



**POLY ANALYTIX**

**WHITE PAPER**

## Contents

INTRODUCTION.....	3
Data Market Context.....	4
VISION .....	7
PRODUCT DEPCRIPTION .....	8
PLA.....	9
TECHNICAL DETAILS .....	10
OVERVIEW .....	10
EVERYONE IS A DATA PRODUCER .....	10
DETAILED DESCRIPTION OF THE PRODUCT.....	11
DISCLAIMER .....	11
DEFINITIONS .....	11
DATA STORAGE PROCESS.....	11
PRODUCT RECAP.....	13
MARKET FOR DATA ANALYSIS .....	13
PLA & SMART CONTRACTS.....	13
SCENARIOS .....	13
TECHNOLOGICAL SOLUTION.....	13
FUTURE POTENTIAL .....	14
BIG DATA ANALYSIS STORE.....	14
DATA PRIVACY CERTIFICATE .....	14
DATA MODEL STANDARDIZATION .....	14
STANDARDIZED PRIVACY LAYER.....	14
CONCLUSION.....	15

# INTRODUCTION

The POLY ANALYTIX – Market for Data Analysis – ecosystem empowers everyone to participate in the data market with their own data, while simultaneously preserving their privacy. The PLA token enables anyone to buy data analysis results from a decentralized pool of information while rewarding data producers and plug-in providers for their contribution.

Individuals create significant amounts of data while using daily devices such as mobile phones, tablets, smart home devices, and computers. In most cases, data producers give up all rights of their data by agreeing to nontransparent terms and conditions. Data brokers sell this data to big corporations making substantial profits. The data producers are left without any control or profit. Currently, most of our data is stored in centralized servers, which are popular targets for hacker attacks, who gain access to sensitive information and leak it. The progress in global data protection policy is slow and the market for big data grows continuously. As the amount of breaches, leaks, and hacks is increasing severely, a new approach to handle data is needed. It is time to regain the control over our own digital identity.

POLY ANALYTIX is building an innovative data analysis market ecosystem based on the Lisk blockchain that allows data producers, data analysis buyers, and plug-in providers to become active in a fair, anonymous and privacy-protecting open data market. For the first time in history, decentralization is bringing transparency, balanced value distribution, and efficiency.

The POLY ANALYTIX ecosystem encompasses the following key elements:

- ▶ Client-sided data storage and easy to handle encryption framework to retake control of personal data
- ▶ Monetization of anonymous data contribution through PLA token
- ▶ An open, blockchain-based data analysis marketplace for everyone
- ▶ Monetization of analytics and data science skills through PLA token
- ▶ Reliable analytics of internal and external data for business purposes
- ▶ A decentralized, non-manipulatable pool of information

## Data Market Context

Data and analytics capabilities have made a leap forward in recent years. The volume of available data has grown exponentially, more sophisticated algorithms have been developed, and computational power and storage have steadily improved. The convergence of these trends is fueling rapid technology advances and business disruptions. Moreover, data and analytics are changing the basis of competition. Leading companies are using their capabilities not only to improve their main operations but to launch entirely new business models. Flows of data have created new infrastructures, businesses, politics, etc. The value of data is increasing as new methods and tools of data analysis are popularized. Furthermore, the extracted value is highly dependent on its ultimate use, and ecosystems are evolving to help companies capture that value.

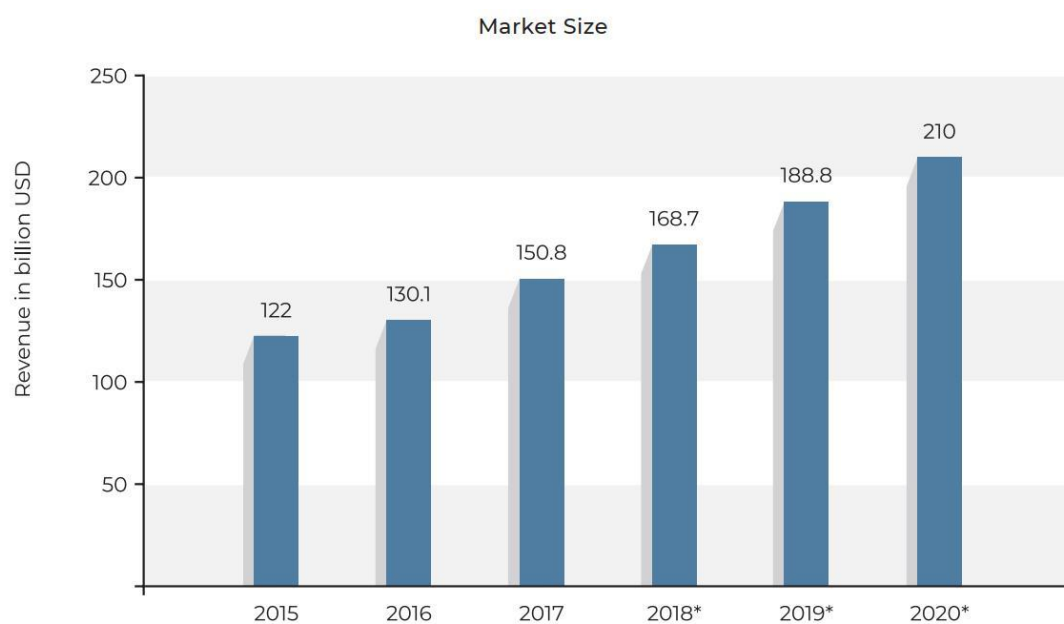


Figure 1 – Revenue from Big Data and Business Analytics worldwide from 2015 to 2020 (in billion USD)

## POLY ANALYTIX

The global data market size was predicted to grow from USD 9.7 billion in 2016 to more than USD 18.2 billion in 2018. The growth is directly connected with the world's data size expansion and with dynamic digital market growth.<sup>3</sup> According to E-marketer<sup>4</sup>, the global value will rise to over USD 375 billion by 2021. The knowledge about customers is essential for effectively tailoring products and services to customers' needs in the programmatic model

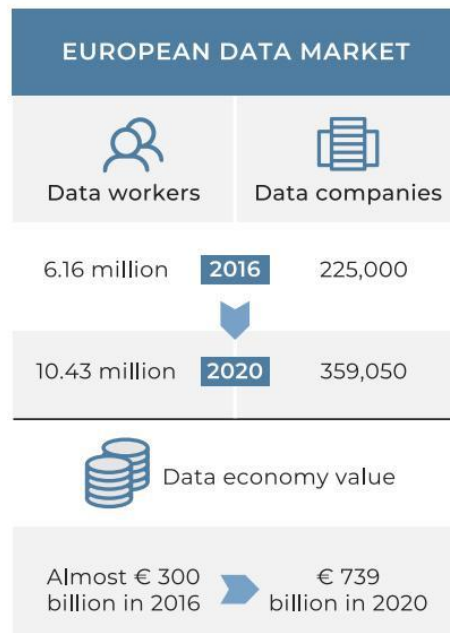


Figure 2 – European Data Market Supply

When it comes to Europe, the value of the European data economy was estimated at EUR 257 billion in 2014, or 1.85% of EU GDP. This increased to EUR 272 billion in 2015, or 1.87% of EU GDP (year-on-year growth of 5.6%)<sup>5</sup>. The European Commission (6) predicts that, if policy and legal framework conditions for the data economy are put in place in time, its value will increase to EUR 643 billion by 2020, representing 3.17% of the overall EU GDP. As shown in Figure 4, the top data markets for 2017 in Europe include the UK, Netherlands, France, Germany, and Denmark.

COUNTRY	DATA MARKET SIZE
UK	\$ 1,285 m
Netherlands	\$ 196 m
France	\$ 181 m
Germany	\$ 133.5 m
Denmark	\$ 78.5 m

Figure 3 – Top 5 Largest Data Markets in Europe in 2017

## POLY ANALYTIX

The data-driven transformation is spreading into every corner of the economy and society, ever-increasing amounts of data are generated by processes or machines based on emerging technologies, such as the Internet of Things (IoT). The new era of connectivity itself changes the way data can be accessed. The enormous diversity of data sources and types, and the rich opportunities for applying insights into this data in a variety of domains, including for public policy development, are only beginning to emerge. To benefit from these opportunities, both public and private players in the data analysis market need to have access to large and diverse datasets. The issues of access and sharing in relation to the data generated by these machines or processes are central to the emergence of a data economy and require careful assessment.

In a new data-driven economy, a new set of rules and regulations are required for the market to operate efficiently and effectively. Currently, data is collected from different parts and then marketed through data brokers. These data markets offer the service of combining data from different sources to generate new insights. This results in two major issues: **Data brokers** are in the position to take the major cut. The earnings only partially go to data gatherers such as app developers and companies. End-users, the true data producers, get nothing, or a fraction of the value, in form of rebates or bonus programs.

Unfortunately, until now **data markets** are rarely designed to be open. Data is kept isolated within the market it belongs to. As a result, valuable information that could be generated across data markets is lost.



Figure 4 - Data Markets Today

## VISION

Digitization enters more and more aspects of our lives. Data-driven technologies and their applications are making our world more complex, and the speed at which it is changing is rapidly increasing. Our digital privacy is also subject to this change. It is not always easy for normal consumers to keep track of the different developments. In the meantime, our Digital You leaves data traces everywhere it goes, which often leads to a significant breeding ground for digital crime

In order to prevent and guard our digital identities in the future, POLY ANALYTIX envisions, that through the use of upcoming technologies and methods a meaningful and social solution can be found. This solution is imaginable by utilizing and logically combining community-driven intelligence, secure encryption frameworks, decentralized applications, the latest data science methods, decentralization enabling technologies, as well as scientifically proven new information technologies and applicable game theory approaches. Balancing these elements in a savvy way has resulted in the creation of thought-provoking impulses and ideas:

- ▶ Providing the same access to data for everyone.
- ▶ Creating the world's open decentralized generic pool of information.
- ▶ Gathering innovative tools for data evaluation.

POLY ANALYTIX ultimately envisions that every single data producer – human or electronic device – is in full and transparent control of their generated data and the Digital You.

POLY ANALYTIX envisions a new data ecosystem approach, where data is the main resource benefitting all participants by being smartly used.

POLY ANALYTIX envisions a data ecosystem, which by itself is of self-regulatory nature driven by all its participants.

POLY ANALYTIX envisions that the world and societies will hugely benefit from a smart and fair applied data-driven ecosystem beyond the borders of various industries.

In order to reach the mentioned approaches and to realize the high-level visions a baseline for privacy-driven data handling must be set initially

# PRODUCT DESCRIPTION

## **POLY ANALYTIX – MARKET FOR DATA ANALYSIS**

The data market is currently dominated by enterprises who are buying big sets of data from third parties that are subject to different data protection laws. In known systems, these enterprises decide which data is available for sale. A problem with this is that all of the data is shared regardless of whether the data, or part of the data, contains information the originator does not want to share. Thus, the originator of data is exposed to an unnecessarily high risk of privacy violations.

POLY ANALYTIX is creating a new approach to provide a system implementing the method to eliminate the stated problems and to extend

the data market with further advantages. The objective is achieved with a system which provides the following key aspects:

- ▶ A method for contributing data anonymously to a remote system while getting paid.
- ▶ Client-sided data storage and an easy to handle encryption framework to retake control of data in general.
- ▶ A method for contributing analytics skills while getting paid in exchange.
- ▶ An open data analysis marketplace for anyone.
- ▶ A decentralized pool of information.
- ▶ A method for getting analysis results of new and sensitive data.
- ▶ Trustworthiness through blockchain and smart contracts.

The most important advantage is that data producer's privacy is respected, no third parties are involved, and entry barriers are minimized. Thus, a new ecosystem can be established in which everybody can participate – through transparency, trust, and monetary benefits.



## PLA

PLA will be the essential token for the POLY ANALYTIX ecosystem and will be based on the Lisk Blockchain. To provide trust, process transparency and to handle rewards for the contributions, the POLY ANALYTIX platform will be running on the PLA token through smart contracts.

POLY ANALYTIX intends to implement an exchange service onto the system, to support a convenient usability for all participants.

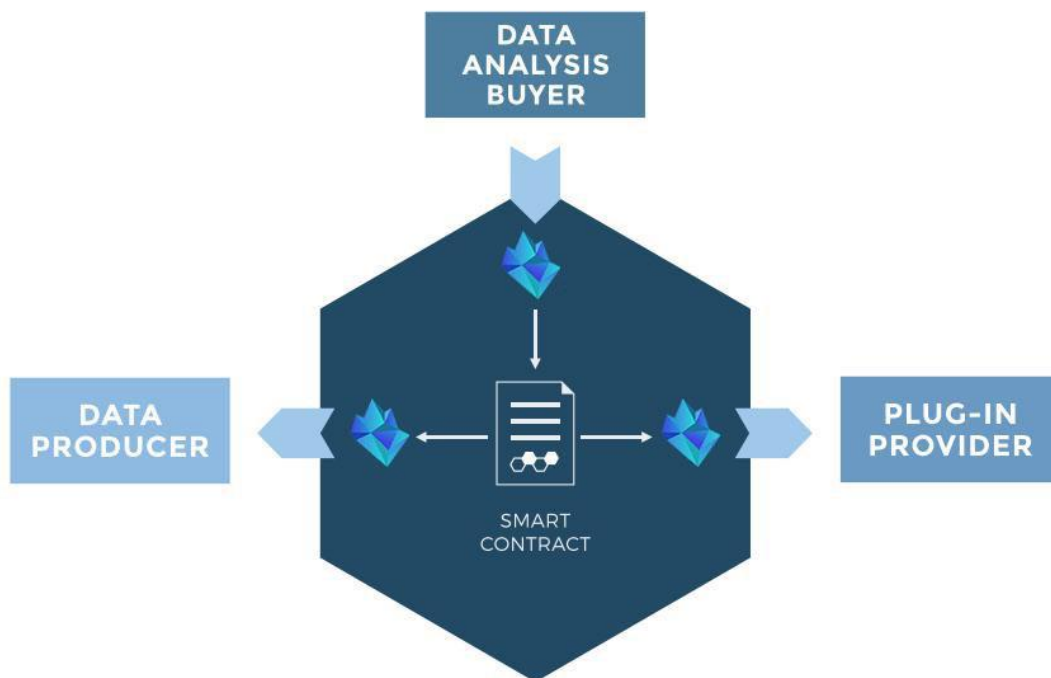


Figure 5 – Token Flow in the POLY ANALYTIX Ecosystem

## TECHNICAL DETAILS

This section provides a detailed description of the technical concepts, objectives, and processes of POLY ANALYTIX which are currently being developed. Additionally, it functions as an insight into the contents of the POLY ANALYTIX patent. Following the description of the ecosystem's participants i.e. data producer, data buyer entities, and plug-in provider in the previous section, these will be referred to as entities. In some cases, the data buyer will also be referred to as the actor.

### OVERVIEW

The product encompasses a method and a system for the protection of electronic data from origination up to data processing by third parties. With such a system, a data analysis buyer without knowing the identity of a data producer can provide funds in exchange for more unique insights.

### EVERYONE IS A DATA PRODUCER

Everyone nowadays is a data producer. One example is the use of digital services such as apps, browsers or the use of social networks such as Facebook, Twitter, Instagram, and others. Every electronic device that helps people, such as temperature sensors in rooms, milling machines, pulse counters, etc., also generates electronic data that can be evaluated and is therefore valuable. However, in most cases data producers cannot freely determine how much data is collected. It is customary for the data producer to transfer the rights to their data to a service provider by agreeing to the service provider's (often non-transparent) terms and conditions in order to use the service. Large companies in particular collect and use this data for profit, without the data producer being able to control or participate in the profit. The data market is currently dominated by enterprises who are buying big sets of data from third parties that are subject to different laws of data protection. In previously known systems these enterprises decide which data is available for sale. A problem with this is that all of the data is shared, regardless of whether the data, or part of the data, contains information the originator doesn't want to share. Thus, the originator of this data is exposed to an unnecessarily high risk of privacy violations.

Acknowledged by many, these centralized data servers on which the data to be analyzed is stored are increasingly being attacked and data is being stolen by hackers. This poses an enormous danger, especially for data producers if personal data such as bank details, addresses, telephone numbers, etc., fall into the hands of unwanted parties.

# DETAILED DESCRIPTION OF THE PRODUCT

This section provides a detailed description of how the components within the system are working together to fulfill the demands stated in the process overview above.

## DISCLAIMER

The product will be described more in-depth with reference to accompanying drawings. These drawings will show, by way of illustration, specific exemplary embodiments by which the product may be designed and implemented. Nevertheless, the product may be represented in many different forms and should not be construed as limited to the realizations set herein. These embodiments are provided so that this description will be thorough and complete, fully conveying the scope of the product to those skilled in the art. Among other things, the future product may be embodied as methods or devices. It may also take the form of an entire hardware embodiment, an entire software embodiment or an embodiment combining software and hardware aspects. The following detailed description is, therefore, not to be taken in a limiting sense.

## DEFINITIONS

**Data** is generally understood to mean information, (numerical) values or formulated findings that have been obtained through measurement, observation, etc. According to the product, data is all electronically recorded information that applies to an object or event. When data is processed, data is defined as characters (or symbols) that represent information and serve the purpose of processing. Data protection law essentially refers to personal data, i.e. information about natural persons, such as gender, date of birth or place of residence.

A **data producer** can be a natural person who enters information about themselves, such as their clothing sizes when shopping online. A data producer may be a legal person or a community of persons. The data producer can also be a machine that either generates data itself by executing instructions or contains sensors that record the temperature of a room, for example.

A **third party** is any person or instance that is not represented as a data producer. For example, it is often a platform (Facebook, Twitter, Google+, etc.) that users can only use by agreeing that certain rights to their data are transferred to the platform (third party). Any broker who transfers data from one provider to a buyer and receives the data is a third party. A direct purchaser of user data is also a third party in this sense.

A **user device** can include any electronic application that has an input and an output or interface to process the data. The user device may also include any type of data processing device capable of receiving and transmitting data over a network. For example, the user device can be a computer, a mobile phone, a laptop, a tablet, a server, a smart-watch or any combination of these devices. The user device is there to record the data of a data producer. In the case of a smart IoT device as data producer, the data producer could also directly be the device.

## DATA STORAGE PROCESS

One part of the product to be developed will be an application which will run on a device. This is the data collecting part of the system: The application gains access to data sources and can trigger storage processes. A device in this case may be any computing device which is capable of receiving and transmitting a message over a network. The category of "device" may

## POLY ANALYTIX

include smartphones, handheld computers, IoT devices, wearable computing devices, tablets, desktop computers, servers or any device combining one or more of the preceding devices.

The main objective of the data collecting part of the system are two software interfaces: a data-viewing interface and a connection interface.

These interfaces will be open-source and can be implemented into every kind of application. To ensure the integrity of all participating processes, all executed code must be signed off by a certificate authority of the Main System. This is done implementing scripts that use a cryptographic hash to validate the authenticity and integrity of the code. The hash is used to verify that the code has not been modified and that the correct version is available. If this is not the case, the proceedings will be suspended.

These interfaces are used to encrypt and save different types of data that the application is generating and update the meta-information that is stored on the Main System. The application itself has to trigger events that activate the start of the processes that follow.

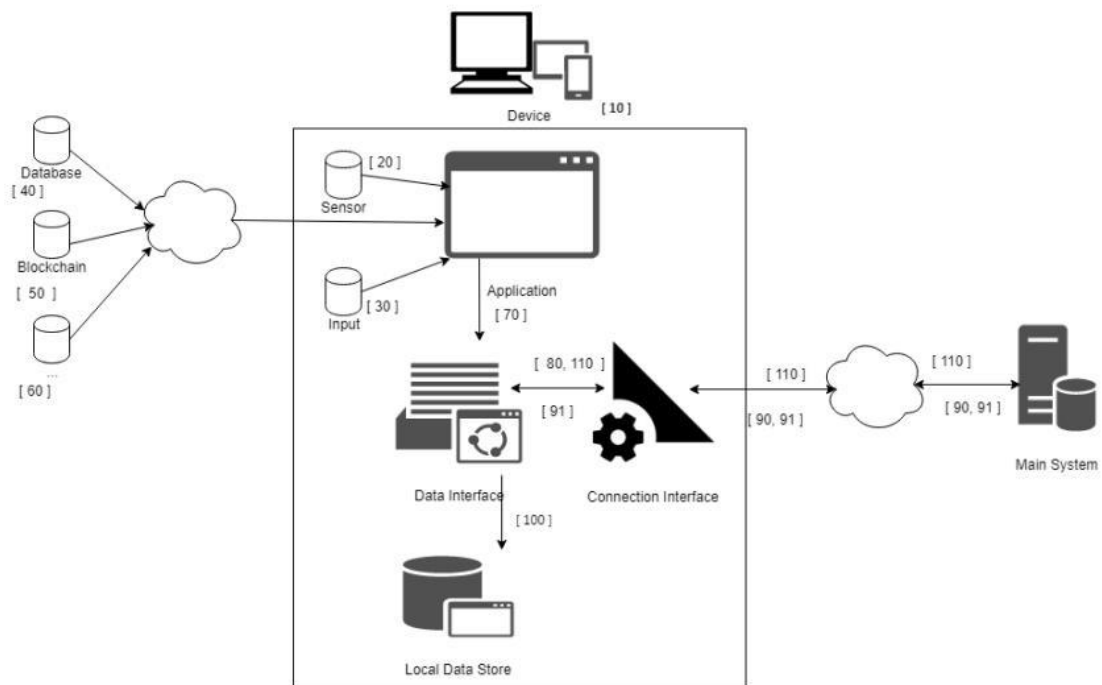


Figure 12 – Generating Data – It depicts an example of a system for collection of various kinds of data which is later used by a remote system for further processing.

## PRODUCT RECAP

POLY ANALYTIX's ecosystem approach is not that easy to understand at first glance. A short product recap summarizes the preceding chapters.

### MARKET FOR DATA ANALYSIS

Using blockchain technology, three main data market participants will be able to operate on the decentralized POLY ANALYTIX ecosystem. Individual data producers will be given the opportunity to provide and monetize their data, which will be encrypted and cannot be manipulated. This data thereafter will flow anonymously into data analyses, which results companies will be able to buy without gaining access to the raw data itself. The analysis plug-ins will be provided by either analytics companies, freelance developers or data scientists, who will monetize their skills and knowledge in a fair way.

### PLA & SMART CONTRACTS

Transactions will run on the basis of smart contracts and the POLY ANALYTIX token – PLA. Data analysis buyers will purchase analysis results with PLA. One PLA stake will flow directly to the individual data producers, another to the plug-in providers and a small one to POLY ANALYTIX for maintaining the system.

### SCENARIOS

The next step for POLY ANALYTIX is to start pilot projects in order to test the marketability of the ecosystem and jointly collaborate with strong partners to gain valuable insights and finalize a well-rounded product. Since the POLY ANALYTIX solution is a generic one, it is applicable in many data-driven industries. POLY ANALYTIX is aiming for positioning itself in industries like machine data, energy, eHealth, data analytics and IoT among others.

### TECHNOLOGICAL SOLUTION

POLY ANALYTIX is developing a decentralized data analysis market. It strives to provide a fair platform, that allows for data conversion into analysis results for anyone; companies, research institutes, and other interested parties, while offering data scientists and developers the opportunity to monetize their skills.

## FUTURE POTENTIAL

The key driver of POLY ANALYTIX's strategy and its decision-making approach is its long-term focus. Recent concerns are not resolved by simply providing a single solution to ensure data security and digital privacy. Instead, a holistic approach must leverage the full potential of POLY ANALYTIX and the data market. As an outlook for the future, POLY ANALYTIX sees multiple opportunities for integrating the product in the everyday life at the social level and in the daily business making of companies.

### BIG DATA ANALYSIS STORE

Uploading various analysis schemes by plug-in providers and making these schemes available to every participant, would ultimately result in a "Big Data Analysis Store". Hence, analysis buyer could browse online for the right analysis scheme, like today's app store. Data scientists on the other hand would have a great community source for extending their analytics schemes as well as getting inspirations for new analytics schemes.

### DATA PRIVACY CERTIFICATE

Companies that are hesitant to monetize their data in fear of customer backlash would be enabled to make use of their data in a transparent and open way. By buying data analysis from the POLY ANALYTIX ecosystem, companies could advance their image by stating a data privacy certificate provided by POLY ANALYTIX guaranteeing data privacy.

### DATA MODEL STANDARDIZATION

Data is widely unstructured. Valuable information is available for free on the Internet. With POL token as incentive experts could focus on aggregating data, restructuring it into standardized data models and feeding it into the POLY ANALYTIX network. Sources for this kind of data could be: public repositories, government statistics, stock market data or data from public websites such as Twitter, Crunchbase or Kaggle. New technologies are expected to provide new datasets which initially will be heavily unstructured, e.g. in the sphere of biogenetic data. By democratizing data model standards, data scientists in the POLY ANALYTIX ecosystem would eventually be able to elect standardized data models for each emerging industry.

### STANDARDIZED PRIVACY LAYER

The technology underneath the POLY ANALYTIX ecosystem basically acts as a privacy layer, which POLY ANALYTIX plans to make available for developers in a framework. By implementing the privacy layer into applications, developers could then create privacy-secured applications. Therefore, POLY ANALYTIX is having a competitive edge by securing data privacy of the end-users along with an additional revenue stream through data monetization. Further, creating POLY ANALYTIX on the Lisk Blockchain bears the opportunity for future Lisk dApps to integrate the POLY ANALYTIX privacy layer into their systems.

The security and openness of POLY ANALYTIX will also allow all kinds of data to be added to the system. The added value will shape the standardization of new data types and combine data in new ways. Completely new cross-analytics unlock unseen value for science and society.

## CONCLUSION

The data markets will change, and data producers will call for adequate compensation for the usage of their data. They will demand transparency and control over the processing of their information. Politics will need to adjust to these needs, because society's mindset is rapidly changing as its awareness of data control and privacy is increasing ever since recent data scandals. POLY ANALYTIX acts on exactly these issues. The system is meant to provide a solution for all parties involved in the data market, which ensures access to data, data privacy and compensation for data producers.

POLY ANALYTIX is developing a privacy layer for future applications in various industries, where data-driven decisions and processes are prevalent. Thus, making it possible to envision and to build a blockchain-based, decentralized ecosystem for a fair market for data analysis,

POLY ANALYTIX's ecosystem participants, the data producers (individuals and electronic devices), data analysis buyers, and plug-in provider (e.g. data scientists and application developers), will receive full benefits in the form of tokens or access to data and highest efficiency thanks to a thought-out design and newest technologies.

POLY ANALYTIX wants to make every individual data producer the controller over their own data. Backed by blockchain technology, POLY ANALYTIX's decentralized data market will be the data analysis solution of tomorrow.

Going beyond that, POLY ANALYTIX wants to address multiple needs in data-driven industries. New partnerships are meant to help POLY ANALYTIX to develop novel use cases and utilize its innovative cross analytics capabilities.

POLY ANALYTIX's uniqueness is defined by the combined features of all parts of the ecosystem. It is about the aggregated potential of all features from each part of the ecosystem. Reaching these potentials will drive the POLY ANALYTIX project towards the bigger vision as an established standard for data privacy and easy data market accessibility – where privacy is guaranteed by design.

POLY ANALYTIX envisions a world, where everybody can use data-driven services without giving up their privacy and disclosing their **“Digital You”**.