective bargaining power.

After discussing the business case and Value Propositions associated with the health pilot the presentation discussed one key challenge for users of the KRAKEN platform, which was how do they work out a realistic price for their data?

It was discussed that with the technology used in KRAKEN, data owners can collect and track vital information to aid them in defining a realistic price for their data. For example when using SMPC in privacy-preserving analytics, data owners have to make available the type and quantity of data they want to share, and they also define the purposes of use or use cases that their data can be used for. They also know the types of computations that would be required for these use cases, which could include basic descriptive statistics right the way up to training of artificial intelligence agents. The

KRAKEN marketplace requires data buyers to identify their purposes of use and the computations they want to perform on data before gaining access to it.

When the SMPC system and the marketplace is combined with the blockchain, it creates an immutable record of all the necessary information a data owner might need to base their pricing decisions on, including: The type of data, the quantity of records in the data, the number of variables in each of the records, the types of computations that are being performed on the data assets, and the types of use cases or purposes of use.

With all of this information tracked, KRAKEN provides a sufficient base for data owners to define and evolve a realistic pricing model.

Finally, the possible business and revenue models for the marketplace were discussed. The first of these was a marketplace-as-a-platform model where the marketplace acts as an open platform that can be used by anyone to share either direct access to their data or share their data for analytics. In this model the marketplace would either receive revenues in the form of transaction fees for each data transaction or subscription fees to access the marketplace. In the case of analytics, the marketplace could also sell use-case specific, pre-built queries, which are sold as premium apps within the marketplace to be used on top of the SMPC Network.

The second business model discussed for the marketplace was more aligned to private hospital networks, and a “computation as a service” model. In this model SMPC network nodes would be hosted locally in hospitals that wish to connect their local data assets and run analytical queries or AI models in private marketplace style frontend environments where only pre-determined partners have access to the data marketplace. KRAKEN would sell licenses to use the software for a specified period of time, and it would again be possible to build further use case specific queries that are programmed on request and sold to the hospitals.

Live feedback from advisors:

During the discussion we asked the following questions to the advisors: How do you currently motivate patients to participate in your app? Are you thinking about any possible incentives such as money or rewards? Do you think your users would be interested in receiving crypto for sharing their data?

Answer from advisors: “The first motivation for patients to use the app and share data is to get feedback, some results some help in their activity, in the management of their pathology and to better manage their therapy. Basically, to get a direct benefit in the interaction with the hospital or to better manage their therapy. The second motivation, collaborating with the patient organization, is to share information in the context of the association to better achieve the association goals. So, the first motivation is not monetary. There might be a possibility that if we used some monetary incentive it might be interesting and might encourage people to be more active. For some kinds of people, especially those involved in associations, there might be a monetary incentive not directly addressed to them but to the association or for instance to some specific projects/groups where there might be a real incentive. Basically, people would be willing to sell access to their data and then in return donate the money generated to these organizations that will invest in research projects.

In my experience with healthcare providers, the first motivation is to better manage the relationship with my healthcare provider and secondly, working with patient associations, the incentive is not directly related to monetization but is more related to supporting a project by creating a collaborative situation. Some patient associations are moving to get data and to involve patient members in this process. There is a need to have some actors who can guarantee the system to support and understand the value of this activity.”

During the discussion the following comments emerged from advisors: “what about the nature and number of data consumers? There is the need to consider the business that will be using KRAKEN, who are they and how many are they? This might have an impact on the value and consequently on prices”.

"Need to consider data quality, whether data are complete and obtained directly from the medical device by the user or by other specific methods.

"With my experience with Hospitals Computation as Service in a specific large research project where are involved multicentric research project could be a good way to start a collaboration, where large multicentric activities maybe could be a part of that project and to be a normal asset/infrastructure for managing that kind of project. For activity where there is not specific large cooperation maybe monetization of queries could be an approach for small research point or where the researcher normally works only to the multicentric large project."

3.2.3 Technical Approach: SSI components, Crypto aspects

SSI components

During this session, the main achievements of WP3 were presented. In the first part of the presentation a brief introduction to the Self-Sovereign Identity approach was presented. Then, a very high-level detail of the main use cases developed by the WP3 was shown, including the health pilot but also the Education pilot. Then an architecture diagram with the main blocks developed was presented, highlighting the main components necessary for the implementation of the use cases commented on previously. Finally, a demonstration of the different components, integrated within the final environment, was presented to show how the different components interact with each other live.

Live feedback from advisors:

The advisors were pleased with the main achievements performed by the WP3 team and encouraged the consortium to keep working this way.

The advisors also commented the necessity of taking into account the new eIDAS draft, eIDAS2. Also commented, the evolution of GDPR assessment from Data Protection European institutions. On the other hand, the use the selective disclosure, privacy-preserving revocation mechanisms and trust registries for recording eligible issuers are worth considering and incorporating. Hence the taken approach is worthwhile.

CRYPTO technologies and components

With respect to KRAKEN's Crypto aspects (mainly WP4) we first presented our approaches and results on **(1) privacy- and authenticity-preserving data analysis**, **(2) designated secure end-to-end data sharing**, and **(3) privacy-preserving aspects of Self-Sovereign Identity (SSI) systems**. Next up, we asked the advisors **specific**, as well as **open**, **questions** for input and feedback.

For **(1) privacy- and authenticity-preserving data analysis** we presented first, the overall goals as well as our chosen main technologies; and second, our general Crypto architecture.

For **(2) designated secure end-to-end data sharing**, we presented first, the overall goals as well as our chosen main technologies; and second, an example data-sharing flow within our Crypto architecture.

For **(3) privacy-preserving aspects of SSI systems**, we presented first, the general research goal as well as our used main technologies; and second, the general flow of an enhanced privacy-preserving showing leveraging an SSI system which is the output of one of our research papers.

After stating our questions, the **input and feedback**, was mainly focused on (1) the deployment of an MPC node by ("bigger") users, such as hospitals, and (2) the handling of the change of MPC nodes within KRAKEN.

Live feedback from advisors:

Advisors gave us valuable feedback in an abstract way on our general Crypto architecture. Specifically, they mainly mentioned how we deal with the scenario, when the MPC nodes in KRAKEN change; especially because our first deployment scenario considers (only) three MPC nodes. For instance, if two or more MPC nodes join the system; this is a very important question. Because in this case, already existing Data Providers in KRAKEN, need to create new encrypted shares for the joined MPC nodes. Furthermore, the Data Providers would need to decide how many shares of their data they want to create, and then, onto which MPC nodes they want to split them. As this can be a difficult task (how many (threshold) shares and which nodes), this would require a user action from Data Providers which we – as KRAKEN – might not want. Moreover, the already chosen MPC nodes, before the new ones joined, might have already accessed their share; which reveals the task of how to guarantee that these nodes indeed delete their respective share – given that they have had it already once. Since this is a very interesting and important follow-up task, with points we have maybe not even realized yet, we agreed on having a follow-up meeting, with respect to this topic. Although, this task is likely only practically relevant when the KRAKEN project is finished – as we first focus on the deployment scenario of exactly three MPC nodes, which are hosted by three different KRAKEN partners (Atos (Spain), TX (Finland), and XLAB (Slovenia)).

Moreover, there was a broader discussion about the case when a (“bigger”) user of the KRAKEN system, such as a hospital, wants to significantly increase their privacy guarantees for data analysis by providing an MPC node on their own. The discussion included almost all advisors. The main input and following outcome of the discussion was, that it probably makes the most sense to first talk to health-related organizations which have enough “tech-savyness” to really self-deploy an MPC node. For instance, for a doctor in a hospital, it might not be feasible to self-deploy an MPC node. While the discussion was focused on these “bigger” users, we want to note here that, theoretically every KRAKEN user could provide an MPC node. Although, the case of this every-user MPC-node self-deployment, is currently not really practical; on a technical level, mainly due to the lack of performance of MPC on the phone.

3.2.4 Demonstrations

Self-Sovereign Identity functionality demo

As part of the WP3’s presentation a demonstration of the main Self-Sovereign functionality was presented. The demo shows the case where the end user connects to the Marketplace and performs the following steps:

- Click on login button showed in the Marketplace website.
- The Marketplace website shows a QR Code with the required information for performing the authentication on it.
- The end user reads the QR Code using the mobile.
- As result the end user receives an invitation from the marketplace on the mobile, as it is the first time connecting to this service, the end user accepts the invitation.
- The Marketplace accepts the connection from the mobile of the end user in a transparent way to the end user.
- The mobile of the end user shows that the connection has been established
- Once the connection is established, the Marketplace shows a form to the end user on the website to gather the information necessary for the registration of the end user to use this service.
- The end user submits the form using the marketplace website.
- This process triggers an issue credential process internally, and then, the marketplace sends a credential offer to the end user

- The end user receives then an offer on his mobile.
- The end user accepts the offer on the mobile.
- The Marketplace, in a transparent way to the end user, sends the credential to the end user
- The end user sees on the mobile that the credential has been received and stores it into the mobile with his choice name/nickname for this credential.
- With this step, the registration process is finished.

For successive login operations, the end user just needs to read the QR Code presented by the Marketplace using the mobile and the system will send the information required for a successful authentication.

Legal Identity Manager (LIM) demo

LIM is designed to implement two functionalities:

- Issuing to a user's wallet of an electronic identity (eID) Verifiable Credential produced from an EIDAS identity assertion, assertion produced by an EIDAS identification phase.
- Signing (PADES) pdf documents with the identity contained in the eID VC issued by the LIM using the remote signature API provided by Infocert. Not showed in the demo because not yet implemented.

Two tools were used in the demo, the KWCT (KRAKEN Web Company Tool), to simulate a user's SSI wallet, and the LIM.

The LIM acts as a bridge between two distinct worlds, EIDAS and SSI, covering both an EIDAS service provider role and an SSI issuer role.

EIDAS provides to the LIM the trust/legal framework during the identification phase, the SSI provides to the LIM by design non-tamperable Verifiable Credentials and privacy and secure communication with the user.

LIM uses the services provided by the Italian eIDAS AGID Proxy node to implement the EIDAS authentication phase.

Demo steps:

The LIM, after that the user connected his SSI wallet to the LIM using standard DIDcomm's DID connection protocol, redirected the user to the eIDAS_AGID_Proxy website. On the AGID node web site, the user selected a country and an identity provider of that country, after that, he authenticated himself (Infocert was the identity provider chosen in the demo) using his real EIDAS identity. After the authentication the LIM, using DIDComm protocols, sent an eID Verifiable credential offer to the user's SSI agent. The user, using the webUI of the KWCT, accepted the credential offer. The LIM automatically issued the eID_VC that, after the user acceptance in the KWCT UI, was saved inside the user's SSI agent.

Marketplace demo

The marketplace is the portal that the users use to publish and purchase access to Data Products. The functionalities it provides include:

- Registration
- Login
- Browsing of Data Products
- Data Products publication for both the health and the education pilot
- Data Products purchase

The marketplace allows users to publish on the marketplace three kinds of data products.

Batch Data Products allow data consumers to access entire datasets. This data sharing modality is enabled also by the integration of the marketplace with the SMPC technology.

Analytics Data Products allows data consumers to access analytics computations performed on datasets. This data sharing modality as well is enabled also by the integration of the marketplace with the SMPC technology. This feature was not shown in the demo as its development was still ongoing.

Real time stream Data Product allows data consumers to access streams of real time data. This data sharing modality is enabled also by the integration of the marketplace with the Streamr network technology. This feature was not shown in the demo as its development was still ongoing.

Moreover, the marketplace performs registration and login of users by integrating with the SSI technology.

Demo steps:

The demo provided by TX started by showing the login in the marketplace, performed by using both an SSI wallet and a Metamask wallet. The login was followed by the publication of a Batch Data Product. The entire process of publication consisted of the following steps:

- Choice of Data Product sharing modality,
- Metadata provision in the edit product page,
- Market sector choice and tags selection,
- Collection of user's statement of not sharing data of persons other than himself or, in case he does, that he collected the consent of these third persons,
- Purposes selection for the eligibility check of data consumers,
- Encryption and cloud storage of a dataset,
- Price selection,
- Publication on the catalogue

The publication was followed by the purchase, consisting of the following steps:

- Browsing of the catalogue,
- Navigation to the Data Product page showing metadata and policies
- Navigation to the Data Product purchase page showing a subset of the metadata, plus a widget for the selection of the purposes and GDPR disclaimer
- Purpose's selection,
- GDPR disclaimer check,
- Eligibility check of the data consumer,
- Payment of the data product,
- Download and decryption of the dataset.

Live feedback from advisors:

The feedback from the advisors highlighted positive features of the marketplace together with some areas of improvement. In particular, the selection by the data provider of the kind of buyers that can be considered eligible to buy a Data Product on the marketplace during the data publication workflow had positive feedback. A similar positive reaction was provided on the part of the Data Product publication workflow where a user must declare that they have obtained consent from data subjects whose data is included in the published Data Product. They did add however, that this might not be enough to prevent the marketplace from being a data processor.

The advisors also suggested a different way of selecting the countries where a data provider decides they would like to share their Data Product. Currently the marketplace gives the user two choices: Countries that are part of the KRAKEN platform (Austria, Belgium, Denmark, Estonia, Finland, France, Germany, Italy, Netherlands, Portugal, Spain, Sweden, UK) or all countries in the world. The suggestion was to give the user three choices: GDPR countries (countries of the European Union), third countries with an adequacy decision and the other countries around the world without an adequacy decision.

3.2.5 Ethic and legal aspect

In terms of progress, most of the legal and ethical deliverables have been submitted. The only deliverable still to be submitted is D7.3 'Ethical and legal evaluation and recommendations', due in November 2022.

The exist several legal and ethical concerns related to the KRAKEN research activities, which are the following:

- the involvement of human participants in the pilots: for this concern the KRAKEN consortium has implemented informed consent procedures and specific partners have obtained approvals by relevant ethics bodies for the involvement of human participants in the research activities (relevant deliverables are D8.1 H - Requirement No. 2 and D8.5 H - Requirement No. 7).
- the processing of personal data: this challenge is addressed by the development of an internal KRAKEN data use policy, the implementation of informed consent procedures, the adoption of a data protection and security by design approach, assessing the need to conduct a DPIA for the research activities, and measures around the further processing of personal data (relevant deliverables are D7.1 Ethical and Legal Management, D8.1, D8.2 POPD - Requirement No. 4, D8.3 POPD - Requirement No. 5, D8.6 POPD - Requirement No. 8).
- the potential use of sensitive personal data in the health pilot: here KRAKEN has obtained recommendations and opinions from the internal Ethics committee and has made use of dummy data in the pilot demonstrations instead of real personal data (relevant deliverables are D8.4 GEN - Requirement No.6).
- the immutability of the blockchain: this concern is addressed by not storing any personal data on the blockchain used in KRAKEN (relevant works in WP2 & 7, D7.2 Ethical and Legal Requirement Specification).
- residual challenges: any residual challenges are addressed through continuous and ad-hoc interactions between partners as well as a risk management process.

There are also legal and ethical concerns related to the KRAKEN system as a platform:

- anonymous data; specifically the question whether personal data is properly anonymized in the light of the requirements of the GDPR: our approach is to provide the possibility to share already anonymized data and also offer the possibility to use third party anonymization tools.
- valid consent:
    - what about the validity of consent when personal data is not published by the data subject directly, but rather by a controller;
    - is consent valid when specific processing activities or purposes are not yet known at the time of consent, for example the transfer of data to specific third countries;
    - in the case of withdrawal of consent, how to ensure a fair repayment of the price to the data consumer, for example through a smart contract.
- the status of responsibilities of actors, such as the status of Data Unions: here we developed a taxonomy of actors and their status in different scenarios;
- data monetization, specifically the unclear legal framework on the monetization of personal data and the need for additional guidelines: here we want to keep the system flexible to adjust to different possibilities, but in principle the monetization of personal data is possible if the relevant legal & ethical requirements are met;
- is a Data Protection Impact Assessment (DPIA) needed for the KRAKEN system: a full DPIA is not feasible due to the fact that many aspects and details depend on the real-life implementation of the system. Instead we opt for a lightweight research DPIA that takes into account the information available to us during the research phase.

Some of the next steps in the ethical and legal activities are:

- monitoring and researching policy & legal developments; such as the upcoming EDPB opinion on monetization of personal data in the EU as well as upcoming legislation (e.g. the Data Governance Act).
- D7.3 'Ethical and legal Evaluation and Recommendations in M36: this deliverable will evaluate and validate the implementation of legal and ethical requirements, formulate policy recommendations, contain a lightweight research DPIA, and analyze the EU data monetization approach and role of the Data Governance Act.

Live feedback from advisors:

Some of the ethical and legal concerns were discussed with advisors, who gave the following feedback:

- Blockchain: no personal data is stored on the blockchain but be mindful that metadata and entity names are not personal data in themselves. Keep metadata and entity names broad (e.g., data subject, controller, etc). He also advises to consult the CNIL guidelines on blockchain.
- Anonymous data: there are new EDPB guidelines on anonymization/pseudonymization in the pipeline.
- Validity of consent: consent should indeed be as specific as possible. In the future we want to move towards an automatic consent flow without continuous feedback loops and consent exhaustion. Also check the proposals for the Data Governance Act, Digital Markets Act, and Digital Services, which could affect the current landscape in terms of actors and how we might deal with consent.
- Transfer of personal data to third countries: reformulate the different categories: EU countries (subject to the GDPR), third countries outside of the EU with an adequacy decision, and the rest of the world.
- Lightweight DPIA: it makes sense to conduct a lightweight DPIA instead of a full-scale DPIA.
- Right to be forgotten: be mindful to remove all personal data, including encrypted personal data.

3.2.6 Education pilot

The KRAKEN education pilot if structured in three use cases:

- A Student receives an education Credentials from their University
  → This use case demonstrates KRAKEN's focus on the principle of sovereignty.

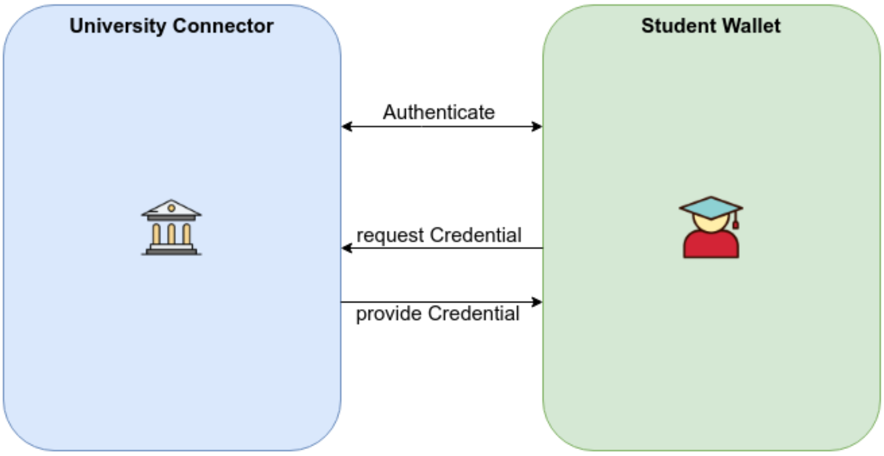


Figure 1 – University connection flow

- A Student provides their Credentials to other University or a HR Agency
  → This demonstrates the verification of the authenticity of the credential. Additionally, selective disclosure of attributes is possible to enhance the students' privacy.

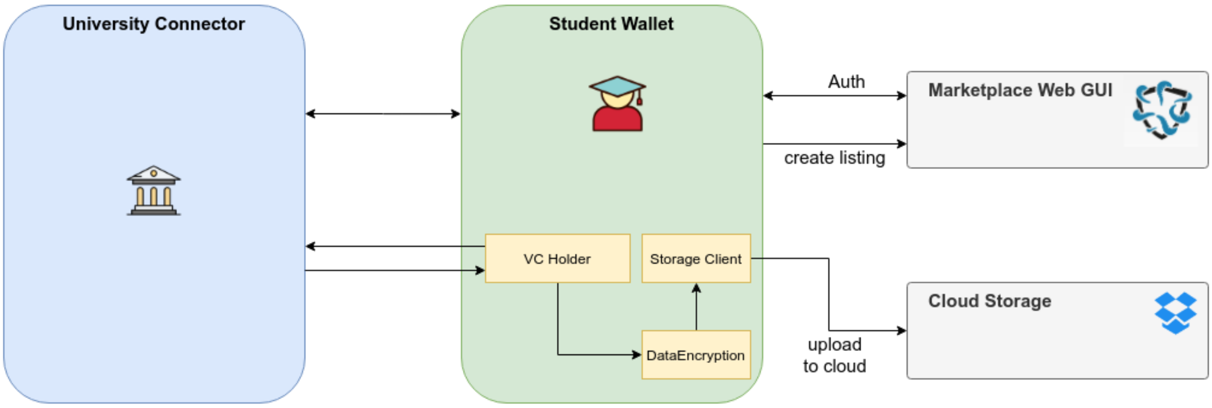


Figure 2 - Data flow between University and students

- A Student submits encrypted Credentials to Marketplace.
  → A statistical agency acquires one or multiple data sets and performs privacy-preserving analytics. This demonstrates KRAKEN's advanced cryptography.

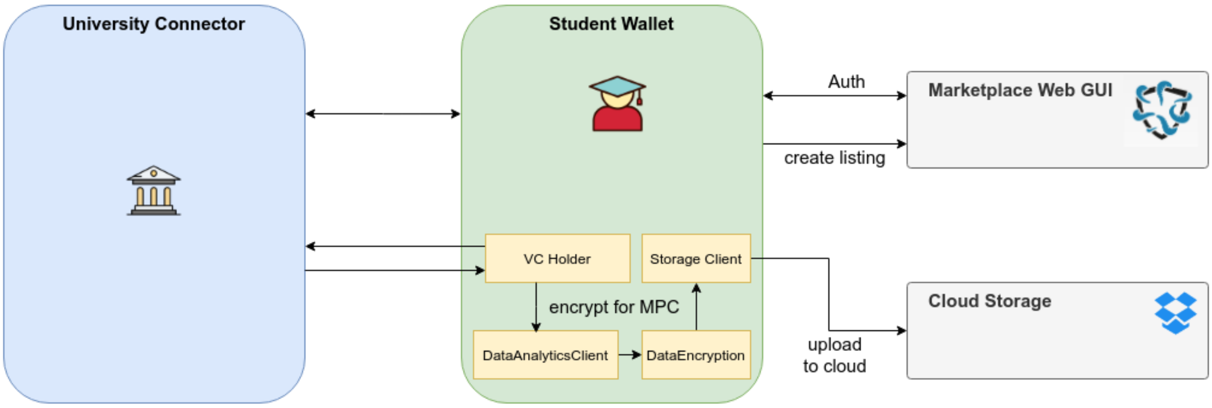


Figure 3 - Data flow between University and Students Wallet

Additionally, the third use case can be modified to combine data from multiple sources into one computation, demonstrating the flexibility of the KRAKEN system and the power of the integrated cryptographic technologies.

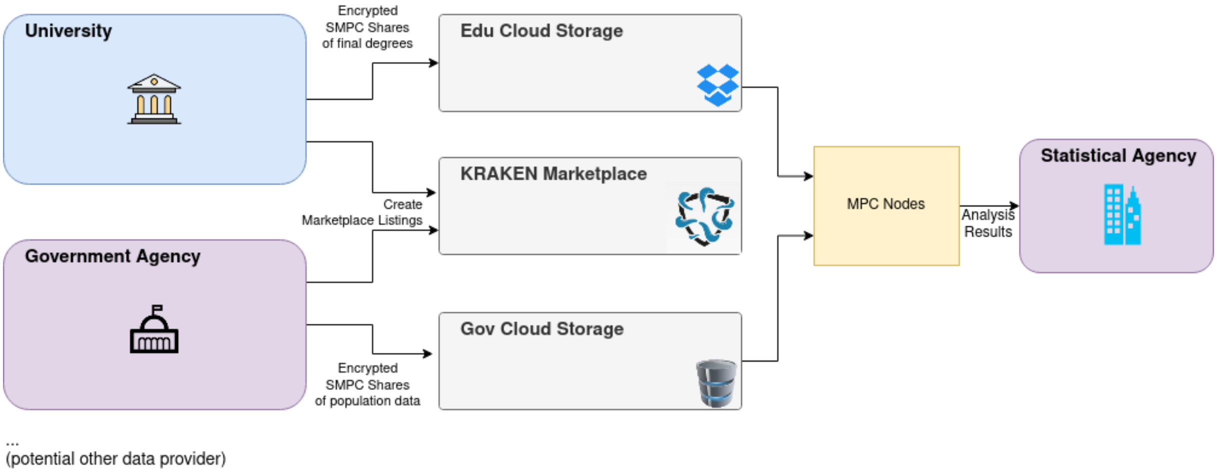

University
Encrypted SMPC Shares of final degrees
Edu Cloud Storage
Create Marketplace Listings
KRAKEN Marketplace
Government Agency
Encrypted SMPC Shares of population data
Gov Cloud Storage
MPC Nodes
Analysis Results
Statistical Agency
...
(potential other data provider)

Figure 4 - Combining Education data from multiple sources

Live Demo: The current status of the implementation of the KRAKEN education pilot, in specific the university connector and the integration with the KRAKEN mobile wallet, was also demonstrated to the advisors during the meeting.

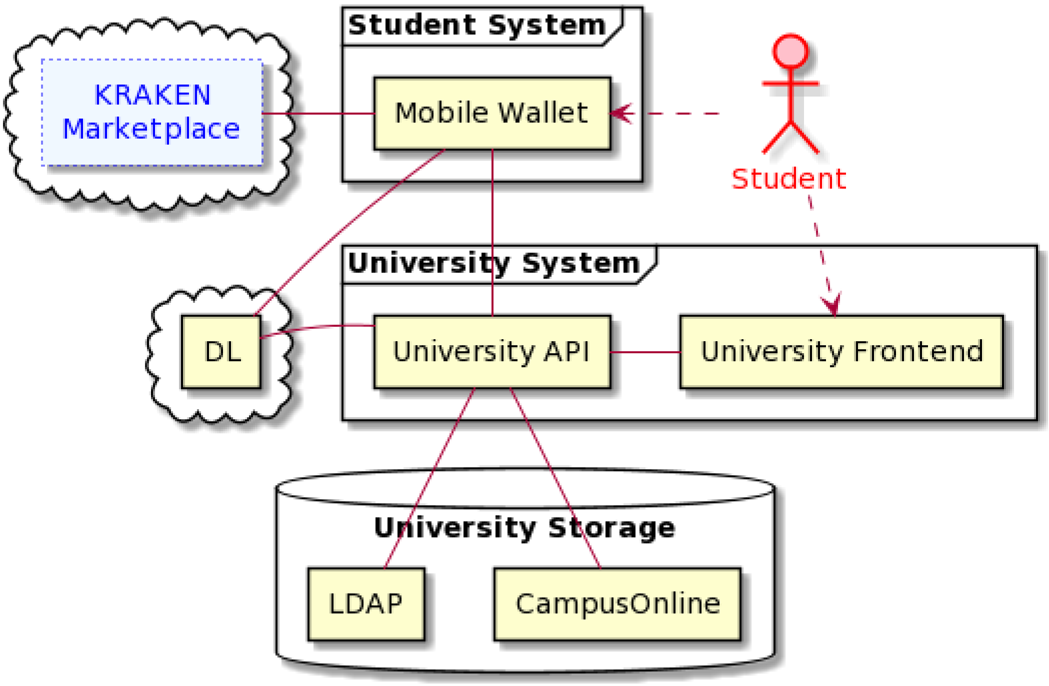


Figure 5 - Educational Pilot flow

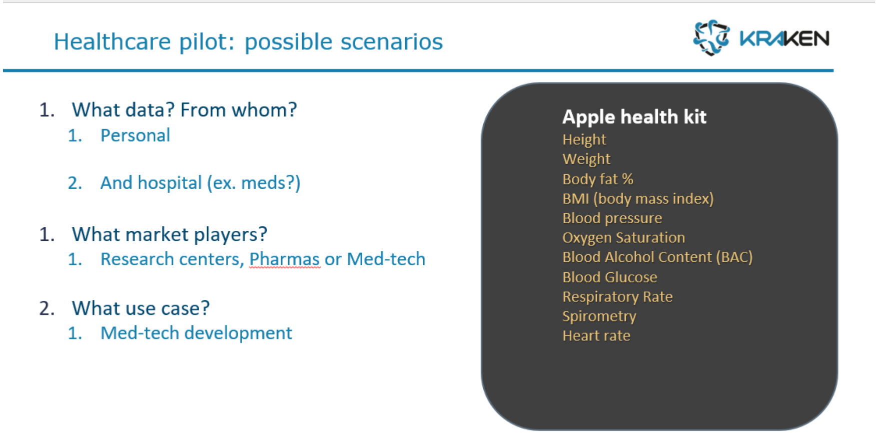
Live feedback from advisors:

No additional feedback.

3.2.7 Health pilot

The health pilot presentation focused on the plan for joining the platform and testing some of its key functionalities in the second period of the project, specifically discussing the details of use cases, goals and focus, such as possible data products, for what type of buyers. In particular Andrea Migliavacca was engaged, as the executive of a clinical data exchange mobile application, in this part of the discussion.

The possibility of engaging with the Apple Health Kit ecosystem through players active in this space was presented to the advisors and positively received.



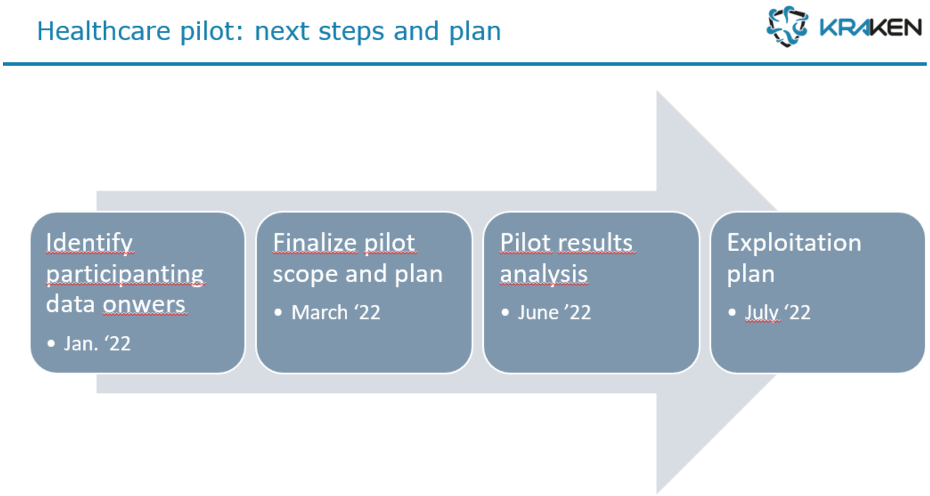
Healthcare pilot: possible scenarios

1. What data? From whom?
   1. Personal
   2. And hospital (ex. meds?)
1. What market players?
   1. Research centers, Pharmas or Med-tech
2. What use case?
   1. Med-tech development

Apple health kit
Height
Weight
Body fat %
BMI (body mass index)
Blood pressure
Oxygen Saturation
Blood Alcohol Content (BAC)
Blood Glucose
Respiratory Rate
Spirometry
Heart rate

Figure 6 - Health pilot: possible scenarios

In conclusion of preliminary plan and timeline for the pilot deployment was presented eliciting interest in participation.



Healthcare pilot: next steps and plan

Identify participanting data onwers
- Jan. '22

Finalize pilot scope and plan
- March '22

Pilot results analysis
- June '22

Exploitation plan
- July '22

Figure 7 - Health pilot: next steps and plan

<u>Live feedback from advisors:</u>

Advisors highlighted that the most relevant use cases are those, in the area of real-world medicine in which patient data can we provide it to pharmaceutical and medical device companies to gather important feedback on medications side effects, efficacy another important clinical information that is not available from randomized trials.

3.2.8 Future of the project and sustainability

The consortium showed to the Advisory Board the two sustainability aspects considered as key for any successful sustainability:

- **Acceptance** (from end users)
  - Stakeholder validation
  - New legal structure
  - Exploitation agreement

The consortium presented a draft version of an exploitation agreement as the most pragmatical approach to a joint exploitation.

This agreement will tackle any potential commercial opportunity that may pop up beyond the project lifespan.

- **Financial** (cost/benefits effective)
  - Pricing structure & compensation schemes

The consortium presented a compensation scheme model as well

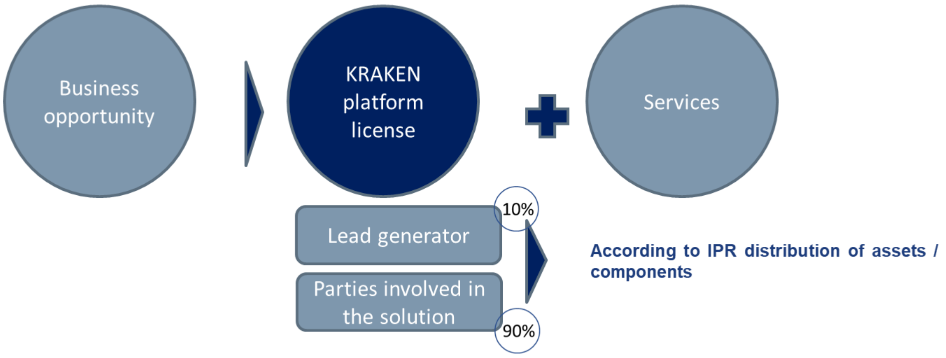
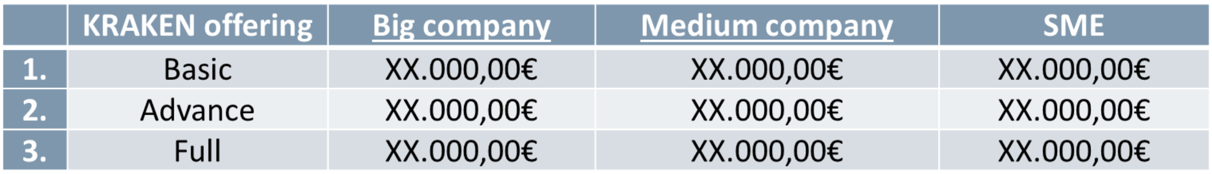

Business opportunity
KRAKEN platform license
Services
Lead generator
Parties involved in the solution
10%
90%
According to IPR distribution of assets / components

Figure 8 - Compensation scheme model

And the option of a modular approach related the platform functionalities and trying to accommodate them to the customer needs.

This different modular offer will have a differentiated price which may result in wider acceptance from a financial point of view from end users.



	KRAKEN offering	Big company	Medium company	SME
1.	Basic	XX.000,00€	XX.000,00€	XX.000,00€
2.	Advance	XX.000,00€	XX.000,00€	XX.000,00€
3.	Full	XX.000,00€	XX.000,00€	XX.000,00€

Figure 9 - Example of modular offer

Live feedback from advisors:

No relevant feedback received

4 Advisors' recommendations

This chapter summarizes, in the following table, the main advisors' recommendations with references to the meeting sessions where each subject was discussed.

Section	Comments/Recommendations from Advisors
Market Analysis Highlights	• Potentially in a future round, the open-source works coming out of related/relevant projects and organizations can be leveraged. Examples are: The polypoly cooperative (https://polypoly.coop) creates a polyPod which allows users to manage with whom they share data in a convenient, empowering way. The Human Colossus Foundation (https://humancolossus.foundation) works on a Dynamic Data Sharing Hub with Consent Flow as part of their Dynamic Data Economy program. • An important role can be played by patient associations which can be an important guarantee vehicle for non-professional users.
Business Model (Education & Health)	• Both use cases are highly appropriate and contemporary: In health, there are large-scale opportunities for leveraging patient data from various sources. In education, having reliable data about candidates is vital for employers. Key consideration should always be "Who benefits (most)?" as this usually helps to direct payment flows. Flexibility in processing these is a critical success factor. Considering the current developments in the "crypto market" it seems inevitable that crypto payments are the future for achieving this flexibility. However, regulatory obstacles are still not resolved despite the crypto market cap having touched 3 trillion USD in Nov 2021. As for the "computation as a service" aspect, I assume security-focused, highly decentralized cloud infrastructures in the sense of "confidential cloud computing" will become available soon. It may not be relevant to host own computing hardware for security reasons. (In such a scenario, one or more micro datacenters may be placed in a hospital which are added to a confidential computing mesh cloud.) • Analyze the future legal alternatives for Kraken, consortium, Joint Venture. • Crypto volatility could be an issue. • Pricing should be based on real market data or at least on market analysis of real potential data buyers. • Health: one segment to consider is the service to multi-center research projects. • Cryptocurrencies: perhaps too innovative. it is appropriate to take into account not only monetary compensation systems but, for example, transparent support for research projects or charities • Price: price transparency is essential. it is a complex issue to define as it intersects with the quality of the data
Technical Approach. Platform Overview	• SSI technology seems the logical choice to integrate into the solution set, as confirmed by the worldwide momentum SSI has achieved (Sovrin network statistics show increasing amount of ledger writes). Particularly selective disclosure, privacy-preserving revocation mechanisms and trust registries for recording eligible issuers are worth considering and incorporating. Hence the taken approach is worthwhile. • The technical approach is innovative and appropriate. One could consider discussing how to add/remove one MPC node in a dynamic manner, but this can be considered as a topic for yet another project. • Visibility of blockchain data can be an issue for commercially sensible data, i.e: even if information of on the queries is not stored in the

Section	Comments/Recommendations from Advisors
	blockchain, the volume of transactions and the date of such transactions can be sensible and valuable information for competitors - I think it is adequate to the problem faced
DEMOS	- I assume the achieved status is not the end of the line yet. The field tackled by KRAKEN is fairly new and all stakeholders will have to familiarize themselves first with the processes and value propositions. Particularly SSI is a paradigm shift which demands a flexible, adaptable mindset. From experience of various go-to-market, production-deployed projects (particularly in SSI) I can only state that most problems surface when many end users put strain on a solution. Particularly critical voices usually first come up if something goes live, which means their merciless scrutiny will reach you only in a go-live scenario. - The demo was very interesting and shows how the marketplace works according to expectations.
Ethics Advisory Board Meeting	- I appreciate that ethical aspects are getting the attention they deserve by KRAKEN. - For a bigger picture in data protection, it gets easier to have very concrete use cases to assess. So, what the data protection partners will need is something detailed. What one will need to provide useful legal feedback is a clear picture on certain details including: - Entities: Which system components are there and who in the sense of natural or legal persons is operating these components? Which data flows are involved? Have this for use case to display and as a potential first blueprint for a later implementation. This is necessary to identify legal needs for contracts between joint controllers, controllers and processors and the necessary information of data subjects. Make clear in which role the Kraken platform will be active - as a controller or on behalf of its customers?
Pilot focus (WP5) Education	- The education use case is contemporary and will generate real-world usefulness. This is confirmed by the fact that the "diploma use case" is one of the most popular ones addressed by SSI communities in various jurisdictions throughout the world.
Pilot focus (WP5) Health	- The health use case is contemporary and will generate real-world usefulness. There is tremendous potential of re-using medical data that has been collected via different methods in a privacy preserving way. Current solutions are far from leveraging it to the possible extent. - The Pilot, with all the limitations of the case, can make it possible to identify and experiment methods of collaboration between the various stakeholders. It will be very important during the pilot to fine-tune the operational aspects that will allow the marketplace to function in the future
Future of the Project and sustainability	- See response under "DEMOS". I suggest transforming the KRAKEN results/technical prototypes into a minimum viable product and go to market with it. I advise putting special attention on legal matters as this innovative approach may be subject to many such barriers/pitfalls. - As always, the challenge of sustainability is very important. There are many market opportunities both as a service to the health system, for example as an infrastructure for multi-center clinical research projects, and for the conscious involvement of the public. Clearly it is a long-term project that must include the ability to sustain itself for a fairly long period of time. This means that it will be necessary

Section	Comments/Recommendations from Advisors
	to equip the Marketplace with adequate capitalization in order to operate with an adequate time horizon.
Final feedback from advisors	• I appreciate the hard work flowing into KRAKEN and the intriguing insights generated. The time is absolutely now to re-empower data subjects and resolve the downside of surveillance capitalism. Projects like KRAKEN are lighthouses for others in the data economy, which can drive shifting away from the unhealthy practices imposed on data subjects for far too long. • The current results and progress seem excellent. • Take into account the new eIDAS draft, eIDAS 2.0. Also, the evolution of GDPR assessment from Data Protection European institutions. • I think the project is of great interest for the future and there may be many applications.

Table 2: Recommendations from advisors

5 Conclusions

The Second Advisory Board meeting was very useful to better address the next steps of the project and constructive thanks to a positive collaboration with advisors.

The consortium received an overall positive feedback regarding the business and strategic approach adopted; at the same time, we received very interesting feedback and suggestions both on business side and technical aspects of the project, for example:

- The suggestion “A potential complexity of the “new” use case, which could be a usability risk” will drive our future choices in terms of design of new processes paying attention to guarantee a high level of usability in the perspective of end users.
- Ensuring a flexibility in processing the payment flows is a critical success factor. Considering the current developments in the “crypto market” crypto payments could be a solution, but at the same time, it is important to consider that Crypto volatility could be an issue. Moreover, in this context, we need to consider that Cryptocurrencies could be too innovative.
- Another relevant suggestion is that there are not only monetary compensations to consider in the reward system, but also transparent support for research or charity projects.
- As regards health pilot, the first motivation for patients to use the app is not monetary: patients are willing to share data to get feedback and some help to better manage their therapy. Basically, to get a direct benefit in the interaction with the hospital.

These recommendations will have an impact on the milestones to be achieved during the next year of the project and will affect almost all KRAKEN’s Work Packages. The deliverables influenced by advisors’ suggestion will be:

- D6.3 – Final market analysis
- D6.5 – Final exploitation Analysis
- D6.8 – Final communication, dissemination, and standardization report
- D7.3 – Ethical and legal evaluation and recommendations
- D5.6 – Marketplace final release

The third Project Advisory Board meeting is scheduled in November 2022, at the end of the project. The purpose of the third meeting will be to collect feedback from advisors that could guide the consortium in the launch phase of KRAKEN to the market.

@KrakenH2020

Kraken H2020

www.krakenh2020.eu

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 871473