Technical Paper

Part I

The STIPS|Oculus platform will accumulate a lot of raw data on each registered asset. We will collect this information with our own instruments from plenty different sources and will receive it from active users of the community who will be rewarded for the provided information.

Then this information will be aggregated to generate integrated reports or the information will be used to calculate other data that can be of use for investors/traders to make analysis.

In total there are 15 different types of data that we collect and analyze. They are:

1. Code activity on GitHub

Oculus will accumulate information about all open-source products associated with the assets on the platform. Thus there will be a software library. Then the platform will analyze software update periods and this information will be recorded in our database. As the result **the development activity index** that can be used further for comparison of different projects will be calculated on the basis of this data by a proprietary algorithm.

Our end user will have access to the software catalogue on an asset, consolidated information about update periods as well as development activity index.

We will apply release dates to the price graph of relevant software.

2. Asset Flow on Major Wallets

This data type is the most voluminous in our database. The community participates in collection of this data. In order to collect it we deploy our own nodes with a complete blockchain of each token. For minor tokens it is possible to use third-party blockchain explorers to receive all data about transactions.

Then all these transactions are transferred to our database for further analysis. Oculus user will be able to make analysis of this data manually, search, select with a filter, view the largest wallets or the largest transactions for a certain period. The user can compare this data with the price graph. For instance, apply all incoming/outgoing transactions of a specific wallet to the price graph of this token.

In addition we will search for such impacts using IBM SPSS (part of IBM Bluemix) and our own infrastructure will be used not only for automated check of hypotheses, but also for real-time monitoring of events that impact on price. In future we will use the results of this work to make profit from our STIPS Smart Fund.

An additional feature of STIPS Oculus based on the collected data will be graphic representation of cryptocurrency flows what will allow users to see all the chain of transactions of specific assets.

3. Marketing Activity Tracking

Using resources of our community we will collect information about marketing activity of different projects, for instance, about their participation in exhibitions, conferences, competitions (e.g. startup contests), roadshows. At the same time we will take into consideration the level of events (country, cost of participation, ticket price for attendees, number of participants and attendees). We will also consider participant status (sponsor, award winner, or winner of a competition).

Any user of the STIPS Oculus platform will be able to view all this information in marketing calendar of a project (past and planned events).

In addition we will collect information about advertising activity on different platforms and other digital advertising.

This information will serve as a basis for calculation of ***the marketing activity index*** that can be used further for comparison of different projects.

4. HR Activity

We will also collect information (with the help of our community) about all participants of all projects. We will collect information about staff members as well as advisors. All these records will be associated with certain dates. Oculus user will be able to view company staff and its history. For instance, staff growth and employees who left a project.

Other information that may be useful is information about participation of staff in different projects. There will be a record on each employee and advisor with the history of his/her participation in various projects. This will allow to see if a person participated in a scam project before.

On the basis of this data two internal indexes will be generated: ***the staffing index*** of a certain project to assess, if all the key competences are covered in a project or there are missing positions and ***the credibility index*** of a certain person.

5. Roadmap History

With the help of our community we will collect information about roadmap of each project. At the same time we will keep all of its revisions. With the help of our community and our own algorithms we will receive information about roadmap revisions. It's pretty difficult to assess these changes automatically, however, we will provide this information to any STIPS Oculus user to perform his/her own analysis. In addition our user will have the opportunity to apply roadmap revision dates to a price graph to analyze dependance.

6. ICO Rating Aggregator

There are plenty of services in the web that rate ICO projects in general or just rate their reliability. We are going to use our own facilities to aggregate all such services and show a single table of all ratings on each project (token). Also, based on this and other data, we will generate our *trust index*.

STIPS Oculus user will be able to filter a list of all projects by this index that is, for instance, to see all projects rated 8 and higher.

The community will not participate in this process, since ratings will be chosen automatically. We will need the help of our community only to form a list of such services (adding new service).

7. Trade Signal Channel Aggregator

There are also plenty of projects in the web that give their trade advice on various assets (so called trade signals). With the help of our community we will create a database of such information sources, then using semantic analysis (it is planned to use IBM Watson NLU Text Analysis) we will sort this information by types (trade signals, market analysis, etc.) and provide it in a structured way to STIPS Oculus users with the indication of the information source. Thus, there will be a single feed of all signals on each project.

8. News Aggregator (Sentiment Analysis)

Another large array of information in our database is various news on each project. We plan to save information about a project from multiple sources to our database. Then we will assess importance and utility of an article and estimate sentiment of this information (positive or negative) using different instruments of sentiment analysis (we are going to use a number of instruments from IBM Watson Natural Language Understanding, and similar solution by ABBY - Compreno, Infoextractor SDK).

Eventually each Oculus user will read all the necessary information on a project from the mass media in one place. At the same time there will be filters to make information more relevant. In general there are a lot of such services in the Internet, but our solution offers a number of advantages.

First, with all this information (absolute number of articles and their contents), we will generate **the writing index** and **the sentiment index**, which can be used to compare different projects.

Second, we will store history of the indexes. Thus, it will be possible to assess dynamics of how they change. So, any STIPS Oculus user will be able either to keep track of a project news or to apply dynamics of the indexes to a price graph .

Considering complexity of such analysis initially it will be made only for the English language, but step by step other languages will be added (Chinese, Russian, etc.)

9. User Sentiment Monitoring (Sentiment Analysis)

In addition to the huge database of publications we will pay special attention to user communication on different platforms: forums, chats in Telegram, Slack, Discord, etc. All this information will be accumulated in our database for further analysis.

We will take into consideration both overall level of communication on each project and topic analysis of this communication. Then this information will be aggregated in the general report. STIPS Oculus user will see communication activity. For instance, the user will see that this week the communication activity went up drastically. He/she will also be able to get a comparative analysis of what topics are most discussed by the community of a project. For instance, in some project users discusses mining, in the other they discuss trading or ICO, the third one constantly needs more technical support to make settings of the application software (wallets, miners, etc).

We will also store history of these general reports, therefore Oculus users will be able to see communication activity dynamics of a project.

10. Project Listing Analysis

One of the most important indicators of project quality and success is its listing on exchanges. This is because top exchanges impose strict requirements on projects for listing. Save that for top exchanges listing cost can exceed one million US dollars, it is compulsory, but not sufficient condition of listing. It is necessary to comply with a number of requirements and undergo a serious check. So, one can get the impression about a project on the basis of a listing.

On our platform we will collect information about exchanges, where a project is traded. The information will be collected automatically with our own infrastructure. We plan to use API for integration with each existing exchange. Thus, we will have all information about listing of each current project.

Any STIPS Oculus user will be able to see all the platforms where a token that interests him/her is traded and compare prices on different exchanges (with an option of buying it).

In addition to this list of exchanges we will generate ***the listing index***, which can be used to compare different projects. We will also store history. Thus, the user can get a history of project listing on an exchange and more importantly when it was delisted. Delisting may have a profound effects on trading instrument price and it can be a sign that all is not well with a project. All this information will be available to our users both in real time and as active notifications by subscription.

Another bonus feature will be notifications about new listings of a project. Along with this our solutions allow to learn about the forthcoming listing several hours before the information becomes public. It can be used for making profit from this information. It's quite often that listing on a large exchanges results in a sharp, but short-term token price growth, that is why the exact date of listing or the forthcoming listing in general is often kept secret. We will have this information.

11. Blockchain Use Dynamics (for Smart Contracts)

For the blockchains that can be used by third parties for service provision (for instance, so called smart contracts), we will collect all information about such sister projects.

First, all our users will see the list of all sister projects.

Second, on the basis of this information we will generate ***the smart contract use index*** to compare different projects.

Third, we will store and show dynamics of this index. Thus, any Oculus user will be able, for instance, to see sharp growth of a blockchain use at a given moment.

We also plan to collect information about blockchains that projects use to hold an ICO and about currencies they accept. This information will be aggregated and applied to a price graph. For instance, surge in price of Ethereum in 2017 is often associated with the large number of ICOs using either its blockchain for their own smart contract or accepting payments exclusively in this currency. After having applied this information to a price graph STIPS Oculus user will check or reject this theory and will probably use it further in his/her trade strategy.

12. Analysis of Token Distribution Across Wallets

With our own blockchain nodes we will be able to collect statistical data on each wallet for any time period.

We will offer our end users the feature of viewing token distribution in blockchain across wallets in the form of a user-friendly report. This report will show the volume of total emission stored in the largest wallet, the volume stored in the ten largest wallets (TOP10), in the hundred largest wallets (TOP100), and in the thousand largest wallets (TOP1000). On the basis of this information Oculus user will be able to draw conclusions about distribution of tokens and relevant risks (if a large volume of tokens is in the hands of one person, it means a potentially high risk of exchange rate drop, if the owner wants to sell them).

In addition it will be possible to view both history and current data. Thus, the user will see dynamics of token distribution.

13. Trade Instrument Volatility

We will calculate ***the volatility index*** with the help of open-source data that we will collect on our own using a special algorithm for each token (listed).

It will be possible to assess the index change and compare it with the price graph, if necessary.

14. Trading Volume Analysis

Integration with each existing exchange will allow us to assess trading volume of each token on all exchanges on the whole. A user of the Oculus platform will be able to look through this information and study trading volume dynamics as well as compare it with the price graph.

15. Project Jurisdiction Analysis

We will collect information about all legal entities of each project. We will validate this information with the help of relevant inquiries to registers of legal entities (the Uniform State Register

of Legal Entities, etc.). Information about open bank accounts (if it's available in open sources and it can be checked) as well as about established offices (for instance, from descriptions of jobs offered in a project).

STIPS Oculus user will have access to such consolidated information to study. We will label projects drawing user attention, for instance, to inconsistency of jurisdiction of incorporated legal entities, accounts, and territory, where their financial activity is being conducted.

Part II: Indexes

Thus, based on 15 types of data Oculus will calculate 10 indexes:

1. Development Activity Index

The index shows how active is the development of various software for this project. Both official software and third party (community) software are taken into account. The index is calculated based on the entire project life cycle, i.e. it shows not the current state, but the average state over a lifetime. So, even with zero development activity recently, the index may still be high due to the past activity. The graph of this index is perfect to see the whole picture. For instance, it helps to see that the activity has been running down recently.

2. Marketing Activity Index

It shows how active is the company in marketing, if it arranges its own events (the way Ethereum does) or participates in some big events. The index will take into account both the scale of the event and the status of the company in it. The data is accumulated and averaged. The graph of this index will show bursts of project activity at different times.

3. Staffing Index

It shows how well a project is staffed, technical department, legal department, its own development and so on. All this will affect the final index. The index shows the current situation, but the user will be able to see the history of this index.

4. Project Member Credibility Index

This index is set individually for each person who participates in a project (or in several projects). If an employee was spotted in serious successful projects, his/her credibility index will be higher. If he/she was seen in scam projects, the index will be lower. The index reflects only the current situation, it will not be possible to view its history, since mistakes and appeal are possible in this assessment and the rating may change significantly over time due to the correction of mistakes.

5. ICO Project Rating Index

This index is an average indicator of the ratings of various services that audit an ICO, adjusted for some of their own metrics. For calculation of the index specific assessments of various companies will be taken into account with different weights depending on the prestige of an auditor. Finally, the index will also be affected by the information about the project, which is not taken into account by third-party auditors. For instance, participation in a

project of people with a bad reputation (who were seen in scam projects), size and staffing, possible conflicts in jurisdictions of the legal entity, place of business and country of residence of key project participants. The index will show the current situation, however, the user will also be able to see the entire history of this index.

6. Writing Index

This index shows how often one writes about a particular project on various resources and in various media. In order to calculate this index the number of published materials and the credibility of a resource or a media will be taken into account . As a result, an average index will be generated that will allow comparing different projects or draw sample. The entire history of changes of this index will be stored, therefore, the user will be able to monitor the burst of maximum activity.

7. Sentiment Index

This index reflects the sentiment of people about a particular project. If they say that a project is a scam and it's rolling into the abyss, then this will definitely affect the value of this index. The ratio of positive and negative comments to the total number of comments will be considered to calculate this index. However, in order to generate the index a critical mass is required. If there is too little communication regarding any project in the community, then this information will be recorded, but the index will not be calculated, because volatility is too high with a small amount of input data.

8. Listing Index

Based on this index, one can understand how easy it is to buy, sell or exchange a trading instrument of a specific project. Calculation of this index will take into account presence of a token on exchanges and in currency pairs. At the same time the fact of listing on a particular exchange will be taken into account with different weights, depending on the exchange credibility. The exchange credibility, in its turn, will be defined by the relative turnover of the exchange in total and other features and criteria of the exchange. For instance, a listing on a fully regulated exchange that has the accompanying licenses will have a greater impact on the growth of this index than a listing on any decentralized exchange with the same turnovers.

9. Smart Contract Use Index

This index shows how much this blockchain is in demand for other projects in order to create their own token. With the help of this index it will be possible to compare different projects. The value of this index is linear in the number of subsidiary projects using a particular blockchain. Increasing number of such projects increases the value of the index. Despite the fact that the number of used projects cannot decrease, the index value of a specific project may decrease due to the growth of this indicator in other projects, since the index shows the relative value in comparison with the general indicators.

10. Volatility Index

This index shows how volatile the trading tool of a project. With this index it will be possible to select, for example, the most volatile or, on the contrary, the most stable tools. The index reflects the current value of the volatility of each tool, however, the user can see the history of its change.

Part III: Architecture and Tools

Below you will find information on the architecture of the analytics platform and options for implementing standard scenarios for analysis and research of financial data. The services of the Platform are designed to support decision-making on investments and financial asset management for STIPS FinTech. Strategic business initiative of STIPS FinTech is to use innovative and analytical information tools for investment management and informed decision-making.

Goals of the current STIPS FinTech projects

1. Create a computational platform for operations with data on decentralized cryptocurrency systems and service provision for smart investment and asset management.
2. Provide mechanisms for collecting and replenishing the repository and data collections with complete, relevant, and unbiased information on each asset.
3. Create services based on cognitive technologies for data analysis, modeling, and evaluation, implement procedures to create, train, and apply analytical models for financial tools and markets.
4. Provide tools for analysts and business users based on cognitive services to make business decisions.

At present STIPS Fintech create and develop a range of high-tech tools and technologies such as the decentralized STIPS Oculus platform with innovative algorithms of fundamental analysis, the Stips20.io cryptocurrency index for traders and private investors, STIPS 20Fund, the passive index fund for investing in cryptocurrencies based on the buy and hold strategy, STIPS Smart Fund.

Operation and development of financial tools requires a high-performance computing platform with a wide and expandable set of analytical and information services with elements of artificial intelligence.

In addition to the wide range of analytical functions, the platform should be flexible and scalable when using computing resources and services, highly reliable and safe, automating the processes to create, maintain, and operate financial services and tools. The platform should store and process unlimited volume of data and documents.

Creation and implementation of an analytical platform based on a scalable and reliable computing infrastructure, both cloud and in-house, is an extremely important task.

General requirements to the platform are:

- reliable software to receive, store, process, and analyze a variety of public and private data related to financial processes, cryptocurrencies and investments
- implementation of analytical services for data analysis and decision making
- providing software and visual interfaces for data management, modelling, and creation of new services, providing access to data and services

Using Data as a Tool for Innovation and Business Development

Data has become a source of business innovation in proportion to the amount of data available. News and streams in social networks, data on financial markets are information-intensive data available for

analysis. A company capable of using these sources, separating valuable information from background processes, seeing relationships and patterns in data, and then acting in accordance with this knowledge, is better prepared for competitive business than any other.

Open and Scalable Solution Architecture

The architecture of the Platform should be based on usage of services with standardized interfaces, multiple use of functional elements, elimination of unjustified duplication of software features, unification of standard operational processes.

Architecture of the Platform should be:

- open to add new applications (subsystems, tasks, etc.) due to change of current processes and emergence of new processes and management techniques introduced in the current rules of a company;
- flexible and customizable;
- fault tolerant to deviations of the parameters of the controlled object (work model), errors of the maintenance personnel and users, failures of hardware and software;
- scalable that is capable of providing functionality and allow expanding or reducing capacity;
- interoperable, that is, capable of conflict-free exchange of information with other information systems and users.
- portable, that is functional when transferred to another software and hardware platform.

The Platform should allow cloud, in-house or mixed infrastructure depending on technological and organizational conditions of the company.

Reliable and Scalable Storage, Data and Content Management System

Requirements to the storage, data and content management system include provision of high reliability and centralized storage of electronic data on cryptocurrencies, financial tools that eliminates or reduces to a minimum risks of data loss and damaging.

High availability of data in compliance with the life cycle of data and document collections as well as saving access history. Support for processing, search, format transformation features and various access layers.

Integration with Contiguous Systems, API and Technological Interface Support

For integration with contiguous system the Platform should have a number of application programming interfaces (API) in the form of software application services.

The Platform should support standard communication mechanisms such as SOAP, MQ, JMS, COM etc.

# Continuous Availability and Reliability of the System

The platform should operate continuously except for periods of maintenance and process works described in the regulations as well as in emergency situations.

There should be a log of the platform on key events of operation that will include registration of errors with indication of date, time, information about errors as well as communication protocols between the nodes of the Platform.

ARCHITECTURE AND SYSTEM CONTEXT

The architecture of the Platform is based on a hybrid cloud infrastructure of IBM Cloud.

The software designed for implementation of the Platform features includes instruments for modelling, analysis, execution of analytics services and their use processes. IBM offers IBM Cloud that is a computational platform for implementation of the service oriented architecture. Below you will see a solution pattern based on IBM Cloud including IBM Analytics Engine, IBM Object Store, IBM Cloudant, IBM Watson Studio, IBM Knowledge Studio, IBM Discovery, IBM Streaming Analytics, IBM API Connect. The solution shows the current components, system environment, data communication channels.
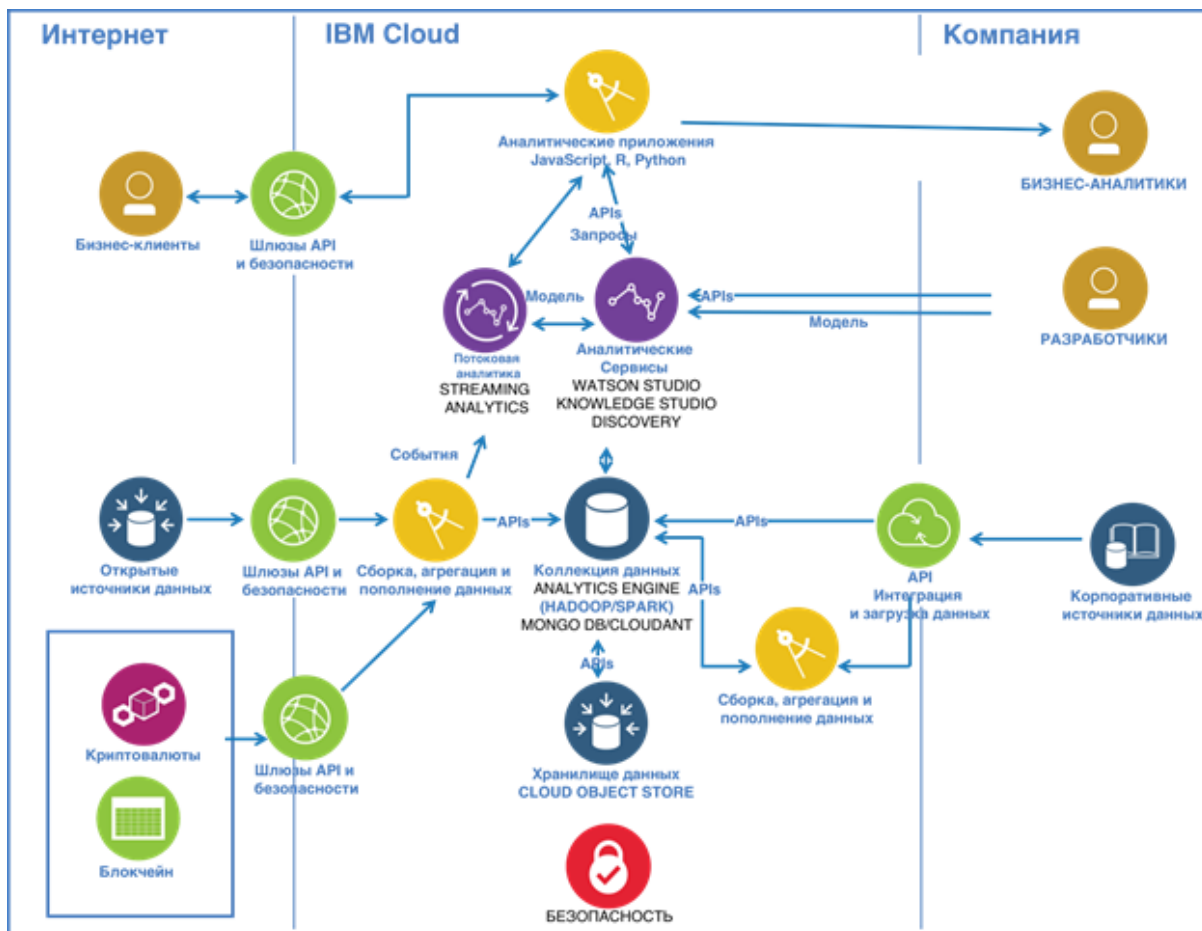
One of the advantages of the platform's architecture is its extended functionality and scalability of the solution, capacity to implement full and automated lifecycle of data and service management, access

to service and data management mechanisms, what allows to see and control the state and operation of the Company's business services with the means of the platform.

Solution Configuration

The solution that we offer is multicomponent and includes key subsystems

- System for collecting and replenishment of external data, including data from distributed cryptocurrency networks, data aggregation and transformation
- System for storage of structured and unstructured data
- System for development, modelling, setting, training, and execution of analytics and cognitive services
- System for development and execution of client and front-end applications
- Administration and monitoring subsystem responsible for status monitoring of the system components, administration, and configuration of the system components
- Safety subsystem responsible for authentication and identification information check, electronic digital signature check, audit
- Components of the system analytics and reporting, and other subsystems.

There is IBM Analytics Engine used as a basic application cluster platform for launching analytics services, application, and server-based computing. IBM Analytics Engine together with IBM Watson Studio serves as a converged platform for computing in distributed environments of Apache Spark and Apache Hadoop, and other subsystem operate on its basis. In order to process data collections for multiple computational and model problems it is convenient to use database services such as MongoDB or IBM Cloudant. Cloud Object Store is used as a scalable data store.

Web applications. The architecture implemented in the digital platform has a unified point of access to internal and external Web applications as well as access to the Company's applications. The portal offers an integrated content and applications as well as unified workspaces for teamwork.

IBM Cloud platform offers opportunities for implementation and execution of analytics and cognitive services, decision-making, collection of content and data for all the aspects of financial transactions by providing necessary compute resources in a flexible way.

IBM Analytics Engine

IBM Analytics Engine is a consolidated service for calculations in two distributed environments Apache Spark and Apache Hadoop. IBM Analytics Engine provides environment for development and deployment of advanced analytics applications. It utilizes HortonWorks Data Platform as an open-source Hadoop distribution, that is why it enables to provision Spark and Hadoop clusters in minutes as well as improve interaction, facilitate work with users, scalability and reliability. Compute and long-term storage infrastructure are separated, that is why each one can be scaled independently to avoid loss of data, if compute cluster goes out of services.

IBM Watson Studio

IBM Watson Studio provides a team of analysts and developers with a single environment for development and deployment of analytics solutions. IBM Watson Studio boosts machine and deep learning necessary for implementation of AI methods in business to drive innovation. IBM Watson Studio provides a toolkit for research scientists, application developers and experts in this field for collaborative and easy work with data and using this data to create, train, and deploy machine learning models.

IBM Watson Studio enables data analysis with RStudio, Jupyter and Python in a customizable shared environment that includes services and extensions from IBM as well as components to manage and use Spark clusters. IBM Watson Studio offers all these functions, including:

- a single workspace for different instruments
- communication and search space for specialists
- environment for collaborative project management of content management function

IBM Streaming Analytics

IBM Streaming Analytics analyzes a broad range of streaming data – unstructured text, video, audio, geospatial, sensor in real time.

Streaming Analytics is based on IBM Streams, an advanced computing platform, that can be used to analyze and correlate information as it arrives from thousands of real-time sources.

IBM Knowledge Studio

IBM Watson Knowledge Studio is designed to create a machine learning model that understands the linguistic nuances, meaning, and relationships specific, or to create a rule-based model that finds entities in documents based on rules that you define.

Knowledge Studio provides tools for annotating unstructured domain literature, and uses those annotations to create a custom machine-learning model that understands the language of the domain. The accuracy of the model improves through iterative testing, ultimately resulting in an algorithm that can learn from the patterns that it sees and recognize those patterns in large collections of new documents.

IBM Cloudant

IBM Cloudant is a fully managed JSON document database. It is designed with regard to scalability, high availability and reliability. It has a wide range of indexing mechanisms including MapReduce, full-text indexing, and geospatial indexing.

Replication capabilities allow synchronizing data between database clusters, desktop PCs, and mobile devices.

IBM Cloud Object Storage

IBM Cloud Object Storage is a scalable cloud storage for a high fault tolerance and safety.

IBM Discovery

Watson Discovery uses IBM Watson Natural Language Understanding, the collection of API on the Watson platform that offers expanded text analysis with the help of natural language processing, machine learning and deep learning. This set of APIs can analyze a text to extract concepts, entities, key words, sentiments, relations, emotions, and taxonomy.

Watson Discovery can also be integrated with IBM Watson Knowledge Studio that provides an integrated development environment to design and deploy user annotators where domain specialized knowledge is used. Knowledge Studio can be used by experts of a subject or an area in the field of data. Trained cognitive services can be used for decision-making purposes.