Data Lakes for Insurance Industry: Exploring Challenges and Opportunities for Customer Behaviour Analytics, Risk Assessment, and Industry Adoption

Bálint Molnár[1][a], Galena Pisoni[2][b] and Ádám Tarcsi[1][c]
[1] *Eötvös Loránd University, ELTE,IK Pázmány Péter 1/C, 1117, Budapest, Hungary*
[2]*University of Nice Sophia Antipolis, Nice, France*

Keywords: Data Warehouse, Data Lake, Enterprise Architecture, Data Science, Insurance.

Abstract: The proliferation of the big data movement has led to volumes of data. The data explosion has surpassed enterprises' ability to consume the various data types that may exist. This paper discusses the opportunities and challenges associated with implementing data lakes, a potential strategy for leveraging data as a strategic asset for enterprise decision-making. The paper analyzes an information ecosystem of an Insurance Company environment. There are two types of data sources, information systems based on a transactional databases for recording claims, as the basis of financial administration and systems policies. There exists neither Data Warehouse solutions nor any other data collection solutions dedicated to utilizing by Data Science methods and tools. The emerging technologies provide opportunities for synergy between the traditional Data Warehouse and the most recent Data Lake approaches. Therefore, it seems feasible and reasonable to integrate these two architecture approaches to support data analytics on several aspects of insurance, financial activities, risk analysis, prediction and forecasting.

1 INTRODUCTION

The rise of Fintech, changing consumer behavior, and advanced technologies are disrupting equally all the financial services industry, among which also it's most prominent member, insurance. The insurance industry is preparing for the process of digital transformation and how it conducts business and big data capabilities lie at the centre of it. The insurance industry has been using data to calculate risks for years, still, with new technology now available to collect and analyze large volumes of data for patterns and better risk prediction and calculation, the value of understanding how to store and analyze it has grown exponentially (Liu et al., 2018).

Insurers are at their early stage of discovering the potential of big data, and multiple technology companies are investigate how to make value of such technology.

Equally important is to note that the majority of insurance companies are left with systems, processes, and practices that would be still recognizable by those that were in the industry in the 1980es. Changes brought by digitalization even more pressure the insurance industry (Traum, 2015; Pisoni, 2020).

Due to the big volume of data generated by insurers, it is especially important to enable the best use of them, and in this paper, we discuss opportunities and challenges in implementing data lakes for insurance companies and a potential strategy for leveraging data as a strategic asset for enterprise decision making, especially in the domain of customer behavior analytics and risk assessments. We present potential data lake reference architecture, it's respective data warehouse, and a mapping between the different components of the Data Lake to Zachman architecture. Also, we identify potential future challenges companies will face due to the use of big data, and we draw what in our view will be the new business logic of insurance companies in the digital era. The organization of the paper as follows: Section 2 an overview and analysis of architectures and applications of Data Warehouse and Data Lake, Section 3 investigates huge data collection and opportunities and limits for data analytics in the domain of customer behavior analytics, Section 4 analyzes how Data Lake and Data Warehouse can

[a] https://orcid.org/0000-0001-5015-8883
[b] https://orcid.org/0000-0002-3266-1773
[c] https://orcid.org/0000-0002-8159-2962

Molnár, B., Pisoni, G. and Tarcsi, Á.
Data Lakes for Insurance Industry: Exploring Challenges and Opportunities for Customer Behaviour Analytics, Risk Assessment, and Industry Adoption.
DOI: 10.5220/0009972301270134
In *Proceedings of the 17th International Joint Conference on e-Business and Telecommunications (ICETE 2020) - Volume 3: ICE-B*, pages 127-134
ISBN: 978-989-758-447-3

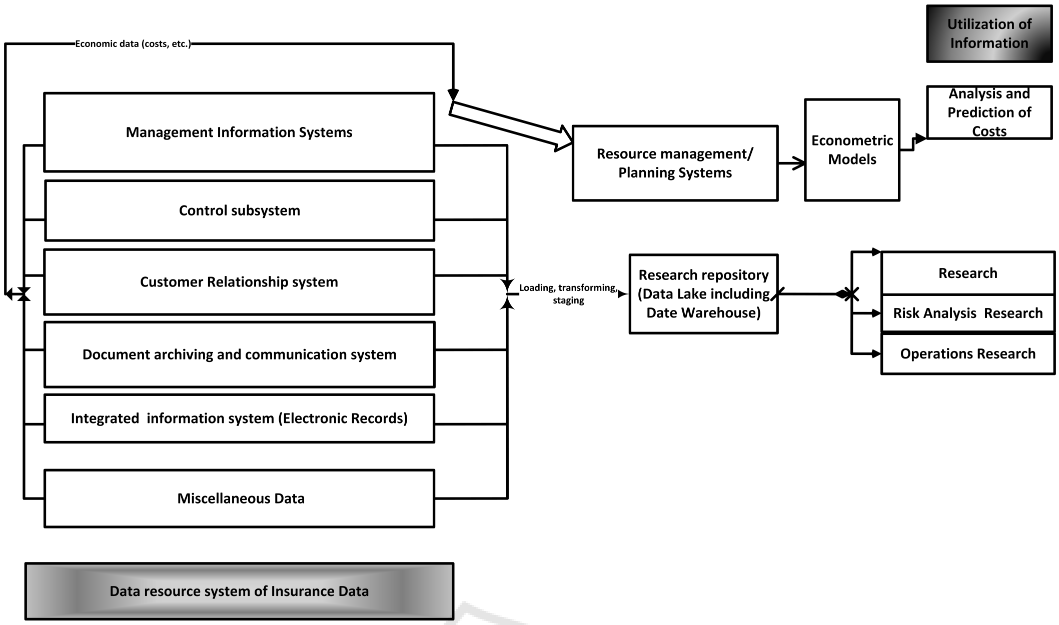
Utilization of Information
Economic data (costs, etc.)
Management Information Systems
Control subsystem
Customer Relationship system
Document archiving and communication system
Integrated information system (Electronic Records)
Miscellaneous Data
Data resource system of Insurance Data
Resource management/ Planning Systems
Econometric Models
Analysis and Prediction of Costs
Loading, transforming, staging
Research repository (Data Lake including Date Warehouse)
Research
Risk Analysis Research
Operations Research

Figure 1: Sources of data and usage, inspired by (LaPlante and Sharma, 2014; Boobier, 2016; Duggal et al., 2015).

help risk calculation, Section 5 examines the industry response and potential new logic for insurers of the future, and then the article concludes with a summary.

2 POTENTIAL APPLICATION DOMAINS

Potential domains for the use of the big data and data lakes in the insurance process include:

- *Pricing and underwriting*, estimating the price of an insurance policy is based on complex risk assessment process, big data can give the ability to more accurately price each customer by comparing individual behaviour compared to a large pool of data, a process that allows the insurance companies to correlate behaviour to risk (Boobier, 2016). This is most visible in the car insurance, where insurers can compare the behaviour of single driver to the behaviours of a large sample of other drivers, especially given the big number of past claims they have data for customers.
- *Settling claims*, a typical claim process starts with an insurer asking to assess the loss or the damage of the insured person or company, and this process can be long and painful for both of the parties. Different solutions have been devised already in this respect: automated claims setting, especially based on the domain of car insurance, where the company has a large amount of data, and already from past data can fully automatize claim report and their response to them. Insurance companies can execute and automate decisions related to big data, such as whether the insurance company should pay the claim or not risks, which can be later only approved by staff member instead of doing the process manually, which can also, in turn, save personnel time and effort (Minelli et al., 2013).
- *Monitoring of houses*, the insurance companies use and distribute IoT devices to monitor a range of activities at home. By comparing the real-time data collected from the delivers about the house risk, the insurance companies can intervene before a claim occurs, by recommending the policyholder to adjust the high-risk behavior, such as forgetting to lock the door or to set alarms. Having integrated data in this scenario is a key, because it allows insurance companies to see which policy-holders have good practices or which will almost close to certainly submit a claim, and like in future the company can reward good behavior of customers (Pal and Purushothaman, 2016).
- *Monitoring of health*, big data can be used to understand patient risks and prevent health issues before they arise. One way may be through drawing insights from big data to better make predictions and recommendations to customers for different insurance policies and their coverage, by connecting medical records and multiple sources or information regarding attitudes towards health like wearable of phone data. One example would

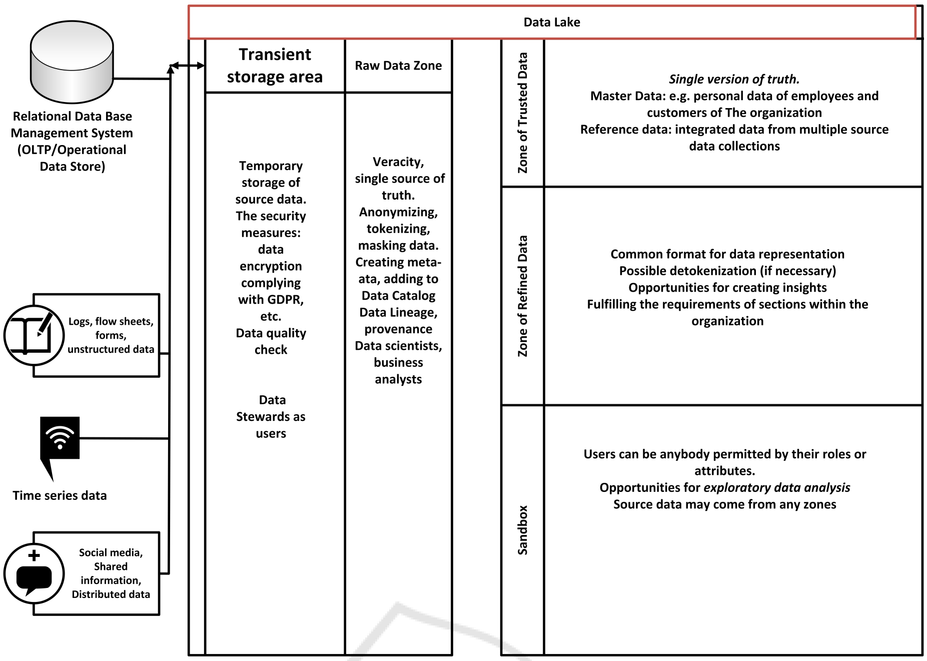
Data Lake
Relational Data Base Management System (OLTP/Operational Data Store)
Logs, flow sheets, forms, unstructured data
Time series data
Social media, Shared information, Distributed data
Transient storage area
Temporary storage of source data. The security measures: data encryption complying with GDPR, etc. Data quality check
Data Stewards as users
Raw Data Zone
Veracity, single source of truth. Anonymizing, tokenizing, masking data. Creating meta-ata, adding to Data Catalog Data Lineage, provenance Data scientists, business analysts
Zone of Trusted Data
Single version of truth.
Master Data: e.g. personal data of employees and customers of The organization
Reference data: integrated data from multiple source data collections
Zone of Refined Data
Common format for data representation
Possible detokenization (if necessary)
Opportunities for creating insights
Fulfilling the requirements of sections within the organization
Sandbox
Users can be anybody permitted by their roles or attributes.
Opportunities for exploratory data analysis
Source data may come from any zones

Figure 2: Data Lake Reference Architecture.

be a diabetic patient who needs insulin and if such intake is traced, the insurer would also be alerted such an irregular use (Spender et al., 2019).

- *Improve customer experience*, insurance companies have started to investigate the use of chatbots to allow for a more rapid response to customer queries and automatizing the response process. To implement the process, an AI algorithm needs to be trained on a big data and huge amount of data from the past on insurance policies, claims and other areas of business of the company, and as a result provide near to instance response to customer questions. The chatbots can be used also by new employees and even the staff can use and query the chatbots, which can be also voice-based, and like this to gain the information needed to serve the customer in front of them(Koetter et al., 2018; Riikkinen et al., 2018).

Other use can be for business process management aims (Rodrıguez et al., 2012), as well as for improved security of services (Ristov et al., 2012).

3 DATA LAKES AND DATA WAREHOUSE ARCHITECTURES

Generally, organizations are full of data that are stored in existing databases, produced by various information systems, data streams are coming from mobile applications, social media, web information systems, and other devices that are linked to the internet (Internet of Things, IoT). The collected data can be considered as heterogeneous in both structure and content, i.e. there are structured, semi-structured, and unstructured data items, some of them are accompanied by meta-data. There is an essential difference regarding data capture in traditional Business Information Systems and Information Systems for Financial Industries especially Insurance. In the Financial Industry, especially the Insurance environment, the initial data gathering is typically manual, and even the structured data contains plenty of textual information. The input data are created typically by human agents who can make errors in data either intentionally or unintentionally. Examples of error-prone data entry are as follows: name, address, identifiers (social security numbers). Thereby, the method of data entry leads to imprecise data that makes difficult any data mining exercises, e.g. association relationships that may have been found prove to be erroneous. While inputting data into Information System of Insurance Policies, Customers Data, frequently use either templates or copy-paste commands to fill in various forms as e.g. claims in forms to comply with policies, standards, and other regulations. The problem with that practice is that it conceals the subtle differences among the electronic customers' records that would be precious for data analytics and knowledge discovery. Information that is produced by automated systems along with

Table 1: Comparison of Data Warehouses vs Data Lakes.

-	**Data Warehouse**	**Data Lake**
Data	Structured, processed	Raw: Structured / semi-structured / unstructured
Processing	Schema-on-write	Schema-on-read
Storage	Requires large storage (architecturally complex)	Requires larger data storages (architecturally less complex). It can be cheaper despite the large amount of data stored.
Agile-aware	Fixed structure	Tailored
Purpose of Data	Fixed (BI, reporting)	Not Yet Determined (Machine Learning, Data analytics)
Target audience	Business Professionals	Data Scientist

auto-fill and edit check options as e.g. speech to text, OCR (optical character recognition) that computerize customers' data may introduce systematic and random inaccuracies that differ from clerks to clerks, a software tool to software tool. These errors are difficult to quantify and forestall.

To ensure that there will be a centralized location that would serve as the single source of truth, the data with different types and structures from various sources should be collected and loaded into a Data Lake. The most recent technologies can yield opportunities for the application of data analytics and models of Data Science. The results of running data analytic algorithms could be actionable knowledge in the clinical research environment.

Generally, it is assumed that the data coming from source systems are in good quality however, the market and administrative forces have not enforced a satisfyingly high level on standards of data quality. The typical life history of data can be seen in Figure 1. On the left part of the diagram, the various major source systems of financial data can be found. The data are represented as customers' electronic personal records, geo-codes for geographical information, and other loosely coupled data related to management, and business administration.

Our paper showcases architecture for a moderate size insurance enterprise environment so that the Vs (volume, velocity, variety, veracity, variability, value) of the Big Data are as follows: A population of a million customers may generate electronic customer records in the order of terabytes yearly. Primarily, the variability of structured, semi-structured, and unstructured data increases the complexity thereby the difficulty of ensuring the single point of truth within the data collection.

Data Lake as a Big Data analytics system allows the continuous collection of structured and unstructured data of an organization in the form of data representation without changing the original data, without data cleaning and transformation of data models - i.e. without loss of information, so the information can be analyzed with the greatest degree of freedom. Data Lake solutions focus primarily on integration and efficient data storage processes, besides providing advanced data management, data analytics, machine learning, and self-service Business Intelligence (BI) and data visualization services as well. A Data Lake offers services for Business Users using BI tools but the typical target audience is the data scientists. Data Lake is effective for an organization where a significant part of the organizational data is structured (and interpreted, stored in several, not yet reconciled sources), complemented by a large amount of unstructured data. The most important is that goal of the data processing is to utilize corporate data assets, exploring further new contexts, typically non-repetitive research questions. Our proposed Data Lake architecture consists of a Data Warehouse as well, to support or the information requirements of the organization (see Table 1). Within a hybrid Data Lake that contains a robust Data Warehouse as well, the life cycle of data commences with transformation, cleansing and integration. The objective to build up a Data Lake is to separate the daily operation, transactional data from the non-production data collection. Historically, the Data Warehouse technology has been employed for that purpose. During the data staging phase, the data are cleansed and filtered for the target data structure within the Data Warehouse, i.e. the fact table and dimensions. The phase of data staging includes data migration, data integration, translation of codes used for data representation, transformation between data base management systems. The Data Warehouse served as a basis for data analysis traditionally. The ETL (Extract, Transform, Load) procedure is applied for feeding data into the Data Warehouse. During that step the general data cleansing and transformation happens, e.g. amputation of trailing and leading white spaces, superfluous zeros, standardization of identifiers/identifying numbers, inflicting constraints on data fields, converting English units into metric units. While the before mentioned data-transformation is carried out relationships among entities may be dropped or harmed. Similar way, the data integration from multiple source systems, can lead to errors that are transmitted into the Data Warehouse

To overcome the data quality limitations of Data Warehouse, the idea of the Data Lake is conceptualized. Touted idea of Data Lake is that it deposits data in their original form, i.e. the *Transient*and/or *Raw Data* (see Figure 2 and Figure 3, (LaPlante and Sharma, 2014)) contains the data after an in-

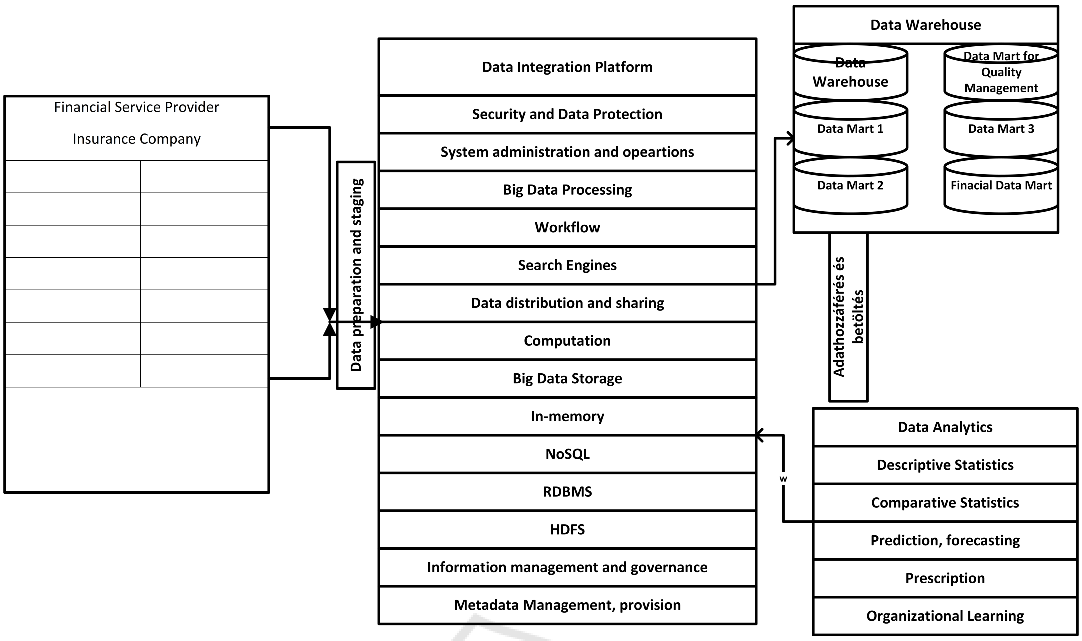

Figure 3: Fintech Data Lake Data Warehouse.

gestion phase so that the Data lake collects structured data from RDBMS (Relational Database Management Systems), semi-structured data (XML, binary XML, JSON, BSON etc.), and unstructured data along with meta-data that are represented typically in semi-structured format. The ingestion phase may mean loading, batch processing, data stream processing of source data while the necessary quality checks are carried out utilizing the MapReduce capability (Duggal et al., 2015). The essential property of the raw data zone is that it is considered as “a single source of truth” as it keeps the data in their original form, however the masquerading, and tokenizing of the data may happen in that zone. Data scientists, business/data analysts can return to that zone when they look for associations and relationships that may have lost during data conversion, transformation, encoding, and encrypting.

The Trusted Zone executes the procedures for data alteration as quality assurance, compliance with standards, data cleansing, and data validation. In that zone, several data transformations happen to correspond to prescribed local and global policies whereby the data can be considered as a “single version of the truth”. This zone may contain master data and the fact data that are governed through the data catalog that is filled in by meta-data automatically or semi-automatically. The data in the Refined Zone undergo several further changes that aim at the usability of data in algorithms of Data Science. These transformations include formatting, potential detokenization, data quality control to fulfill of requirements of algorithms so that models of the subject area (e.g. healthcare) and data analytics can be created. Thereby, knowledge discovery exercises can be carried out and understanding of data collections may be achieved. The access rights of users within each zone should be strictly maintained in the form of Role-Based Access Control, and for the temporary deviations from the baseline can be solved by Attribute-Based Access Rights. For researchers, managers, and other subject area experts who would like to pursue exploratory data analytics (Myatt, 2007), the Sandbox provides opportunities to set up models, discover associations and relationships among attributes without involving members of IT department and other extra costs. The researcher can bring data into the Sandbox from any other zones within a controlled environment. It is even allowed to export interesting results back to the raw data zone for re-use. The claimed advantage of the Data Lake is that the data extracted from the source systems are transformed before the actual use data for analysis. This approach permits more adaptability to requirements than the controlled, structured environment of Data Warehouses. (see Table 2, (Zachman, 1987)).

4 INSURANCE DATA AND CUSTOMER

The Financial Sector, especially, the Fintech companies are consumer-centric enterprises. The Insur-

Table 2: A mapping schematically between Zachman architecture and component of Data Lakes.

Aspects / Perspectives	***what***	***how***	***where***	***who***	***when***	***why***	***model view***
Contextual	Fact, business data / for analysis	Business Service	Business Intelligence, Workflow	Business entity, function	Chain of Business Process, Workflow	Business goal	Scope
Conceptual	Underlying Conceptual data model / Data Lake structured, semi- and unstructured data	Business Intelligence with added value originated	Workflow	Actor, Role	Business Process Model	Business Objective	Enterprise Model
Logical	Class hierarchy, Logical Data Model structured, semi-structured and unstructured data	Service Component Business Intelligence,Data Analytics	Hierarchy of Data Analytics Service Component	User role, service component	BPEL, BPMN, Orchestration	Business Rule	System Model
Physical	Object hierarchy, Data model (Object Data Store)	Service Component	Hierarchy of Service Component	Component, Object	Choreography	Rule Design	Technical Model.
Detail	Data in SQL/NoSQL, other file structures	Service Component Business Intelligence,Data Analytics	Hierarchy of Data Analytics Service Component	Component, Object	Choreography, Security architecture	Rule specification	Components
Functioning Enterprise	Data	Function	Network	Organization	Schedule	Strategy	

ance Industry that tries to follow the mainstream technology and not to lag put emphasis on data analytics and the application of Big Data related technologies. Traditionally the Financial Industry generally and the Insurance Industry especially are interested in the consumer groups that can be defined by age, gender, on-line activities, social and economic status, geographical distribution. In the Financial Industry, the enterprises customize the services and products to encounter the requirements of every customer fragment. Naturally, consumers are not treated equally, and those who have bigger affordances are traditionally being treated better. The identification and special treatment of these affluent consumer groups are important and the use of data lakes and data warehouses can significantly improve the process of *detection of different target groups*. Another of the benefits of exploitation of the Big Data technologies in the Financial Industry is to offer *fraud detection* approaches. The Insurance Industry gravitates towards on-line services on a large scale so there is a high chance to be the victim of a fraud. The Data Science assists the companies in the Financial Industry to understand their consumers behaviour, even the patterns of on-line actions. The key to fraud detection is in making use of more contextual data. Not just data about the immediate transaction and session, data about user is needed from his/her patterns of activities, biometric data about the person involved in the transaction, and background data on the user involved. Again, the advantage of use of data lakes is in the fast response with adequate information from the pool of data. In the Financial Industry, banking, payments, and insurance the concept of personal touch in services acquires importance and it becomes the most significant tool for marketing. The rivalry among the competing companies enforces the the enterprises to adopt such *personalized type of services*. In the Insurance Industry, the personal recommendation may include how to save money by combination of insurance policies, which combined insurance policy with investment has benefits for the consumer, etc. Once more, such data can be easily generated from the data lake of the company.

5 RISK CALCULATION

Risk management is an important task in all industries. In the Financial Industry, Data Science provides the opportunity to recognize the potential risks, as bad payers or incorrect decisions of investments. The application of data analytics methods cannot make the potential risk avoidable, however, it offers the identification of potential risks in a timely fashion. The techniques of Data Science can assist enterprises in the Financial Industry to adapt their company strategies and business processes to minimize the risks. It has recently become commercially viable to create risk profiles for individual customers, defined by factors such as age, gender, health, work activity, place of residence, driving behavior, etc., and the willingness to pay and make the categories corresponding individual offers. One of the frontiers in applying the data lakes will be exactly in this, and how to do it fast given that the data is transformed during analytics. Developing new, or more sophisticated, risk models can enable insurers to offer more competitive rates, or to offer insurance for previously uninsurable risks, due to information gaps which today are filled in by the increased

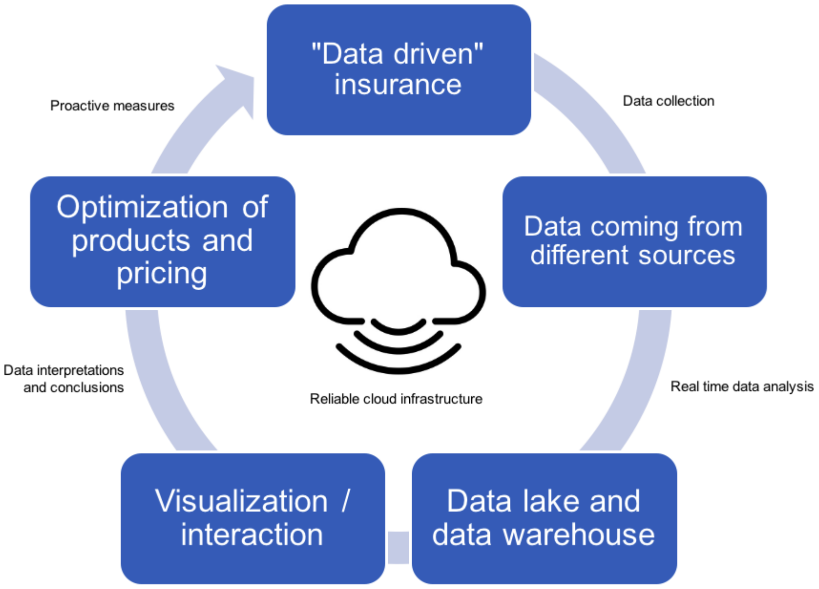
"Data driven" insurance
Data collection
Data coming from different sources
Real time data analysis
Data lake and data warehouse
Visualization / interaction
Data interpretations and conclusions
Optimization of products and pricing
Proactive measures
Reliable cloud infrastructure

Figure 4: Business logic of insurance company in the digital era.

availability of data. Data lakes in this regard make the work of a data analyst in a company much easier and the use of new analytical methods would be faster with the use of data lakes.

6 INSURANCE COMPANIES

Big data and new ways to elaborate data available from multiple sources brought to raise the possibility for the insurance companies to offer *usage-based insurance*(Thiesse and Köhler, 2008). For the moment it is only for niche customers, and as technology advances are significant, it is expected to significantly increase market penetration. It might also increase accessibility to insurance, and will allow to categories that previously did not insure to get coverage through the monitored coverage programs. New business models, based on a unique personal situation tied to the consumer will rise. Potential examples of the new scenario for which insurance companies neither think nor have an offer yet, especially in the context of usage-based insurance:

- Frequent air travelers, and people who travel often. Should travelers insure for all the flights or only for the flights that are likely to be canceled or delayed? Should this decision be made on information about weather, the company profile of canceled flights, or other events that may impact flight cancellations?
- Homeowners, and owners of personal possessions. How to insure a house while on a rental contract? What about Airbnb renting? In general for personal possessions, should companies insure only new possessions, or also old ones? Again, should companies insure the rents for a longer duration of time, or insure for single rents?
- Car insurance and drives that drive frequently or infrequently. Should there be insurance coverage for part-time driving, or based only on the numbers of km customers drive? Does the price that now customers currently pay truly take into account the driving history of the person or is determined based on the real risk of your car being stolen? If someone borrows the car for 2 days, should he have separate insurance, especially if the current insurance policy is based on the owner's behavioral profile? What about rides in Blablacar?

In general, the benefits for insurers will be also in reduced costs of underwriting and administration of insurance policies, and using the data lakes approach will have *better integration* with data from other industries, and significant improvement in the use and making sense of such data.

An important obstacle and limitation for the adoption of such approaches by insurance companies are to avoid risk prediction that is 'too thin' and leaves the company with big financial exposure to a potential unlikely event. The lower the cost of the coverage for high-risk events, the poorer the insurer will be. Future works will need to investigate how to build new risk models that will support usage-based insurance offerings.

In Figure 4 we present the basic structure of the "data-driven" insurers. The data collected will be analyzed in real-time in the data lake, visualizations will be presented to insurers, and the insurer can decide on optimization in offering and product or pricing. This information can be presented further in knowledge graphs (Tejero et al., 2020) and can be used for the generation of reports with new insights, the execution of advanced data analysis task between business.

7 CONCLUSIONS

In this paper we presented our first attempt to tell the ongoing problems in the insurance industry, what would be potential advantages and challenges for use of data lakes in the insurance industry, reflected how they will influence customer behavior analytics, and risk assessment, and lastly we discuss industry adoption and challenges that may arise. Besides, we present the basic structure of the digital insurance ecosystem, with data lakes and data warehouses being at the center of it.

Big data and new ways of integrating data in the digital transformation, such an integrated approach will foresee the development of new open business intelligence models, to better detect similar cases among data and stress similarity and explore more the use potential re-use of the same effort for different businesses.

The future of insurance will be data-driven and there will be the need to manage risk even if data makes it easier to estimate it. The pace of change however if big and modern data structures like data lakes and data warehouses will give rise to new business models for their functioning, that will change the way they work as of today.

ACKNOWLEDGEMENTS

The project was supported by no. ED_18-1-2019-0030 (Application-specific highly reliable IT solutions) program of the National Research, Development and Innovation Fund of Hungary, financed under the Thematic Excellence Programme funding scheme.

REFERENCES

Boobier, T. (2016). *Analytics for insurance: The real business of Big Data*. John Wiley & Sons.

Duggal, R., Khatri, S. K., and Shukla, B. (2015). Improving patient matching: single patient view for clinical decision support using big data analytics. In *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions)*, pages 1–6. IEEE.

Koetter, F., Blohm, M., Kochanowski, M., Goetzer, J., Graziotin, D., and Wagner, S. (2018). Motivations, classification and model trial of conversational agents for insurance companies. *arXiv preprint arXiv:1812.07339*.

LaPlante, A. and Sharma, B. (2014). *Architecting Data Lakes*. O'Reilly Media Sebastopol.

Liu, Y., Peng, J., and Yu, Z. (2018). Big data platform architecture under the background of financial technology: In the insurance industry as an example. In *Proceedings of the 2018 International Conference on Big Data Engineering and Technology*, pages 31–35.

Minelli, M., Chambers, M., and Dhiraj, A. (2013). *Big data, big analytics: emerging business intelligence and analytic trends for today's businesses*, volume 578. John Wiley & Sons.

Myatt, G. J. (2007). *Making sense of data: a practical guide to exploratory data analysis and data mining*. John Wiley & Sons.

Pal, A. and Purushothaman, B. (2016). *IoT technical challenges and solutions*. Artech House.

Pisoni, G. (2020). Going digital: case study of an italian insurance company. *Journal of Business Strategy*.

Riikkinen, M., Saarijärvi, H., Sarlin, P., and Lähteenmäki, I. (2018). Using artificial intelligence to create value in insurance. *International Journal of Bank Marketing*.

Ristov, S., Gusev, M., and Kostoska, M. (2012). Cloud computing security in business information systems. *arXiv preprint arXiv:1204.1140*.

Rodrıguez, C. et al. (2012). Eventifier: Extracting process execution logs from operational databases. *Proceedings of the demonstration track of BPM*, 940:17–22.

Spender, A. et al. (2019). Wearables and the internet of things: Considerations for the life and health insurance industry. *British Actuarial Journal*, 24.

Tejero, A., Rodriguez-Doncel, V., and Pau, I. (2020). Knowledge graphs for innovation ecosystems. *arXiv preprint arXiv:2001.08615*.

Thiesse, F. and Köhler, M. (2008). An analysis of usage-based pricing policies for smart products. *Electronic Markets*, 18(3):232–241.

Traum, A. B. (2015). Sharing risk in the sharing economy: Insurance regulation in the age of uber. *Cardozo Pub. L. Pol'y & Ethics J.*, 14:511.

Zachman, J. A. (1987). A framework for information systems architecture. *IBM systems journal*, 26(3):276–292.