Advanced Analytics in Central Banks: Basic Assumptions and Preliminary Results of a Research Project

Matthias Goeken and Leni Hirdes
University of Applied Sciences of Deutsche Bundesbank, Hachenburg, Germany

Keywords: Central Banking, Advanced Analytics, Systematic Review.

Abstract: This position paper presents methodological considerations and ideas as well as preliminary results of a research project. We argue that integrating design science (DS) and review research is useful for developing artifacts that serve, for example, to share knowledge within a domain. Our project aims to support knowledge sharing in central banking by means of systematic reviews. Such a consolidated body of knowledge can also be considered an artifact in the sense of DS research. We demonstrate how such a review can be conducted and what results it yields. Since it is not clear whether such an application of systematic reviews is relevant and feasible, problem definition and objectives of the project are discussed.

1 INTRODUCTION

Central banks are knowledge-intensive organizations. To fulfill their mandate, a critical task is to observe and monitor markets and the economic environment, to consider economic policy measures, to make forecasts and to draw conclusions for potential actions. To do so, they traditionally rely on statistics and econometrics in connection with economic model building. These approaches and methods are well established and have been successfully applied in the domain of central banking for many years.

In fact, central banks tend not to contest with other organizations or with each other. Therefore, unlike many players in the financial industry, they are not dependent on gaining competitive advantages e.g., by means of (technical) innovations. Nevertheless, they are interested in expanding their analytical competencies in order to be able to best fulfil their mandate. And, like other players in the financial industry, they strive to do so by adopting innovative and advanced technologies and approaches, e.g. big data, machine learning, AI, and data science (Bholat, 2015).[1]

Seven years ago, a survey of central banks by the Irving Fisher Committee (IFC, 2017) found that there was already strong interest in big data, machine learning and related topics in the central banking community. In addition to the general trends frequently noted for the economy as a whole (data availability, availability of superior computational power, and more mature methods and procedures) and the related general desire to improve analytics capabilities in traditional areas, there are specific drivers that are pertinent to the growing interest of central banks. E.g., the Great Financial Crisis (GFC) of 2007-09 laid bare the necessity of more disaggregated data (Doerr et al. 2021). In this respect, macroprudential policy has been established as a new field of activity alongside monetary policy and banking supervision. This is where the improved use of granular data on individual economies for financial stability analysis is a key task (Buch, 2019). Figure 1 illustrates the growing interest among central banks in recent years. It displays that the number of central bank speeches mentioning “advanced analytics” and similar related topics increased considerably from 2015 to 2020.

However, actual use by central banks has so far remained limited. The IFC Annual Report 2017 identifies three reasons for this: first, a number of operational challenges, in particular the availability of significant (technical and human) resources to deal with big data. Second, making adequate arrangements for the management of data and information for

[1] In this paper we use “advanced analytics” as an umbrella term to refer to big data (BD), machine learning (ML), data science (DS), and artificial intelligence (AI). As the major parts of the paper are about these approaches in general, a clear distinction is not necessary in most parts of our paper.

Goeken, M. and Hirdes, L.
Advanced Analytics in Central Banks: Basic Assumptions and Preliminary Results of a Research Project.
DOI: 10.5220/0011306800003280
In *Proceedings of the 19th International Conference on Smart Business Technologies (ICSBT 2022)*, pages 137-143
ISBN: 978-989-758-587-6; ISSN: 2184-772X

central banks; and third, that the use of big data for policy purposes is not without risks, such as creating a false sense of security and accuracy among the audience or the public (IFC, 2017).

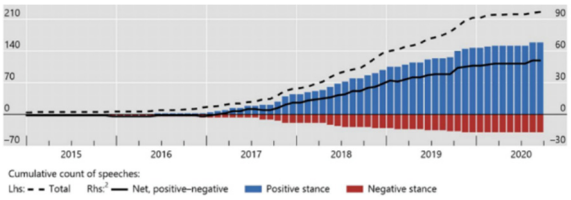

Figure 1: Central banks' interest in big data (Doerr et al., 2021).

These and other problem areas open room for improvement. One sensible idea is to improve knowledge sharing and avoid reinventing the wheel multiple times. Hence, our approach is to summarize and evaluate certain works, where experiences from projects, case studies and applications are already available. In this way, exchange of experience could be strengthened – beyond the usual formats of central bank cooperation.

For the purpose of consolidating existing and published results, so-called systematic reviews have become increasingly accepted as effective instruments in the information systems discipline in recent years (Webster & Watson, 2002; Wainwright et al., 2018). Reviews are seen as powerful instruments to enable the synthesis of prior findings, reconciliation of inconsistent findings, and resolution of relationships. In this realm, reference is also made to the practical relevance of reviews.

For many reviews, the motivation is already given by the fact that primary research exists. Here we are faced with a domain for which it is not entirely obvious whether a review that consolidates the body of knowledge is meaningful, relevant, and feasible. For this reason, the combination of review research and design science seems appropriate.

In this paper, we first outline the rational and the motivation for (part 2) and the methodological background of our research project (part 3). Part 4 outlines first and priliminary results. We finish with conclusions and and give hints for future research.

2 RELATED WORK

AI, Data Science and related topics have already been the subject of review articles. Collins et al. (2021) examine the extent to which primary studies identify business value of IT through AI and specific contributions made by AI, and they consider the scientific and practical implications of using AI in organizations. Abdel-Karim et al. (2021) investigate in their review, why the "transfer of theoretical knowledge from machine learning to applications that solve (industry-)specific problems" is only partly successful. They do so with the goal of advancing IS theory as well as to help practitioners in applying ML methods. Other reviews explore the practical understanding of AI to better identify areas of application and improve adoption (Bawack et al. 2019).

Likewise, papers appear that investigate the "state of the art" with a focus on methodological issues of advanced analytics (Angra & Ahuja, 2017). Some high-level reviews study the application of ML and AI in general by (Warin & Stojkov, 2021). Occasionally, but more rarely, there seem to be works that examine the application of advanced analytics in relation to a specific technical topic or domain. Our – so far non-systematic – literature review e.g. identified reviews on the application of ML and AI in supply chain management or in the context of manufacturing (Fahle et al. 2020; Younis et al., 2021). However, such works are rather rare. It seems that the focus on a domain or specific cases is much more common in other disciplines, e.g. in medicine, probably, because the benefits are more obvious.

We go with Abdel-Karim et al. (2021), who note a low consideration of ML in the IS discipline. However, like them, we assume that after remarkable progress has been made in Computer Science from a technical and methodological point of view, it is a fundamental task of IS research to investigate the transfer into companies and organizations and to support it with appropriate methods and artifacts. Hence, we focus on the transfer of theoretical knowledge regarding advanced analytics to applications. Accordingly, the focus is on sharing knowledge for a practical purpose and action, which is not the primary goal of existing reviews.

Since we can identify a concrete and relevant use case in the central banking domain (see above), our research project aims at the idea of conducting a review of the application of advanced analytics for this rather narrowly defined field. Last year, Kinywamaghana & Steffen (2021) highlighted examples of the application of the aforementioned technologies and methods. To our knowledge, however, there is currently no review that examines the literature for the field of central banking systematically and with a practical stance.

Phase	Outcome
0. Building the Infrastructure	Framework to represent the subject matter (taxonomy, ontology); classification system
1. Defining the research question	Research Question
2. Searching the literature	Preliminary inclusion of studies based on database research
3. Selecting the studies for inclusion	Set of final eligible studies
4. Assessing the quality of included studies and structuring of their results	Assessed and structured studies
5. Combining the results	Representation of the gained and integrated results, e.g., in a 'Summary of findings' table
6. Create a structured report	Report on the findings and the evidence gained

Figure 2: Process for a Systematic Review (based on Cooper 2016; Goeken 2011).

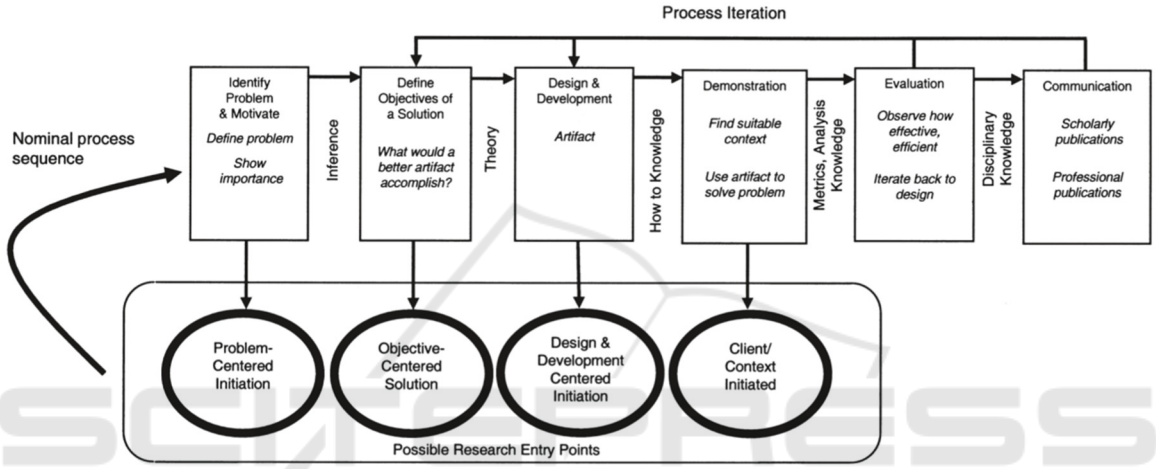

Figure 3: DSR methodology process model (Peffers et al. (2007).

3 METHODOLOGICAL BACKGROUND

Our research endeavour can be sensibly classified in two ways. First, it is review research, since it addresses the stated goals (above all: synthesizing the existing knowledge; representation of the "state of the art" at a given point in time). Accordingly, it is based on the now established and widely accepted process for reviews, which is illustrated in figure 2 – phase 1 to 6. On the other hand, it is also design-oriented in parts, since the result is intended to be an artifact that provides support for practitioners (presenting the "body of knowledge" for practical application and practical usefulness). In this respect, it should also be guided by approaches from Design Science. In this field, the DSRM by Peffers et al. (2007) is considered the "mostly widely referenced model" guiding the process (Vom Brocke et al., 2020) – see figure 3.

Thus, we envisage to combine both approaches to have the two sides of the coin in view – consolidation of knowledge and representation of knowledge for practical application. Due to page restrictions and because we are presenting research in progress here, emphasis is put in the following not on those aspects that are already well discussed (e.g. the search strategy) but on methodological considerations and preliminary results of ours that are already somehow advanced and that could lead to a fruitful discussion.

In short and certainly somewhat simplified terms, we state that, while reviews are primarily required to be transparent and reproducible in order to yield purposeful and reliable results, the development of artifacts focuses on making justified design decisions, but also on ensuring that the final result is useful and easy to use. Accordingly, the procedure for reviews tends to be rather linear, while DSRM provides for explicit feedback loops that not only allow but also stipulate the return to previous phases, so that continuous improvement is possible in an iterative process. If a result is available in phase 3 (DSRM), the process does not end, but the artifact is evaluated after its demonstration. I.e. it is also adapted, changed etc. after (initial) completion if necessary. Although

this is not explicitly discussed in the literature, a step-by-step extension of the artifact might be introduced to improve usefulness and maintain applicability of a given artifact.

Although the definition of a research question is relatively close to the first two phases of the DSRM (problem motivation and definition of objectives of a solution), a phase is added upstream here (phase 0): Due to the fact that – compared with other disciplines, e.g. medicine – IS lacks established nomenclatures and accepted taxonomies of the subject matters, this phase initially deals with the construction of a framework (or multiple frameworks) for the research topic at hand. The goal of this framework is to better conceptualize the research question and potentially decompose it in such detail as to allow for an evolutionary approach. Due to the different nature of this task and its importance, this is treated as a separate phase. This construction is a design-oriented task which itself could be based on design science methodology accordingly (Peffers et al. 2007).

Intention of this phase is also to conceptually atomize the research question so that the searching for relevant literature as well as its structuring and presentation can be supported. It also becomes meaningful with a view to possible iterations. To foresee multiple iterations makes sense against the background of practicability (even in the quite delimited domain of central banking, there are many papers that may not be manageable due to their sheer volume) and because in this way subject matter experts for the different topics can be involved (money laundering is just completely different from financial stability).

4 SELECTED PRELIMINARY RESULTS

In our approach, phases 0 and 1 were carried out iteratively. Problem identification and definition of objectives (DSRM) flow into this and lead to a high-level research question, which is shaped with the creation of a framework (conceptual model or taxonomy). One focus of the research project is the motivation (feasibility as well as the potentials of a collection of case studies and use cases on the application of advanced analytics in central banks). Since central banks traditionally have access to a significant amount of data and statistics that are regularly collected while fulfilling their mandate and inform their decision-making processes, the project will explore whether and to what extent such a collection would be useful for central banks and whether it could help to improve the decision making.

It might help central banks in their knowledge sharing to gain a better understanding of what others are researching or what is being done in general in the topic area of advanced analytics. In classifying research findings, a unified knowledge base could help identify new datasets, associated methodologies, relating projects, and assist central banks with analyses and to better understand them. Thus, for example, the assessment of the impact of policies could take place more accurately (Tissot, 2014). At the same time, a case study collection can serve as an overview so that exchanges and collaborations among central banks can be facilitated. In addition, Tissot (2014) mentions another advantage: the added value compared to "traditional" statistics when central banks run respective innovative projects – e.g., nowcasting and the use of advanced analytics. In our research, beyond the literature positions, we conducted in-depth interviews that elaborated potential benefits of a collection and we identified relevant value dimensions. These will be applied as objectives in a future evaluation.

Hence, we argue, that our research idea addresses a relevant problem in the real-world environment. From the description of the idea, we derive a high-level research question that is at the core of the systematic review: *Where in central banks (in which application areas and departments) are advanced analytic methods (and which ones) applied to which specific issues*?

The structuring and decomposition of the research question is based on the literature, software libraries (like Scikit-Learn and Tensor Flow), organizational charts, library taxonomies, and interviews. Figure 4 shows an excerpt of the structuring of typical task areas of a central bank on the left. On the right, methods and techniques of AI and ML are outlined (also in excerpts).

These frameworks allow a detailing of the research question to work on it in a decentralized way (in subprojects and by respective experts). Research questions can now be derived from the two frameworks by combining domain-specific areas and methods. For example, monetary policy in combination with one or more regression methods. If necessary, it may be appropriate to limit this further, for example, to analyses in this area that refer to the real estate market and examine related effects on price stability (if studies are available at this level of detail). In either case, the detailed research questions ultimately remain anchored in the high-level research question, which allows the results to be combined at that level (or an intermediate level).

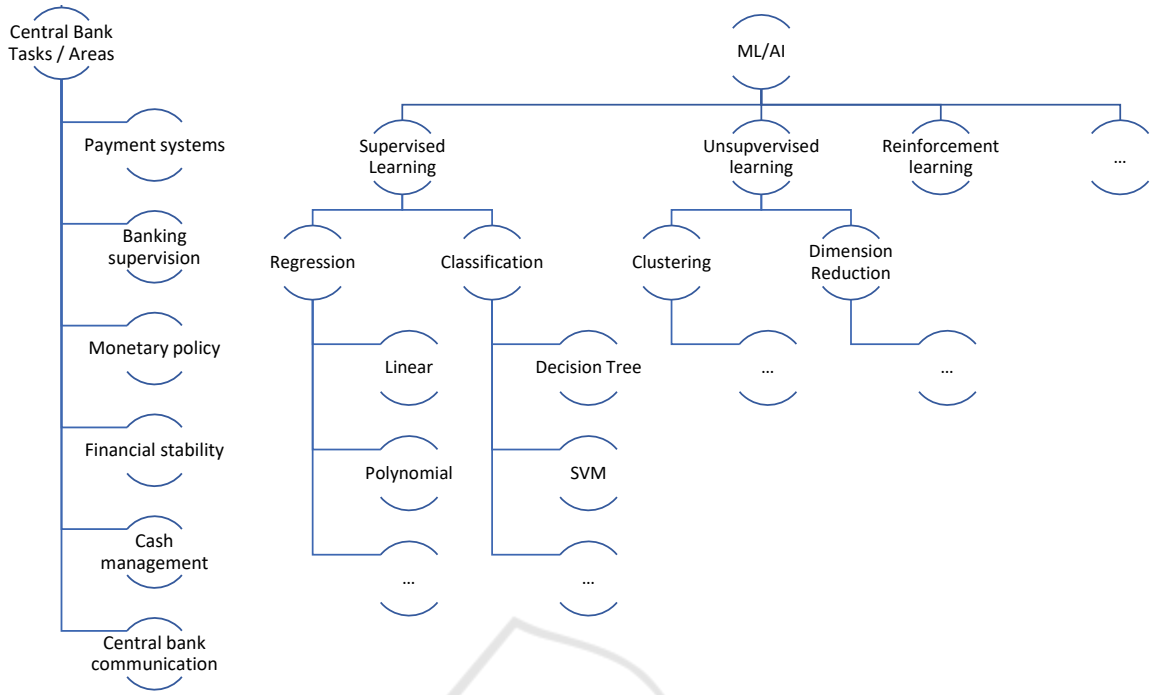

Figure 4: Frameworks for decomposition of the research question.

Along the procedure shown in Figure 2, sub-research questions have already been addressed in each of several subprojects. The recommended databases (e.g., Vom Brocke et al. 2009) were searched, although manual searches were also carried out on special databases and websites, for example those of the Bank for International Settlements (BIS), due to the special domain. After selecting and assessing the quality of studies, a manageable number were selected for deeper analysis.

Combining the results was done in "Summary of Findings tables" (SoF). Other possibilities, e.g., statistical evaluation and combination in a meta-analysis, were not considered, since mainly qualitative results are reported in the respective studies. It became apparent that there is no "one fits all" format for the SoF tables, i.e., that researchers have made modifications in each case, so that a consolidation of the various tables in terms of the research question and the research objective is a further task in the overall context. The table in the appendix shows an excerpt of a SoF table, here choosing from the framework "monetary policy" on one side and "ML/AI" on the other. The table also shows that a clear assignment to a task area is not always possible. In various places, for example, "monetary policy" is combined with other areas of responsibility (e.g., central bank communication).

Already with the results available, which in a way could represent the final outcome of a Systematic Review, it is now possible to switch again to the procedure according to the DSRM. The SoF tables are prototypes and thus, as artifacts, results of phase 3. These can be used to carry out a demonstration and, against the background of the objectives, an evaluation. For the demonstration, this publication is a first step. It is planned – in addition to an evaluation against the background of the objectives – to involve possible addressees and stakeholders from central banks in order to obtain their assessment of the usefulness, ease of use and applicability. As is known from system development, a prototype can be used as a stimulus to make verification and validation as realistic as possible. In this respect, the reviews as partial results also fulfil an essential purpose within the framework of the design-oriented approach.

5 CONCLUSION AND FURTHER RESEARCH

In this paper, a research project has been presented that aims to demonstrate and investigate how and whether systematic reviews can contribute to a better sharing and exchange of knowledge. This is investigated for the specific case of the central banking domain.

The context is to identify and summarize knowledge about the application of advanced analytics in central banks. A presentation of this knowledge in the sense of a consolidation of the "body of knowledge" can provide an overview of existing work, stimulate own projects, and promote exchange and cooperation among data scientists and developers in central banks.

Our work also consists in a (methodological) combination of DS and review research, as the result should not only bring a progress of knowledge and insight but also a very practical progress in the sense of a case study database that should promote the exchange of interested organizations and participants.

In this position paper, only the basic idea has been presented due to the fact that the methodological approach has not yet been fully defined and established, and also in the sense that only preliminary results are available so far. Reflections on the research process have been presented above and it has already been gone through several times. It is founded on established work on review research and established approaches to design science. Against the background of the special question pursued, a combination is aimed at, which currently cannot yet be regarded as well-founded in every respect. Further methodological work is needed here. It is possible that the combination of design science and review research may prove fruitful beyond the domain under consideration.

Furthermore, beyond the still small-scale "prototypes", which each covers quite delimited thematic areas, larger-scale reviews of studies and use cases need to be carried out, so that not only the feasibility is demonstrated, but also a concrete practical benefit.

REFERENCES

Abdel-Karim, B. M., Pfeuffer, N., Hinz, O. (2021). Machine learning in information systems - a bibliographic review and open research issues. In *Electronic Markets*, 31(3), 643-670.

Angra, S., & Ahuja, S. (2017). Machine learning and its applications: A review. In *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)* (pp. 57-60). IEEE.

Bawack, R. E., Fosso Wamba, S., & Carillo, K. (2019). Artificial intelligence in practice: Implications for IS research. In *Proceedings of Twenty-fifth Americas Conference on Information Systems*, Cancun, 2019

Bholat, D. (2015). Big data and central banks. Big Data & Society, 2(1), 2053951715579469.

Buch, C. (2019). Welcoming remarks. International Seminar on Big Data, Building Pathways for Policy-Making with Big Data, Bali, 26 July 2018.

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. In *International Journal of Information Management*, 60

Cooper, H. (2016). Research synthesis and meta-analysis: A step-by-step approach (Vol. 2). Sage publications.

Doerr, S., Leonardo Gambacorta, Jose Maria Serena (2021). Big data and machine learning in central banking. *BIS Working Papers* 2021, https://www.bis.org/publ/work930.pdf

Fahle, S., Prinz, C., & Kuhlenkötter, B. (2020). Systematic review on machine learning (ML) methods for manufacturing processes – Identifying artificial intelligence (AI) methods for field application. In *Procedia* CIRP, 93, 413-418.

Goeken, Matthias (2011). Towards an Evidence-based Research Approach in Information Systems. ICIS 2011 Proceedings. 10.

Irving Fisher Committee on Central Bank Statistics. (2015). Central banks' use of and interest in "big data". IFC Report.

Irving Fisher Committee on Central Bank Statistics. (2018). *IFC Annual Report 2017*.

Kinywamaghana, A., and S. Steffen (2021). A Note on the Use of Machine Learning in Central Banking, FIRE Research Paper.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. In *JMIS*, 24(3), 45-77.

Tissot, B. (2018). Big data for central banks. Bali: Bank for International Settlements.

Vom Brocke, J., Hevner, A., & Maedche, A. (2020). Introduction to design science research. In *Design Science Research. Cases* (pp. 1-13). Springer, Cham.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2009). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. In *CAIS*, 37(1), 9.

Wainwright, D., Oates, B., Edwards, H., & Childs, S. (2018). Evidence-based information systems: A new perspective and a roadmap for research informed practice. In *JAIS* 19(11), 4.

Warin, T., & Stojkov, A. (2021). Machine Learning in Finance: A Metadata-Based Systematic Review of the Literature. In *Journal of Risk and Financial Management*, 14(7), 302.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. In *MIS* quarterly, xiii-xxiii.

Younis, H., Sundarakani, B., & Alsharairi, M. (2021). Applications of artificial intelligence and machine learning within supply chains: systematic review and future research directions. In *Journal of Modelling in Management*. Vol. ahead-of-print

APPENDIX

Sub-results showing combination "monetary policy" and "ML/AI" (all procedures)

#	Title	Authors	Year	Central Bank	Task Areas	Methods and Techniques of AI and ML	Research Question	Outline / Results	Research Setting
5	Between hawks and doves: Measuring central bank communication	Tobback, Ellen; Nardelli, Stefano David Martens	2017	ECB	Monetary policy, Central bank communication	Text Mining, Support-Vector-Machines (SVM), Latent Dirichlet Allocation (LDA)	none explicitly mentioned	Hawkish-Dovish (HD) indicator, which measures the degree of "hawkishness" or "dovishness" of the media's perception of ECB's tone at each press conference.	Case Study
10	Central Bank Communications and the General Public	Haldane, Andrew; McMahon, Michael	2018	Reserve Bank of Australia	Monetary policy; Central bank communication	Lineare Regression, SVM & Random Forest	What does a high-quality ZBK look like? Perception of different addressees.	There is no consensus on what high-quality central bank communication should look like. To shed light on this, 3 important aspects are examined.	Case Study
21	Text mining for central banks	Bholat, David; Stephen, Hansen; Santos, Pedro; Schonhardt-Bailey, Cheryl	2015	Bank of England	Monetary policy; Central bank communication; Financial stability	Text Mining, Supervised Learning: Boolean techniques, Dict. techniques, weighting of words, vector space models, LDA	How can text mining be useful for addressing research topics of interest to central banks?	Discussion & step-by-step guide to text mining, including an overview of unsupervised and supervised techniques.	Report
22	The Evolving Scope and Content of Central Bank Speeches	Siklos, Pierre L.; St. Amand, Samantha; Wajda, Joanna	2018	US Federal Reserve and the Bank of Canada	Monetary policy, Central bank communication	NLP, LDA	How have the use of central bankers' speeches and the content of those speeches changed over time?	Assessment of Change. Comparison between the use of speeches by the Federal Reserve and the Bank of Canada	Case Study
24	Topic classification of Monetary Policy Minutes from the Swedish Central Bank	Cedervall, Andreas; Jansson, Daniel	2018	Central Bank of Sweden	Monetary policy, Central bank communication	Neural Network, LDA	none explicitly mentioned	Analysis of Swedish Central Bank protocols and collection of information using latent Dirichlet allocation and a simple neural network.	Case Study