

EU-U.S. INSURANCE DIALOGUE PROJECT

BIG DATA ISSUE PAPER

The EU-U.S. Insurance Dialogue Project (EU-U.S. Project) began in early 2012, as an initiative by the European Commission, the European Insurance and Occupational Pensions Authority (EIOPA), the Federal Insurance Office of the U.S. Department of Treasury (FIO), and the National Association of Insurance Commissioners (NAIC) to enhance mutual understanding and cooperation between the European Union (EU) and the United States for the benefit of insurance consumers, business opportunity, and effective supervision. In 2018, the EU-U.S. Project's members continued the work focusing on the use of big data besides the other focus areas relating to cybersecurity risk, the cyber insurance market and intra-group transactions.

I. Introduction

The Big Data working group of the EU-U.S. Project is focusing on the relationship between innovation, technology and insurance, specifically the increasing use of advanced data analytics in the insurance sector. The use of new technology and large datasets in the pricing and underwriting process presents a series of potential benefits for insurers and consumers, such as more granular segmentation of risks and risk-adequate pricing, increased objectivity in the underwriting process, and a reduction of the overall underwriting costs. However, there are concerns that the use of new technology and large datasets in pricing and underwriting may potentially lead to excessive segmentation of the risks, unfair/unlawful price discrimination, overreliance on third party vendors, and the inability to verify data accuracy.

This paper will not specifically focus on the benefits and/or concerns of the use of big data. Rather, the paper will focus on providing the reader with a better understanding on the (1) type, quality and means for collecting big data and how this data is then used by both insurers and third parties for underwriting, rating, marketing and claims handling in both the U.S. and the EU and (2) how U.S. and EU supervisors' are addressing their data needs to appropriately monitor the insurance marketplace and evaluate underwriting, rating, claims, and marketing practices in their respective markets given the increased use of big data in the market place.

II. Discussion of what data is collected, how it is collected, data portability, data quality and how it is made available and used by both insurers and third parties in the context of marketing, rating, underwriting, and claims handling.

For insurance purposes, big data refers to the collection of structured data (referring to data in tables and defined fields) and unstructured data (most data referring to things such as social media postings, typed reports and recorded interviews) being used for underwriting, rating, pricing, forms, marketing and claims handling. The data used may include personal data of customers from the companies' own internal database, individual data from internet of things (IoT), e.g. vehicle telematics, cellphones and credit

scoring data from external providers including software.¹ Data collection processes include a number of techniques, including data mining, statistical modeling and machine learning. It should be noted that given the complexity of the algorithms used, and the breadth of data being collected, data quality and accuracy may vary by data sources.

U.S. and EU insurance supervisors continue to expand their knowledge about the availability and use of big data by insurers, which have been reluctant to publicly share this information because of claimed, competitive trade secrets. At the same time, U.S. and EU insurance supervisors have seen and are aware of the penetration of new data-driven technologies in almost every segment of the insurance value chain, particularly as predictive analytics allows insurers to use big data to forecast future events. Insurers collect data from different sources and use them for underwriting optimization, risk selection, technical pricing, product innovations, innovation and automation of processes in interacting with the customer and for own purposes, e.g. management of solvency, capital and liquidity.

Based on the data collected by third party providers or insurers itself, insurers have started to build their own big data and Artificial Intelligence (AI) systems, e.g. to produce behavior forecasts. Using big data and AI is considered a significant competitive advantage providing important insights into the personal behavior of consumers, requiring the implementation of strong data protection systems. While the level of adoption of fully developed techniques involving big data and, particularly, AI is still limited, utilization is expected to expand in the years to come.

For example, through the use of telematics, insurers track mileage and driving behaviors using odometer readings or in-vehicle telecommunication devices that are usually self-installed into a special vehicle port or already integrated in original equipment installed by car manufacturers. These devices provide a significant amount of data for insurers including miles driven, time of vehicle operation, location of vehicle operation (GPS), acceleration speed, braking, cornering, and air bag deployment. The level of data collected generally reflects the telematics technology employed and the policyholders' willingness to share personal data. This allows insurers to collect real-time driver behavior data and combine it with premium and loss data to provide premium discounts for insureds who drive shorter distances at slower speeds than insureds who drive longer distances at higher speed.

Nevertheless, questions relating to data accuracy and quality remain. Insurers should be aware that cell-phone based telematics programs may record trips when the insured was a passenger in someone else's vehicle and of potential gaps in data if cell phone reception is less than perfect. While it could seem appropriate that consumers should be rated on how they drive, questions remain as to whether consumers really understand how certain driving behaviors are affecting their rates in these programs because insurers do not necessarily disclose to consumers how certain behaviors are weighted.

¹ For further examples of types of data used in insurance, see a recent Geneva Association report, *Big Data and Insurance: Implications for Innovation, Competition and Privacy*, (Zurich, Switzerland 2018), 36.
<https://www.genevaassociation.org/research-topics/cyber-and-innovation/big-data-and-insurance-implications-innovation-competition-and>

Insurers may also use data to improve marketing and product development effectiveness by tailoring products to individual preferences. For example, an insurer may compile and use publicly available information such as home valuation, location, and internet usage to create consumer profiles for targeted marketing efforts. This may be considered a continuation of commercial transactions in which a merchant knows its buyers and decides what to offer a buyer based on this knowledge. A key distinction with data and predictive analytics is that an insurance company may collect data without the consumer's knowledge or that consumers do not fully understand the potential reach and consequences of the use of their data. This may lead to insurance companies using data to influence a consumer's decision or not offering products to certain segments of the market. While insurance companies increase efficiencies around distribution, insurers should continue to be held to a standard of delivering information that will improve a consumer's ability to select a product that meets their needs. The availability of an insurance product should not be limited to a segment of the population if this limitation is based on unfair/unlawful discrimination. An open, competitive and fair marketplace helps mitigate this concern.

Data analytics can also be used to enhance the speed in which claims are settled and reduce fraud through the identification of relationships that would go undetected. For example, text analytics can identify potential "red flag" trends across adjusters' reports. However, supervisors will need to more fully understand and analyze whether evolving uses of big data in claims settlement (for instance the use of big data in predicting what settlement offer a specific individual may be willing to accept) are considered fair claims settlement practices.

In the United States, the NAIC's Big Data (EX) Working Group is discussing marketplace activities surrounding the use of big data to obtain a clear understanding of what data is collected, how it is collected and how it is used by insurers in the context of marketing, rating, underwriting, and claims. The Working Group began their work in 2016 by holding public hearings on the use of data and the perspectives of supervisors, consumers, and industry. The Working Group has primarily focused on the use of complex rating models used for personal lines property and casualty insurance. In 2018, the Working Group is shifting its attention to the use of data for underwriting of life insurance products. The Working Group will use their findings to determine what appropriate changes, if any, should be made to the current regulatory framework used to oversee insurers' use of consumer and non-insurance data. For example, U.S. state insurance supervisors recently addressed the issue of price optimization through the NAIC Casualty Actuarial and Statistical (C) Task Force which drafted a White Paper analyzing price optimization and its use in insurance ratemaking, with a primary focus on personal lines². A number of states subsequently issued notices prohibiting or restricting the use of price optimization or the concept of rating based on price elasticity in personal lines indicating that price optimization results in rates that are unfairly discriminatory.

Insurers and insurance producers must protect the highly sensitive consumer financial and health information collected as part of the underwriting and claims processes. This Personally Identifiable Information (PII) is entrusted to the industry by the public. Pursuant to the federal Gramm-Leach-Bliley Act, state insurance supervisors have also adopted model laws for protecting consumer data. In 2000, the

² NAIC Casualty Actuarial and Statistical (C) Task Force, *Price Optimization White Paper*, (2015).
https://www.naic.org/documents/committees_c_catf_related_price_optimization_white_paper.pdf

NAIC adopted the Privacy of Consumer Financial and Health Information Model Regulation to protect the privacy of insurance consumers' personal information. The model regulation provides protection for financial and health information about consumers held by insurance companies, agents, and other entities engaged in insurance activities. In general, the model regulation requires insurers to: (1) notify consumers about their privacy policies; (2) give consumers the opportunity to prohibit the sharing of their protected financial information with non-affiliated third parties; and (3) obtain affirmative consent from consumers before sharing the protected health information with any other parties, affiliates and non-affiliates. In 2002, the NAIC adopted Standards for Safeguarding Customer Information Model Regulation, establishing standards for developing and implementing administrative, technical and physical safeguards to protect security, confidentiality and integrity of customer information. Recognizing privacy issues are applicable across all industries, some states have adopted general consumer data privacy and protection laws.³

US state insurance supervisors have traditionally completed a review of insurance company information security programs during a routine financial examination. With the rising prominence of this risk exposure, regulators have taken the time to enhance the supervisor work program to make sure it appropriately supports a thorough review of such security programs. The NAIC Financial Condition Examiners Handbook provides guidance that supervisors use as part of the financial examination process and includes a review of whether and how the insurer is addressing its cyber risk. Guidance is based both on the COBIT 5 framework and the NIST Cybersecurity Framework.

The NAIC also completed the NAIC Roadmap for Cybersecurity Consumer Protections and the NAIC Insurance Data Security Model Law (#668). The Model Law requires insurers to implement an information security program and investigate and notify the state insurance commissioner of cybersecurity events⁴. All states have consumer data breach notification laws. In addition, the NAIC adopted a Cybersecurity Insurance Coverage Supplement for the P/C annual financial statement to collect information about US cybersecurity insurance markets. The NAIC is also considering creating a Cybersecurity Insurance Institute and an Anti-Fraud Depository.

An emerging consumer issue related to privacy and data sharing notifications is data portability. U.S. state insurance supervisors are aware of this important consumer issue as the behavior of consumers are tracked and captured through new data sources and utilized by the insurance industry. Arriving at the appropriate policy direction is complex and dependent upon multiple factors, including data ownership, balancing cybersecurity data protection with consumer control, and the use of varying technology platforms across companies. It is still uncertain whether the insurance industry will move toward data

³ See, e.g., Calif. Civil Code 1798.100 et seq., http://leginfo.ca.gov/faces/codes_displayText.xhtml?lawCode=CIV&division=3.&title=1.81.5.&part=4.&chapter=&article;; Colorado House Bill 18-1128, signed by Governor May 29, 2018, https://leg.colorado.gov/sites/default/files/documents/2018A/bills/2018a_1128_signed.pdf

⁴ Effective March 2017 the New York Department of Financial Services promulgated a regulation establishing Cybersecurity Requirements for Financial Services Companies, ,Cybersecurity Requirements for Financial Services Companies, 23 *NYCRR* 500, <https://www.dfs.ny.gov/legal/regulations/adoptions/dfsrf500txt.pdf>

portability standards or if an insurance regulatory standard or broader governmental standard will be deemed necessary. Additional discussion among industry, supervisors, consumers, and data vendors is still needed to arrive at a policy direction and standard that properly addresses consumer data portability and protects insurer trade secrets, while considering the potential consequences of increased costs and constraints on innovation.

In the EU, a number of initiatives have been undertaken to review the use of big data and potential implications for the financial industry, consumers and supervisors. The European Supervisory Authorities (EIOPA, EBA and ESMA) conducted a joint cross-sectoral review on the use of big data and published a report with key findings in February 2018⁵. Also in 2018, EIOPA set up a multi-disciplinary InsurTech Task Force⁶ and a thematic review is under way on the use of big data specifically by insurance undertakings and intermediaries with the objective of gathering empirical evidence on the impact of big data across the insurance value chain⁷. This thematic review will assess how big data affects the financial inclusion and/or exclusion of consumers; risks related to a fair treatment of consumers as a result of enhanced consumer profiling techniques and more granular risk assessments; as well as other risks and benefits arising from big data. The thematic review will also collect detailed information on what types of data, data sources and innovative tools are being used by insurance undertakings and intermediaries in the context of big data. At the national level in France, Germany and the UK⁸ studies or surveys on AI and big data use are currently being conducted and in some cases, first findings have already been published. In addition, supervisory authorities in Europe are in discussions with the industry on this topic and in some cases complement this with additional pricing and underwriting reviews.

In the EU the new General Data Protection Regulation (GDPR)⁹ requires of EU firms (i.e. not only insurance undertakings) increased transparency and creates new rights for consumers, additional records, application of enhanced security measures, compliance checks and impact assessments. Similar to the also recently approved Insurance Distribution Directive (IDD)¹⁰, the GDPR recognizes the overarching principle of fair treatment of consumers when it comes to the processing of personal data. Furthermore, the new rules enable consumers to demand the removal of their data from insurers' databases (commonly known as "right to be forgotten"). Finally, the GDPR also recognizes consumers' right to receive the personal data concerning him or her and to transmit that data to another provider (commonly known as "right to data portability").

⁵ Joint Committee Final Report on Big Data, 15 February 2018, <https://esas-joint-committee.europa.eu/Pages/News/ESAs-weigh-benefits-and-risks-of-Big-Data.aspx>

⁶ EIOPA InsurTech Task Force mandate, 13 April 2018, <https://eiopa.europa.eu/Pages/Working%20Groups/InsurTech-Task-Force.aspx>

⁷ EIOPA Big Data thematic review press release, 6 June 2018, <https://eiopa.europa.eu/Pages/News/EIOPA-seeks-evidence-on-the-use-of-Big-Data.aspx>

⁸ In the UK the study only covered Big Data and not AI

⁹ General Data Protection Regulation (GDPR), 27 April 2016, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679>

¹⁰ Insurance Distribution Directive, 20 January 2016, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L0097>

In addition to the GDPR, the Solvency II insurance regulatory framework¹¹ in Europe also foresees a number of provisions regarding the governance arrangements that firms need to have in place when dealing with big data, which are then complemented at the national level. For example in Germany the supervisor published its circular concerning supervisory requirements for IT in the insurance sector (Versicherungsaufsichtliche Anforderungen an die IT – VAIT). The VAIT are set to become the cornerstone of IT supervision for all insurance undertakings and Pensionsfonds in Germany. VAIT primarily targets the level of senior management. The objective of the VAIT is to provide the senior management of insurance undertakings with a clear and flexible framework, particularly in relation to IT resource management, information risk management and information security management. The VAIT are also intended to help increase awareness of IT risks in insurance undertakings and in relation to their IT service providers. The VAIT clarify what BaFin expects from these undertakings with regard to the management and control of IT operations, including the required access rights management. In addition, the VAIT lay down the requirements for IT project management and application development, which also encompasses end-user computing in business units. Overall, the VAIT address all the issues that BaFin considers to be particularly significant based on the findings of its IT supervisory activities and inspections.

III. Supervisors' data needs to appropriately monitor the insurance marketplace and evaluate underwriting, rating, claims, and marketing practices.

Throughout the past few years, there has been a noticeable change in the insurance marketplace, driven by consumer expectations and competitive pressures, as well as the growth of innovation and emerging technology.¹²

Regulatory frameworks need to cope with the speed of innovation through agile oversight using technology-neutral and principle-based approaches. By most accounts, the pace of these changes in the marketplace is expected to accelerate in the foreseeable future bringing with it considerable changes to supervisors' data needs to appropriately monitor the market place. As rating algorithms become more complex and insurers are increasingly challenged with filing complete and accurate algorithms, it will be increasingly important for supervisors to have the tools to assess the data being used and to confirm the consistent result/output of the algorithms used for underwriting, rating, claims and marketing practices. Supervisors will need the appropriate authority and skills to review the use of new data variables, the method of data collection, data completeness, accuracy, timeliness, stability, and business area of use. There is also a necessity to understand the interaction of big data technologies with regulations and to understand if the latter have implications as to how certain big data technologies should be used. Strengthening international cooperation is also important, because changes are often triggered by global and potentially less strictly regulated companies through their use of big data.

¹¹ Solvency II Directive, 25 November 2009, <https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32009L0138>

¹² US Department of Treasury has addressed Fintech and InsurTech generally in a recent report. See U.S. Department of Treasury, *A Financial System that Creates Economic Opportunities: Nonbank Financials, Fintech and Innovation* (July 2018), https://home.treasury.gov/sites/default/files/2018-08/A-Financial-System-that-Creates-Economic-Opportunities---Nonbank-Financials-Fintech-and-Innovation_0.pdf

To address these concerns, in the U.S. state insurance supervisors are continuously pursuing opportunities to enhance current regulatory market-related data, tools and processes including assessing the appropriate skills and resources required in conducting their review of insurers' predictive models to protect consumers in the insurance marketplace. U.S. state insurance supervisors, through the NAIC Casualty Actuarial and Statistical (C) Task Force are (1) drafting proposed changes to the NAIC Product Filing Review Handbook to include best practices for the review of predictive models and analytics filed by insurers to justify rates; (2) drafting and proposing state guidance for rate filings that are based on complex predictive models; and (3) facilitating training and the sharing of expertise through predictive analytics webinars. These efforts will benefit both insurance supervisors and industry by enhancing the collective knowledge of insurance supervisors and providing for more uniform and efficient reviews of insurers' rate filings that include complex predictive models.

The NAIC in 2018 adopted a State Ahead strategic plan in recognition of the evolving marketplace fueled by shifts in consumer behavior and technological advances. The strategic plan specifically recognizes the need for insurance supervisors to promote stable financial markets by optimizing data collection and analysis.

For decades, state insurance supervisors have collected a significant amount of electronic financial data from U.S. insurers to support solvency oversight, financial analysis and examinations in particular, of legal entity insurers and insurance groups. The NAIC utilizes this data in automated analysis and examination tools for supervisors to use in these oversight activities. More recently, macro prudential oversight has increased supervisors' needs for data and analytics. To meet these needs and as part of the NAIC State Ahead initiative, life insurers will begin reporting more data by product category in their annual statutory financial statements as of year-end 2019. The NAIC will also collect mortality experience data from life insurers as part of the full conversion to Principle-Based Reserving in 2020. However, standardized data collection does take lead time for setup and rollout, so the NAIC is also enhancing tools to make supervisors' data calls and surveys, used for more immediate and/or short-term data needs, more streamlined and standardized, including adding data quality features. To make better use of these and other data sources for state insurance supervisors, the NAIC's State Ahead Initiative is also developing a new data platform and business intelligence framework to allow more self-service analytics and facilitate more sophisticated predictive analytics and AI projects to support market, solvency and macro prudential surveillance needs. These initiatives are already impacting existing analysis and examination tools, enhancing functionality with additional calculations and visualizations; and as State Ahead matures through 2020, the NAIC will experiment with an AI solvency tool to complement these tools. Supervisors need innovation in their technology and processes to be effective today and going forward, and the NAIC's State Ahead initiative will ensure U.S. state insurance supervisors are able to meet the challenges of innovation in the industry.

Further, state insurance departments have begun collecting market data from insurers through filings of the NAIC's Market Conduct Annual Statement (MCAS) leveraging NAIC technology to collect and store more market data (focusing on optimizing the collection and use of existing market data). MCAS provides market supervisors with market conduct information not otherwise available for their market analysis initiatives. Companies report claims and underwriting data to each jurisdiction through the NAIC's on-line MCAS application. For the 2017 data year, MCAS data is collected for the following lines of business: Health, Homeowners, Individual Annuity, Individual Life, Long-Term Care, Private Passenger Auto. The MCAS and the centralization of MCAS data have given insurance supervisors another tool to protect insurance consumers and claimants. Through the cooperation of insurance supervisors and the insurance

industry, the MCAS will continue to grow and promises to be the primary source of market conduct data for use in effective market analysis.

In addition, as part of an enterprise approach in support of insurance supervisors, the NAIC is creating a data analytics warehouse which will allow state insurance regulators to more easily identify data across all NAIC functions. This effort will also create a market data lifecycle and governance framework, which will lead to an enterprise approach to retire obsolete market data collection systems and centralize data collection to a modern and flexible platform. Another key to accomplishing this goal is the implementation of self-service business intelligence tools. This will allow insurance supervisors to drill down from high-level information into detailed records, creating and modifying filters to highlight a regulatory issue or concern, and building visualizations of data to see trends and anomalies. Lastly, cloud capabilities being built for the NAIC will enable insurance supervisors to explore opportunities with big data.

In the EU, EIOPA and the national insurance supervisory authorities are currently in the process of reviewing the use of big data/AI and the potential implications for insurance supervisors. The new Solvency II reporting framework has provided European insurance supervisors with a comprehensive, harmonized and high-quality prudential data about the activities of insurance undertakings, which is complemented with the conduct of business reporting frameworks existing at national level. EIOPA's Big Data thematic review is collecting the views of national supervisory authorities on whether these reporting frameworks are sufficient to allow them to adequately perform their supervisory tasks in view of the new technological developments. Furthermore, the mandate of EIOPA's InsurTech Task Force explicitly foresees the promotion of convergence on supervision of algorithms amongst supervisory authorities, which among other things will explore the need for capacity-building activities for supervisors such as workshops and seminars. EIOPA's InsurTech Task Force is also mandated to assess how big data and other innovative data-analytical tools can be used for supervisory purposes in order to capitalize on the new data-reporting requirements introduced by recent European financial services legislation such as Solvency II, MiFID II / MiFIR and EMIR. Indeed, supervisory systems in Europe increasingly make use of modern technology, for example, to enable supervisors to identify unregulated risk takers at an early stage and react quickly.

Supervisors in the EU consider as a pre-condition for appropriate supervision that the auditability and transparency of the data and/or algorithms used is ensured; black boxes which do not allow supervisors to fulfill their supervisory tasks should be avoided, and the reasoning behind the statistical results and complex algorithms should be explainable. Important areas of supervisory attention are increasing model risks arising from BD use and advanced analytics, appropriate documentation of policies and processes and clear allocation of management responsibilities. In addition, EU supervisors are considering the impact on the regulatory perimeter and potential risks resulting from outsourcing of critical IT functions to unregulated Cloud Service Providers. The dependence of insurers on such unregulated external service providers could also lead to operational resilience issues in the event of failure. Moreover, insurers could also become dependent on external data providers, such as large (mostly unregulated) tech firms.

Supervisory authorities in the EU are concerned that the increased role of technology may lead to greater interconnectedness across the market and the potential for common points of failure. Although in Europe monitoring the implementation of data protection regulations is primarily the responsibility of the data protection authorities, if abuse occurs then the financial supervisory authority can also be called upon. Indeed, the GDPR imposes requirements which are compatible with the mandate of insurance

supervisors; EU supervisors will do further work on the precise implementation and interpretation of several GDPR requirements as they pertain to big data/AI applications in the financial/insurance industry, for instance regarding the degree of granularity insurers have to comply with when it comes to explaining results to consumers of AI-models.

V. Conclusion and Next Steps for 2019

The Big Data Working Group has identified some common insurance industry practices across the U.S. and EU markets. While insurance supervisors want to encourage innovation to meet consumer demands, the Working Group also recognizes necessary consumer protections must remain in place. With a sound regulatory standard firmly established and responsive to both the insurance industry and consumers, insurance supervisors in both the U.S. and EU will continue to apply existing standards, monitor how the insurance industry is using big data and AI, understand consumer expectations, and, ultimately decide if regulatory standards and practices may need to change as the insurance marketplace changes. With this, the Big Data Working Group in 2019 will continue to discuss opportunities and challenges relating to these issues and recommends looking further into the following areas:

1. Further exploration of Third-Party Vendor Issues. Discuss current regulatory oversight of insurer use of third-party vendors and whether or how this framework focuses on issues surrounding big data accuracy concerns, and the extent to which the current regulatory perimeter is addressing the ability for regulatory oversight.
2. Discuss disclosures to applicants and policyholders specifically about how rating factors and third-party reports are being used and the opportunities for applicants and policyholders to adopt informed decisions and correct potential errors, while at the same time respecting insurer's intellectual property rights (see use of credit score reports in U.S./EU as recent example).
3. Discuss insurers' use of AI-models, including governance, data quality and how to reduce the opacity of complex ('blackbox') models, for instance through human-intervention, implementation of explanatory models (e.g. LIME-models), or potential restrictions on in- and outputs.