SPA/2021/01 February 2021

SEACEN POLICY ANALYSIS

WHAT WE TALK ABOUT WHEN WE TALK ABOUT PAYMENTS: AN INTRODUCTION TO TEXT ANALYSIS TOOLS FOR CENTRAL BANKERS

Brian Begalle

The SEACEN Centre

The South East Asian Central Banks (SEACEN) Research and Training Centre

SPA/2021/01 February 2021

SEACEN POLICY ANALYSIS

WHAT WE TALK ABOUT WHEN WE TALK ABOUT PAYMENTS: AN INTRODUCTION TO TEXT ANALYSIS TOOLS FOR CENTRAL BANKERS

Brian Begalle

The SEACEN Centre

The South East Asian Central Banks (SEACEN) Research and Training Centre

Level 5 Sasana Kijang, Bank Negara Malaysia,
2 Jalan Dato' Onn, 50480 Kuala Lumpur, Malaysia
Tel. No.: +603 9195 1888
Fax. No: +603 9195 1801
Email: enquiries@seacen.org

For enquiries, please contact:
Brian Begalle
Senior Analyst
Financial Stability and Supervision & Payment and Settlement Systems (FSS&PSS)
The SEACEN Centre
Email: brianbegalle@seacen.org

The ***SEACEN Policy Analysis:*** **What we talk about when we talk about payments: an introduction to text analysis tools for central bankers** reflects the analysis and views of SEACEN staff and do not represent the views of its member central banks and monetary authorities.

Notes:

The SEACEN Centre recognizes "China" as People's Republic of China; "Hong Kong, SAR" as Hong Kong, China; and Korea as "Republic of Korea".

FOREWORD

This paper is the fifth in a series of publications titled SEACEN Policy Analysis. The series is intended to provide in-depth analysis of topical policy issues in macroeconomics, monetary policy, financial stability, and payments systems, with a particular emphasis on contextualizing these issues to the SEACEN stakeholder space. The papers look at the contours of cutting-edge issues that arise with ever-changing macroeconomic environments and technological possibilities and focus more on policy options than on more technical analysis such as econometric modeling.

The current paper, "What we talk about when we talk about payments: an introduction to text analysis tools for central bankers" authored by Brian Begalle, a visiting scholar at SEACEN, is an introduction to how regulatory authorities (central banks, stand-alone regulatory authorities, and deposit insurance agencies, collectively RAs) may consider thinking about how new technologies related to text analysis are being used in financial markets and also how RAs may incorporate these new technologies in their own communication strategies. We know that communication is a powerful tool that, when used effectively, can aid RAs in reaching their policy goals. We also know that communication, when unclear or interpreted incorrectly, can undermine policy goals. As the private sector increases its use of tools to understand financial markets and drive decision-making, regulatory authority intent and other communications, it is imperative that RAs gain a better understanding of these tools and how the tools may be utilized by their organizations. I wish to emphasize that the views expressed in this and all issues of the SEACEN Policy Analysis series are those of the author and do not represent the views of SEACEN's member, associate member, and observer central banks and monetary authorities.

It remains a very difficult time as the world tackles this unprecedented health crisis and its toll on human lives along with its economic and financial consequences. At the SEACEN Centre, we continue to maintain a flexible strategy by providing online learnings of the pandemic, while carrying out policy analysis of the responses on the macroeconomic, monetary, and financial front. We stand ready to provide assistance to members in building and strengthening their capacity during this time.

Mangal Goswami
Executive Director
The SEACEN Centre

February 2021

ABSTRACT

The paper explains some of the emerging technology tools that can be used to analyze text and applies them to a series of speeches by central bankers. The paper seeks to demonstrate how central bankers might begin to adopt use of these tools as a means of ensuring communication matches intent, in the moment and over time, as these tools become more readily adopted by the private sector.

TABLE OF CONTENTS

WHAT WE TALK ABOUT WHEN WE TALK ABOUT PAYMENTS: AN INTRODUCTION TO TEXT ANALYSIS TOOLS[1]

- Communication is a powerful tool and is increasingly used by central banks.
- Understanding how the consumption of communication by market participants is changing and adapting to the evolution of technology driven analysis and its sophistication can strengthen central banks' own communication as an effective tool.
- This paper explores the use of natural language processing (through word frequency) and machine learning (through sentiment analysis) to demonstrate potential uses of these tools, notably in the payment systems space.
- The efficacy of these tools is demonstrated through the analysis of selected speeches on payments from central banks, including some from this region.
- These tools serve to primarily reveal rudimentary conclusions rather than concrete ones.
- Central banks may find value in applying these tools to content they produce, the content of other regulatory authorities, commercial banks and other financial institutions, the financial press and any other communications deemed relevant to their work.

Introduction

Communication is a powerful tool and is increasingly used by central banks. Often, the words of a central banker can help move markets, drive behavior and influence the adoption of innovation. Given this power, it must be wielded carefully and effectively.

Part of success in communication is understanding how the intended audience is consuming information. Historically, individuals have consumed communications by central banks through individual reading and perhaps group discussion. As machine learning tools are being adopted by private market participants, it is worth exploring how central banks might also employ some of these tools, not in a competitive or counteractive way, but to ensure that the central banks continue to improve the sophistication of their communication to match the evolution of how recipients consume it, including its feedback effects.

This paper will explore the use of natural language processing (through word frequency) and machine learning (through sentiment analysis) to demonstrate potential uses of these tools.[2] Both tools can be used for a variety of reasons. It is important for central bankers to understand how market participants are using these tools while exploring application of the tools within their own organizations. Simple application of these tools in this paper will demonstrate how central banks may begin to think about how to incorporate these tools into their communication strategies in a variety of ways. In this paper, these tools are applied to a series of speeches given by central banks on the topic of payments from January 1, 2019 through September 10, 2020.[3]

The tools have a variety of potential uses, both in an ex-post and an ex-ante manner. There are opportunities to use them in an ex-post manner, meaning looking at things that have been already been said and written to help inform or validate what one intends to say in the future. There are also ex-ante opportunities to use these tools to increase the likelihood that future communication will have its intended impact. Both can be used with small and large volumes of unstructured data.

1. This paper is intended to lay out issues and options for regulatory authorities. It is not intended to provide technical advice or advocate for the adoption of laws, regulations, and policies. The word "may" is to be construed in the subjunctive sense and not the permissive sense. The author thanks Glenn Tasky and Mangal Goswami for helpful comments.
2. Word frequency clouds were built using the tool available at wordclouds.com and sentiment analysis was conducted using MonkeyLearn's online sentiment analysis tool.
3. A list of the content analyzed is available in Appendix A.

The sophistication of these tools is continuing to improve, but many shortcomings remain. And even as their quality improves, the value of the tools isn't to drop words in a machine and receive a clear and concise result. These are tools that should help foster reflection and dialogue within central banks to help central bankers understand what they have been saying over time as well as what they may wish to say in the future and how those communications might be received.

As consumption of communication is changing, central bankers must explore ways to change with it to ensure communication remains an effective tool. As the quality of these tools evolve and their use proliferates within the private sector, it is important that central bankers seek to better understand these tools and the ways in which the tools can be exploited to maintain or improve the quality of communications by central bankers.

So much information, same amount of time

The volume and pace of information generation are two key dimensions of the challenges posed. The current Information Age is defined as "the modern age regarded as a time in which information has become a commodity that is quickly and widely disseminated and easily available especially through the use of computer technology."[4] One of the challenges in the so-called Information Age is the sheer amount of content being produced. This can make it difficult to follow the narrative of a complex and dynamic topic over time. Reading every academic paper, speech, white paper and news article on digital currency is more than a full-time job. And that's even before one tries to think across the communications to synthesize themes, seek contrasts, etc. The second challenge is that because information moves quickly, the potential impact of incorrect or misinterpreted information can be significant before clarification can be provided.

With respect to the volume of information being produced, even if one could read all the content produced, one's ability to fully analyze and synthesize the content is compromised by one's analytical capacity/memory and bias. A person can only keep track of so much information. Perhaps more importantly, when you listen to or read a speech, you walk away with a set of assumptions and conclusions. Armed with that knowledge, you move on to future content and it influences how you consume and process the information of the next speech. Some of what you bring to future content might be called expertise, but some of it is also bias. For example, people are often subject to confirmation bias, which means you will be conditioned to be looking for information in subsequent content that conforms to the conclusions and assumptions you have derived from previous content. That is just one example of how consuming content in the traditional way may drive you to form implicit assumptions, views or conclusions that might miss things. This isn't to say you should stop reading! Rather, it raises an issue that needs to be addressed - how might one seek to counteract these potential challenges? These tools may help central bankers to reach a higher level of objectivity when they review information.

There may be significant strategic value in using these tools for review of large quantities of information over time. This paper covers selected speeches on payments by central bank leaders that can help demonstrate application of these tools because payments are a topic of keen interest to central banks over time for a variety of reasons. Payment innovations can increase financial inclusion. Payment innovations can stimulate cross-border economic activity. Absent sufficient controls, payment systems become a mechanism for money laundering and the financing of terrorism. Payment system outages can disrupt market functioning. While this variety of reasons exists over time, a specific reason may have heightened importance at a point in time. The value of reviewing historical content is that it can help inform whether the central bank is covering the range of issues effectively over time, as well as specific issues over shorter time intervals, in a manner that is commensurate with the central bank's strategy regarding payments. Absent this kind of objective review, central bankers may be left with an incorrect or even incoherent message over time regarding payments goals and objectives for themselves and those they seek to influence. Review

4. https://www.merriam-webster.com/dictionary/Information%20Age

over a longer time horizon might ensure that central bankers are giving sufficient airtime to challenges, benefits and risks as the payments system continues to evolve.

With respect to the speed at which information is consumed, the challenge is that we continue to live in a world where speed is deemed paramount. Faster is always seen as better in financial markets. You want to buy the next big idea before other investors see it and you want to exit losing investments before others do the same. This is reaching the point where we are beginning to replace human judgment seeking to understand market sentiment with algorithms that determine sentiment and execute trades accordingly.

The power of words

We already know that the words of central bankers have the ability to move markets. Whether we are talking about Greenspan's "irrational exuberance" comment, Draghi's "whatever it takes" statement, or Bernanke's words that drove the 2013 "Taper Tantrum", central bankers' words matter.[5,6,7] And in those specific cases, it's not clear if the market reaction that followed was the intended effect. However, a more recent example is more conclusive. In July 2019, the words of John Williams, president of the Federal Reserve Bank of New York, seemed to not have their intended effect, to the point that an official "clarification" was deemed necessary. In that case, the phrase "act quickly" as a general statement for central bankers, based on many years of academic research, was interpreted to be about upcoming policy actions, and market expectations for a larger than expected future interest rate cut jumped.[8]

Central banks can harness technology to enhance their communication. This is all to say that the words of central bankers carry tremendous power and the more that can be done to ensure that the chosen words match the intent, the more successfully central banks can use communication as a tool. Given that technology is driving changes in how communication is consumed and analyzed, it is worthwhile for central banks to explore how they might take advantage of these tools for their own communication strategies as well as other means.

Word frequency and sentiment

Two potentially useful tools for text analysis are word frequency and sentiment analysis. These tools are relatively simple to understand in terms of what they do. This simplicity makes them relatively easy to use. However, their degree of sophistication is still under-developed in many cases, which also limits what they can achieve. It is important to understand the benefits they can provide as well as the challenges in reliance on them.

The first potentially useful tool is text analysis focused on word frequency. Word frequency helps to reveal "what" is being communicated or talked about. This can be particularly useful when seeking to assess large volumes of content in an objective way. There are numerous tools available to conduct this analysis and display the results. One method is to use word clouds, which are a visual representation of content that displays words by frequency - the more often a word appears in the content being analyzed, the larger it appears in the cloud. This is a different way to consider the content and may reveal things that are difficult to see through a traditional reading of the material. This tool, when applied to a central bank's own content, may help to more easily reveal what the central bank has been communicating about over a time horizon.

Word frequency can be viewed as an indicator of importance – the more one uses a word, the more important that word, or the concept represented by that word, is to the message being conveyed. For example, if the goal is to raise awareness of the need for strong cybersecurity controls in the payment system, a speech in which the word "cybersecurity" is used only one time is probably generally less effective than a speech in which that word is used ten times. Using the term

5. https://en.wikipedia.org/wiki/Irrational_exuberance
6. https://www.ecb.europa.eu/press/key/date/2012/html/sp120726.en.html
7. https://www.investopedia.com/terms/t/taper-tantrum.asp#:~:text=Taper%20tantrum%20refers%20to%20the,quantitative%20easing%20(QE)%20program
8. https://www.cnbc.com/2019/07/18/fed-clarifies-williams-speech-that-market-took-as-signal-of-a-rate-cut.html

"cybersecurity" twice across five payment speeches in a year is likely to be less effective than using the term fifteen times across five payment speeches in the same year. Of course, this will not always be true. There is almost certainly a speech where fewer uses of the word describing the key subject may be more effective than a speech using more instances of that word. The point is that word use conveys importance to an audience. In general, the more one uses a word, the more important the concept represented by that word, compared to things not discussed. Or, in the alternative, the absence of certain words might be interpreted to convey a lack of interest in the concept represented by those words. This idea is more relevant over larger bodies of content and longer periods of time. Continuing with the example, one speech to a particular audience on a single day may be tailored to that audience and not address cybersecurity risks at all because that topic is not appropriate for that audience and event. However, ten speeches about payments over three years with no mention of cybersecurity risk may drive a consumer of that content to believe that cybersecurity risk is not a concern to that speaker or organization.

Another example can provide more insights into word frequency analysis. The Federal Reserve, the central bank of the United States, is generally understood to have three priorities – monetary policy implementation, the supervision of banks and the provision of payment services. There is nothing explicit to suggest that the Federal Reserve thinks any one of these priorities is more important than the others. Given that, would you expect each of these topics to get equal time in speeches by leaders of that organization? What might a review of five years of speeches reveal? Perhaps for very logical reasons different parts of the mission are emphasized at specific points in time - maybe governors speak about bank supervision more frequently during and in the immediate aftermath of a financial crisis? Do some governors with certain responsibilities prefer public speaking more than others? Do they consciously believe certain parts of their mission are more important than others? Is there no strategy to their speeches over time that maps back to their priorities? There is no straightforward answer to these questions. However, looking at word frequency of communications over a long time horizon might begin to reveal answers to these questions as well as shortcomings in the organization's communication strategy over time.

The second potentially useful tool is sentiment analysis, which is a useful companion to word frequency analysis. The way that the tool works is that it reviews the inputted text and assigns ratings for words based on a pre-defined scoring system. For example, words like "excellent" or "great" would be assigned a value associated with a positive sentiment and words like "terrible" and "disappointing" would be assigned a value associated with a negative sentiment. These tools assign values to words and then calculate a confidence interval to assign to the overall sentiment for the text. For example, the end result may be something like "94% confident that this text has a positive sentiment." Just as word frequency seeks to show "what" is being talked about, sentiment analysis helps reveal "how" something is being talked about. How one speaks about a topic is perhaps as important as the topic itself, therefore, assessing content for tone (positive, neutral, or negative) can be useful for understanding the communication.

There are a variety of sentiment analysis tools available and, as one might imagine, the quality continues to evolve and improve. The complexity of the content (use of colloquial language, use of complex grammar such as double negatives, etc.) can present challenges to the tool. Nevertheless, it may be useful to understand the sentiment associated with communications from central banks over time. In fact, some work has been done in this regard.[9] Sentiment analysis is a concept gathering momentum across market participants and it is critical that central bankers understand how, and to what degree, financial market participants are making use of these tools. Sentiment analysis is being used in a variety of places such as customer service, equity trading and other areas where it can be useful to analyze large amounts of unstructured data (customer feedback, speeches, quarterly and annual reports, etc.) As noted earlier, speed is paramount. Accurate automated sentiment analysis can provide a trading edge due to speed. A computer can process a central bank leader's speech, the minutes of a monetary policy meeting, a company's press release, etc., and quickly assign a favorable or unfavorable sentiment and execute trades before a person reading the

9. Sentiment in Central Banks' Financial Stability Reports by Correa, Ricardo, Keshav Garud, Juan M. Londono, and Nathan Mislang available at: https://www.federalreserve.gov/econres/ifdp/files/ifdp1203.pdf

content has gotten through the first paragraph, much less figured out its meaning and executed trades based on that interpretation. Under this use, a greater reliance on automated sentiment analysis, may serve as an amplifier of market volatility. Would the market have dropped more significantly in December 1996 after Greenspan's "irrational exuberance" comment if more automated sentiment analysis had been driving trading activity?

Sentiment analysis is not just used for speed. It can also look across significant amounts of unstructured information over time to determine sentiment to support longer term investing decisions. In some cases, these tools are proving effective for investing.[10] Should these tools drive more investing over time, they may serve as an amplifier of bubbles and volatility, creating perhaps larger crowded trades more quickly than we observe today.

The key question is whether the sentiment attached to the content is aligned with the communication goals. The sentiment a central banker may be seeking to convey can legitimately be very different from speech to speech. While we often think of central banks as being neutral or objective, one can imagine that a specific speech may be assessed as being very negative (perhaps when speaking about money laundering) or very positive (perhaps when speaking about financial inclusion.) There is no right or wrong sentiment, but striking the proper tone is critical to the efficacy of the communication.

Word frequency – what are central bankers saying?

In order to demonstrate how word clouds can function as an analytical tool, we play a matching game. Figure 1 contains four different word clouds. Each word cloud represents speeches given by central bank leaders from four organizations on payments topics from January 1, 2019 through September 10, 2020. The organizations represented are the European Central Bank, The Board of Governors of the Federal Reserve System, the Hong Kong Monetary Authority and the Monetary Authority of Singapore.

Can you identify which word cloud represents the speeches from each organization? Perhaps before guessing, it would be worthwhile to think about what you might expect to see. Do you think you know how leaders in these organizations are talking about payments right now? What would you expect them to be focused on? The benefits of financial inclusion? The risks of money laundering? The increasing role of fintech companies in payments? Perhaps you would expect them to be emphasizing these things equally?

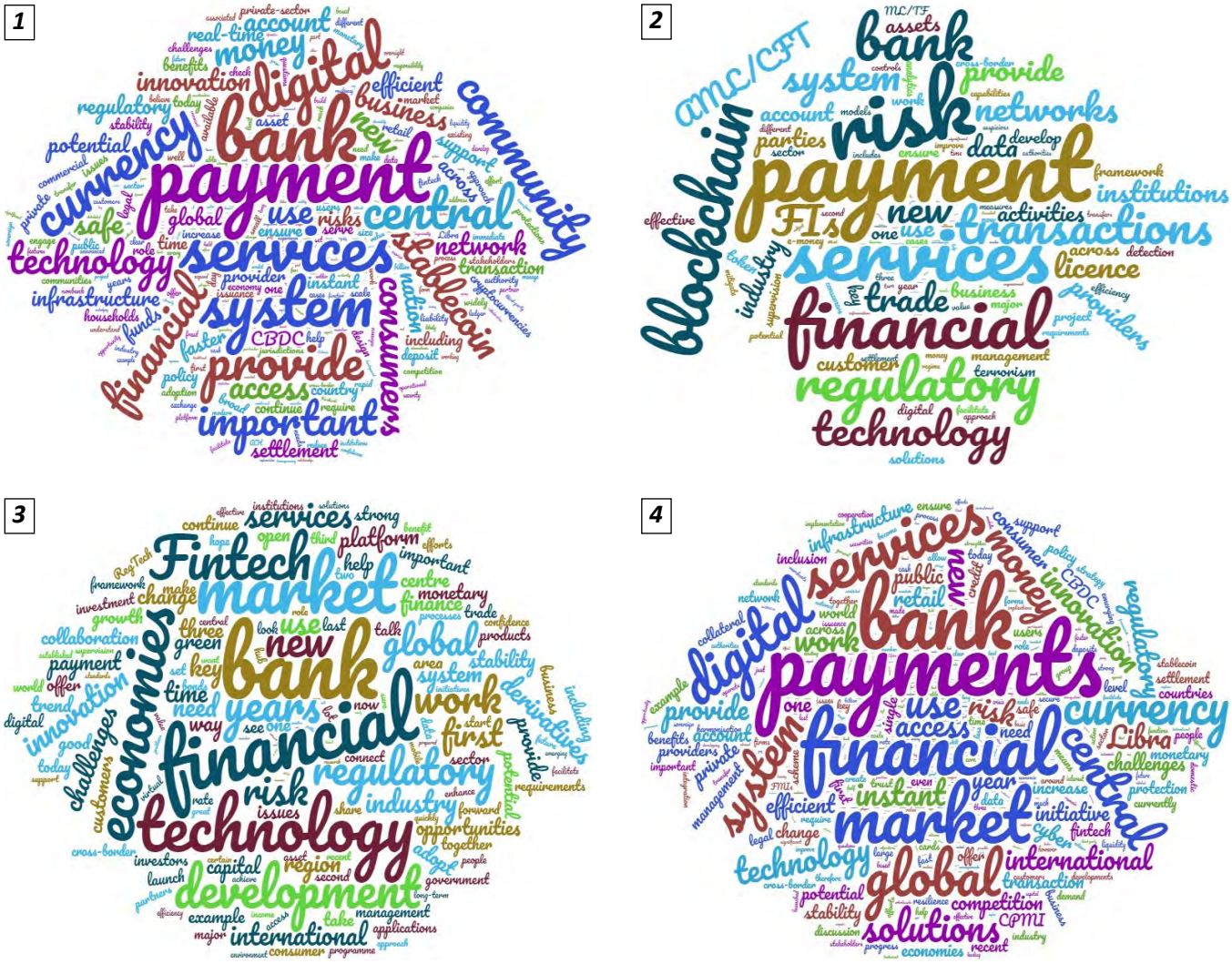
Before moving on, one can think about (or even write down) the words one would expect leaders at each organization to most frequently use when giving a speech about payments. An example (perhaps not surprising) is that when central bank leaders talk about payments, they use the word "payment" or "payments" often. What other words would you expect to see? Also, would you expect leaders from these organizations to be using different words? What might some of those differentiating words be? Now look at the word clouds and see if you can figure out which word cloud represents each of the organizations. What words provide clues?[11,12]

10. https://www.bloomberg.com/news/articles/2020-10-21/market-beating-robot-fund-manager-shys-away-from-cyclical-bets?sref=s1nXlOUg

11. The point of this exercise is to make you think about how central bank leaders speak about the topic of payments. Because of that, key words or phrases that are not relevant to the topic of payments, but might have provided clues as to which organization is represented have been removed. For example, terms like HKMA, Singapore, ECB, and other proper nouns which easily identify the organization have been removed.

12. The word clouds are fit to page so you can easily compare them, which means you may not easily see all the words. While larger word clouds may allow one to see more words, it is important to remember that this is a relatively simple tool and focus should be placed on large words you see and large words you don't see, but would expect to see. Focusing on all the words is not an effective or efficient use of the tool.

Figure 1



Word frequency analysis

The similarities across some of the word clouds are what one might expect. As noted earlier, there are often many words that would always be used when talking about payments. Bank is an example. Although to be fair, if the word bank was missing, it might indicate that banks are not an important part of the payments' ecosystem in a country or region, which is certainly a possibility. In the alternative, significant use of the word "bank" might be indicative of a presumed reliance on banks in the payments system, which may reveal an assumption or potential bias regarding the many non-bank fintech solutions currently under exploration around the world.

The differences in words are probably what is more interesting. For example, the second word cloud is the only one of the four that suggests frequent use of the word "blockchain." Does that reveal a strategic priority with this central bank relative to the others? That is also the word cloud that has most prominent use of the words "risk" and "AML/CFT." The first word cloud suggests frequent use of "digital", "currency" and "stablecoin." It's worth noting that "digital" and "currency" are also frequently used in word cloud four. Word cloud four also appears to have frequent mentions of "Libra." Which two of the organizations might you expect to be speaking more frequently about digital currency? Why?

There is no single, correct way to do this analysis. The point of creating word clouds or looking at word frequency in some other form is to approach the content from a different perspective to see if it reveals something that might otherwise go unnoticed. Sometimes, such as in this case, you can see meaningful differences in how a set of individuals is speaking differently about a topic over time. Perhaps those distinctions were already known to you. However, if you're having a hard time deciding which word cloud belongs to which institution, why might that be the case? Did you expect them to be speaking similarly? Did you have incorrect assumptions about how these institutions talk about payments?

The absence of certain words can be as useful and interesting as the frequent use of certain words. When a person approaches this content, she or he has a set of assumptions. Some were proven true and others were not. First, one may have expected to see many of the "buzz" words currently present in the payments world as a result of considerable fintech resources going into payments. References to technology, Fintech, digital currency, Libra, etc. would have been consistent with one's expectations. Absence of these types of words might suggest that a central bank is not really following the investment, potential for progress and evolution within the world of payments.

The use of word frequency analysis can be useful for raising questions for further exploration that may lend itself to better communication. While one would also expect some of the more traditional words associated with payments to be more prevalent, they were not. There are three examples. The first example is the term "interoperability." This word has been associated with payments for a long time as central bankers have desired for years a payments infrastructure to match the evolution of the global economy. While it was used in some of the content, the term never appeared frequently enough in any of the organizations' speeches to be easily visible. While it may be a technical term, it is often viewed as one of the largest obstacles to cheaper, more efficient cross-border payments. Given the G20 objective focusing on cross-border payments, should one expect to see it featured more prominently in the word clouds? The Stage 1 Financial Stability Report to the G20 entitled "Enhancing Cross-border Payments" uses the term "interoperability" six times.[13] While that may not seem like often, it is worth nothing that the report is only five pages long. Further, the report itself notes that "As catalysts and facilitators of improvements by the private sector, public authorities can encourage interoperability of private sector improvements to payment systems." What might insignificant use of this word by leaders in central banks mean with regard to achievement of this objective? It should be noted again that prominence of a word, or lack thereof, isn't necessarily a problem. Rather it reflects a potential misunderstanding. In this case, perhaps it was incorrect to assume that central bankers would be talking about this issue. A case can be made that problems regarding interoperability are well understood within the world of payments and therefore central bankers need not use their influence in this manner. The value in conducting this type of analysis is to reveal these types of gaps in perceptions or understanding. Within the specific central bank, the question becomes "Is this a conscious choice or something unintentionally overlooked?" While that cannot be determined here, as indicated, it raises some useful questions. For the leader of an organization, the question is "Am I talking about interoperability not enough, too much, or in an amount that seems consistent with our strategic priorities?"

The second example may be more surprising and perhaps troubling. While it may be heartening for a bank supervisor to see the word "risk" featured fairly prominently across all the word clouds, it is surprising to see that only one word cloud contained "AML/CFT" within easy view. Given the enduring problems with money laundering and the financing of terrorism, this specific risk receiving less attention, particularly in the context of potential massive technological change in payments, likely seems counter-intuitive to a seasoned bank supervisor. As fintech companies move into the payments space, there is some potential evidence that they are not managing this risk any better than traditional banks.[14,15] Does less focus on this risk in speeches by leaders result in less focus on this risk by market

13. https://www.fsb.org/wp-content/uploads/P090420-1.pdf
14. https://www.bbc.com/news/technology-47751945
15. https://www.pymnts.com/news/security-and-risk/2020/wirecard-linked-to-mafia-money-laundering/

participants? Perhaps a good use of communication by central bank leaders would be to seek to redirect some fintech energy into solutions which measurably reduce money laundering and the financing of terrorism? To be clear, this is not to suggest that central bank leadership doesn't care about AML/CFT risk, but perhaps one contributing factor to the persistence of this issue is a perceived (or actual) impression that this is not a priority, given how little it is raised in speeches by leaders of some central banks?

And what about cyber? Finally, the near-complete absence of the term "cyber" in the word clouds suggests little mention of the concept of cyber risk across the speeches. While cyber risk transcends the business of payments, it continues to be one of the key risks in payments right now.

The main takeaway of this exercise is to highlight that time is valuable and leadership cannot focus on every possible issue with the highest level of priority, but having easy tools to validate how communication is being used over time can be critical to the ongoing effectiveness of communication. Indeed, the point here is not to tell central bankers what to talk about but to show that these tools will allow an audience to analyze their words in a new way and may raise questions or draw conclusions based on the results of the analysis. What matters for central bankers is that strong alignment exists between their words and their priorities over time because this tool makes it easier to "see" whether that alignment exists. Use of this tool in this capacity can begin the process of validating the degree of alignment.

It is worth emphasizing again that this tool serves to primarily reveal rudimentary conclusions rather than concrete ones. For example, simply because the leaders of the organization don't seem to be using certain words or speaking about certain topics, it doesn't necessarily mean these things are not important. For example, maybe it is other officials at the organization that speak publicly about cyber risk or AML/CFT risk. However, the individual chosen to speak to a topic, in terms of seniority in an organization is also indicative of priority. The governors of central banks get more press coverage of their speeches than subordinates and press coverage can be an effective amplifier of priorities. The point is that this tool provides the view from 30,000 feet. It reveals, through word use by some of the most senior leaders in the respective organizations, some possible conclusions about their priorities as well as perhaps some explicit and implicit assumptions of the analyzer. To the degree that there is dissonance, further work will be necessary to determine the degree to which the central bank communications are aligned with the organization's strategy.

Word frequency challenges

The tool for word frequency analysis may be simple but it can present challenges. It's useful to discuss how one may need or wish to manipulate the data. For this technology to be effective, one wants a usable set of data. In the case of word clouds, that means having word clouds that are readable so one can draw conclusions. One potential problem is that the larger the volume of content, the more individual words in the cloud and one can end up with a word cloud that is hard to read and from which to draw conclusions. There are ways to modify use of this tool, but this must be done carefully and with knowledge of the pitfalls.

Careful data manipulation may be needed to make the word cloud more manageable for analysis. Consistent with this concept, the word clouds one sees in this paper are not perfect representations of all the words across all the documents. Manipulation was done to make the word clouds more readable. For example, some different words have been merged into a single word with the result being one appearance of the single word to the exclusion of others merged into it. An example of this is merging the singular and plural forms of a word. If someone has used the word "risk" 87 times across a body of text and the word "risks" 19 times, the question becomes, for purposes of what you're doing, do you want the word to just show up once representing the combined 106 uses of the word and its plural form? Another example is to merge all variations of a root word to capture the concept in a single word. For example, combining the terms "efficient", "efficiency" and "efficiently" to display as the single root term "efficient" or the concept of "efficiency" may improve readability without compromising integrity and emphasis of the meaning. This combining can improve use of the sizing feature of the word cloud as well as improve its overall readability. However, one cannot be certain of the context in which variations of a word or even

a single word may be used because many words have multiple definitions. For example, consider the words "deposit" and "deposits." Deposit could be referring to deposit insurance, whereas deposits could be referring to cash deposits. In this case, combining these words may undermine the goal of the exercise.

There is no single right or wrong way to do the data manipulation. And sometimes, how you choose to do it may ultimately have little impact on final results. For example, combining two words of little usage to create a single word of still limited usage may not change the conclusions you draw. It is just important to note that combining forms of a word must be done cautiously, with knowledge of the potential for misinterpretation. The artificial intelligence attached to some of these tools continues to improve and so the scale of this problem will probably be reduced over time as machines learn to more effectively use contextual clues (other words next to or near a specific word) to understand the meaning of a specific word that may have multiple possible meanings.

You may also want to exclude low frequency words. If you're looking across a body of documents that are comprised of tens of thousands or hundreds of thousands of words, you may conclude that one occurrence of a certain word or even a few occurrences of that word are not significant. You may also want to exclude high-frequency, commonly used words. If it's an unimportant word, like "the", deletion doesn't really affect the conclusions one may draw. If it's an important word to your analysis, like use of the word "risk", only one or two occurrences across so much text may be telling you something very important about what isn't being discussed.

For this paper, a series of choices in how to merge words was made and some commonly used and infrequently used words were deleted to illustrate how the data may be usefully manipulated to improve the efficacy of the tool in this case. Others may have made a different set of choices or opted to not change the data in any way. The changes made here were for illustrative purposes to showcase some of the value and shortcomings of this tool. The value of word clouds and other word frequency tools isn't to understand the use of all the words. The objective is to see if this method of looking at the content reveals something you otherwise may not see or raises questions that may be worthy of further exploration within the content. For example, if you see a word cloud of 50 speeches about financial stability, a word cloud can make visually obvious to you that the word "oil" has been used frequently or that the word "debt" has been used frequently in a way that a traditional reading of the content may obscure.

Sentiment analysis – how do central bankers feel?

Given that technology may allow for sentiment analysis to become an increasingly powerful tool, how might central banks be using it for their own communications? Similar to "what" central bankers decide should be a topic of focus, "how" that topic is discussed is critical. Central bankers are generally perceived or want to be perceived as "neutral" in many communications. However, there can be value in strategic use of sentiment in a specific communication or a series of communications over time. With respect to how to think about sentiment over time, central bankers may want to be upbeat at certain times (perhaps when speaking about innovation), but more negative at other times, (perhaps when speaking about certain risks.)

Sentiment analysis

Similar to word frequency analysis, sentiment analysis can be an effective starting point for raising questions for further thought and analysis. When looking at the speeches previously analyzed for word frequency, the vast majority of them are deemed "positive" with a high confidence interval for that rating. Specifically, 25 of the 29 speeches are assigned a positive rating.[16] This is not necessarily a bad thing. It is easy to imagine that much of the conversation regarding payments would be positive. Innovations with regard to speed, lower cost and increased financial inclusion are all aspects of payments discussions that would likely be framed using positive words. But, similar to the word frequency analysis, it raises the question of whether all the various topics related to payments, particularly the more negative aspects (AML/CFT risk, cyber risk, interoperability challenges), are getting the necessary attention in the speeches of central bank leaders.

16. A list of the speeches with assigned sentiment is provided in Appendix B.

The relatively positive sentiment may link back to the topics being discussed. As the word frequency analysis revealed, much of the content was focused on fintech/technology and digital currency. It makes intuitive sense that speeches might view technology favorably in the world of payments as technology has the potential to reduce costs and improve speed. Also, other academic work has concluded that central bank sentiment regarding digital currency has shifted to a more favorable view.[17] Once such a change in sentiment has been identified, that then allows one to seek to understand why it has occurred. Have central bank views on digital currency become more positive because the merits are better understood? Perhaps COVID-19 is driving sentiment because it created a new need or benefit to digital currency that was not previously contemplated?

It is worth noting that the three speeches assigned a negative sentiment were from the European Central Bank. Two of the speeches, "An ECB digital currency – a flight of fancy" and "Money and private currencies – reflections on Libra" are about digital currency. A negative sentiment is not a bad thing. It may be that the ECB, by delivering some speeches on digital currency with a negative sentiment and others with a positive sentiment, it is providing a more balanced view of the topic.

Similar to what was noted in the word frequency analysis, this tool isn't meant to conclude that something is correct or incorrect. What the tool can do is help one understand how important topics are being discussed. It is up to each central bank to determine the degree to which it is framing these topics in a positive, negative or neutral sense. For example, a risk topic may not need be framed negatively. A speech about progress being made on money laundering issues may result in positive sentiment rating.

As noted earlier, a better understanding of the degree to which these tools are used within financial markets may be critical to the efficacy of central bank communication if there is more widespread adoption. While speeches about payments tend to not be market moving events, speeches on monetary policy have the potential to be. The speech by John Williams, president of the Federal Reserve Bank of New York, may be a canary in the coal mine. Failure to better understand how markets will interpret communication may result in more frequent clarifications which may ultimately undermine communication as a central bank tool.

Sentiment analysis challenges

Similar to word count tools, sentiment analysis tools have potential pitfalls. First, the sentiment may be a function of the overall text being analyzed, so many neutral words may have the effect of reducing the positivity or negativity in the eyes of the tool in a way that may not conform to how one is interpreting the content through reading or listening. Also, the efficacy of the tool can also be challenged by complex grammar usage such double negatives because many of the tools assign sentiment to words in isolation. However, the technology continues to improve and will undoubtedly increase its accuracy over time potential resulting in even more widespread adoption by financial market participants over time.

Conclusion

There are a variety of uses of these tools for text analysis within a central bank. One might use one or both of these tools ex-ante to understand how a speech or paper might be interpreted by those in the financial markets using such tools for analysis. This would help ensure that the tone desired by the communication is effectively received by the audience. At times, results might warrant changes to word choice, emphasis, etc., particularly if market actions become more automated and driven by these same tools.

While this paper analyzes a set of content, specifically speeches on payments, from central banks, use of these tools for central bankers is not limited to that purpose. Central bankers could apply this to any of their own content. They could look at financial stability reports or other public documents over time to illuminate and validate changes or identify the absence of them. Perhaps an exercise that looks across these reports for the last few years is a valid starting point for production of the latest report? The specific examples provided are meant to demonstrate how this type of analysis can raise questions for further consideration or serve as a starting point to more objectively measure aspects of one's communication strategy.

17. https://www.bis.org/publ/work880.pdf

Separately, these tools can also be used beyond public communication by a central bank. For example, a supervisor might find it useful to apply these tools across a variety of content from a bank it supervises (procedures, policies, meeting minutes, etc.) to understand priorities, risks and strategy evolutions over time. Supervisors might also apply these tools across content from multiple firms to look for similarities and differences with respect to discussions by Boards of Directors, Risk Committees, etc.

There may also be value in applying these tools to similar content at peer central banks, to begin to understand how your communications are similar and different. To the degree they are different, are there deliberate, logical reasons that may support these differences? In this case, perhaps things such as scope of authority, scope and design of payments infrastructure, role of payments in economy, and other factors might warrant speaking differently than peers. There may also be domestic events (a payments outage or money laundering scandal) that might affect how payments are discussed at a point in time. It is worth noting that the European Central Bank is making content easily available for parties to conduct this type of analysis to "stimulate natural language processing research on the impact of our speeches on the market and beyond."[18]

Finally, it may also be useful to apply these tools across the body of content produced by financial news agencies to further one's understanding of how "the market" might be talking about certain issues over time. For example, in the payments world, understanding the degree to which articles speak to blockchain or digital currency over time, coupled with sentiment analysis might help to reveal high-level insights about enthusiasm, shifts in prioritization, etc. Changes in what "the market" is focused on or talking about may be important inputs into a central bank's communication strategy. Beyond payments, sentiment analysis can be an effective tool for "taking the temperature" within the financial system about markets such as housing, commodities, equities, etc. How people talk about markets at a point in time and over time affects human behavior that drives the future of those markets, so tracking sentiment can be an effective tool to supplement one's understanding of markets. Also, looking at how the financial press talks about your organization may also shed light on the efficacy of your communications.

With both tools, a user should think of them as a complement to how content might be consumed and created today. These are not meant to replace the traditional reading of content, but, instead, to potentially reveal new insights or confirm one's understanding, particularly when it comes to large volumes of content. Overall, central banks may find value in applying these tools to content they produce, the content of other regulatory authorities, commercial banks and other financial institutions, the financial press and any other communications deemed relevant to their work.

As participants in the financial markets explore adoption of these tools, it seems important that central bankers should be rising to meet the moment and thinking about similarly exploiting these technologies to help them to better achieve their mandates and missions. If we are truly moving to a system reliant on central banks choosing the right words and striking the right tone, as determined quickly by algorithms, as opposed to more slowly by humans, central bankers must adapt their approach to communication if they expect it to continue to be an effective tool. In the absence of doing so, it seems there is higher risk of more unintended volatility and perhaps issuance of more "clarifying statements" as these tools are more readily adopted within financial markets. Such measures will only harm the credibility of the central bank over time. These tools are imperfect, but that doesn't mean that they cannot provide some value now. More importantly, learning about their capabilities and shortfalls now, as the private financial sector embraces them, will help ensure regulators aren't playing catch-up later on, a situation in which regulators seem too often to find themselves.

18. https://www.ecb.europa.eu/press/key/html/downloads.en.html

Appendix A – Speeches analyzed

The speeches in the list below comprise the content analyzed. I tried to find all relevant speeches (speeches about payments) by leaders of the respective central banks. I would note that while some central banks' websites have filtering tools that allowed to filter on speeches about payments specifically, other websites simply have a list of speeches in chronological order which required me to undertake a more subjective process. As noted earlier in the paper, this set of content is more to demonstrate how one might use the text analysis tools, as opposed to drawing any specific conclusions about the content. In the event a relevant speech was missed, I regret the error.

Institution	Speech title	Link
ECB	Fintech for the people	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190131~24b8e3fb49.en.html
ECB	The possible triangle: frictionless movement of payments, securities, and collateral across Europe	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190131_1~8d17d2ff95.en.html
ECB	Promoting innovation and integration in retail payments to achieve tangible benefits for people and businesses	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190207~f900d9105b.en.html
ECB	Lending and payment systems in upheaval – the fintech challenge	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190226~d98d307ad4.en.html
ECB	Future-oriented approach to innovative pan-European retail payment solutions	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190506~14299d5b80.en.html
ECB	Cyber resilience as a global public good	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190510_2~2e988cb439.en.html
ECB	Payments for the people	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190527~dc0760b772.en.html
ECB	Transformation of the retail and wholesale payments landscape in Europe	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190614~509788164a.en.html
ECB	Money and private currencies – reflections on Libra	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190902~aedded9219.en.html
ECB	Welcome address: Challenges in the digital age	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190704~975f757478.en.html
ECB	Digital challenges to the international monetary and financial system	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190917~9b63e0ea23.en.html
ECB	Introductory remarks: Digital currencies, focusing on Libra	https://www.bis.org/cpmi/speeches/sp190925.htm

Institution	Speech title	Link
ECB	Toward the retail payments of tomorrow: a European strategy	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp191126~5230672c11.en.html
ECB	An ECB digital currency – a flight of fancy	https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200511~01209cb324.en.html
ECB	Payments in a digital world	https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200910~31e6ae9835.en.html
FRB	Fast Payments	https://www.federalreserve.gov/newsevents/speech/brainard20190805a.htm
FRB	Digital Currencies	https://www.federalreserve.gov/newsevents/speech/brainard20191016a.htm
FRB	Update on Digital Currencies, Stablecoins, and the Challenges Ahead	https://www.federalreserve.gov/newsevents/speech/brainard20191218a.htm
FRB	The Digitalization of Payments and Currency	https://www.federalreserve.gov/newsevents/speech/brainard20200205a.htm
FRB	Direction of Supervision: Impact of Payment System Innovation on Community Banks	https://www.federalreserve.gov/newsevents/speech/bowman20200227a.htm
FRB	The Future of Retail Payments in the United States	https://www.federalreserve.gov/newsevents/speech/brainard20200806a.htm
FRB	Update on Digital Currencies	https://www.federalreserve.gov/newsevents/speech/brainard20200813a.htm
HKMA	Opening Remarks at the Hong Kong- Paris Financial Seminar	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/01/20190116-1/
HKMA	Speech at the ISDA 34th Annual General Meeting in Hong Kong	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/04/20190410-1/
HKMA	Keynote Speech at HKAB Distinguished Speaker Luncheon	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/04/20190410-1/
HKMA	Keynote Speech at HKMA Fintech Week	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/11/20191106-1/
MAS	Payment Services Bill	https://www.mas.gov.sg/news/speeches/2019/payment-services-bill
MAS	Is the Blockchain a Solution Looking for a Problem?	https://www.mas.gov.sg/news/speeches/2019/is-the-block-chain-a-solution-looking-for-a-problem
MAS	Combatting Financial Crime through New Technologies Built on Strong Fundamentals	https://www.mas.gov.sg/news/speeches/2019/combatting-financial-crime-through-new-technologies-built-on-strong-fundamentals

Appendix B – Sentiment analysis

Institution	Speech title	Sentiment	Link
ECB	Fintech for the people	Positive – 89.9%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190131~24b8e3fb49.en.html
ECB	The possible triangle: frictionless movement of payments, securities, and collateral across Europe	Positive – 74.9%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190131_1~8d17d2ff95.en.html
ECB	Promoting innovation and integration in retail payments to achieve tangible benefits for people and businesses	Positive – 97.8%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190207~f900d9105b.en.html
ECB	Lending and payment systems in upheaval – the fintech challenge	Positive – 98.8%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190226~d98d307ad4.en.html
ECB	Future-oriented approach to innovative pan- European retail payment solutions	Positive – 92.5%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190506~14299d5b80.en.html
ECB	Cyber resilience as a global public good	Positive – 100%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190510_2~2e988cb439.en.html
ECB	Payments for the people	Positive – 98.9%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190527~dc0760b772.en.html
ECB	Transformation of the retail and wholesale payments landscape in Europe	Negative – 91.1%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190614~509788164a.en.html
ECB	Money and private currencies – reflections on Libra	Negative – 53.4%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190902~aedded9219.en.html
ECB	Welcome address: Challenges in the digital age	Positive – 92.7%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190704~975f757478.en.html
ECB	Digital challenges to the international monetary and financial system	Positive – 99.9%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190917~9b63e0ea23.en.html
ECB	Introductory remarks: Digital currencies, focusing on Libra	Positive – 84.6%	https://www.bis.org/cpmi/speeches/sp190925.htm
ECB	Toward the retail payments of tomorrow: a European strategy	Positive – 95.3%	https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp191126~5230672c11.en.html
ECB	An ECB digital currency – a flight of fancy	Negative – 88.6%	https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200511~01209cb324.en.html
ECB	Payments in a digital world	Positive – 81.5%	https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp200910~31e6ae9835.en.html

Institution	Speech title	Sentiment	Link
FRB	Fast Payments for All	Positive – 98.5%	https://www.federalreserve.gov/newsevents/speech/brainard20190805a.htm
FRB	Digital Currencies	Positive – 59%	https://www.federalreserve.gov/newsevents/speech/brainard20191016a.htm
FRB	An update on Digital Currencies, Stablecoins, and the Challenges Ahead	Positive – 99.9%	https://www.federalreserve.gov/newsevents/speech/brainard20191218a.htm
FRB	The Digitalization of Payments and Currency	Positive – 98.9%	https://www.federalreserve.gov/newsevents/speech/brainard20200205a.htm
FRB	Direction of Supervision: Impact of Payment System Innovation on Community Banks	Positive – 99.9%	https://www.federalreserve.gov/newsevents/speech/bowman20200227a.htm
FRB	The Future of Retail Payments in the United States	Positive – 96.7%	https://www.federalreserve.gov/newsevents/speech/brainard20200806a.htm
FRB	Update on Digital Currencies	Positive – 94.7%	https://www.federalreserve.gov/newsevents/speech/brainard20200813a.htm
HKMA	Opening Remarks at the Hong Kong-Paris Financial Seminar	Positive – 97.8%	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/01/20190116-1/
HKMA	Speech at the ISDA 34th Annual General Meeting in Hong Kong	Positive – 97.4%	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/04/20190410-1/
HKMA	Keynote Speech at HKAB Distinguished Speaker Luncheon	Positive – 92.8%	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/04/20190410-1/
HKMA	Keynote Speech at HKMA Fintech Week	Neutral – 55.8%	https://www.hkma.gov.hk/eng/news-and-media/speeches/2019/11/20191106-1/
MAS	Payment Services Bill	Positive – 68.9%	https://www.mas.gov.sg/news/speeches/2019/payment-services-bill
MAS	Is the Blockchain a Solution Looking for a Problem?	Positive – 91.3%	https://www.mas.gov.sg/news/speeches/2019/is-the-block-chain-a-solution-looking-for-a-problem
MAS	Combatting Financial Crime through New Technologies Built on Strong Fundamentals	Positive – 97.5%	https://www.mas.gov.sg/news/speeches/2019/combatting-financial-crime-through-new-technologies-built-on-strong-fundamentals

REFERENCES

Bank of International Settlements, 2020. *Rise of the central bank digital currencies: drivers, approaches and technologies,* August.

Bloomberg, 2020. *AI-Powered Fund Thrashing S&P 500 Is Wary of Cyclical Bounce*, October.

Board of Governors of the Federal Reserve System, 2017. *Sentiment in Central Banks' Financial Stability Reports,* March.

BBC, 2019. *Revolut whistleblower had concerns over CEO conduct and compliance*, April.

CNBC, 2019. *NY Fed clarifies Williams speech that market took as signal of a rate cut,* July.

European Central Bank, 2012. *Speech by Mario Draghi, President of the European Central Bank, at the Global Investment Conference*, July.

Financial Stability Board, 2020. *Enhancing Cross-border Payments - Stage 1 report to the G20*, April.

PYMNTS.com, 2020. *Wirecard Linked To Mafia Money Laundering*, August.

The SEACEN Centre

Since its inception in the early 1980's, The South East Asian Central Banks Research and Training Centre (the SEACEN Centre) has established its unique regional position in serving its membership of central banks in the Asia-Pacific region through its learning programmes in key central banking areas (including Macroeconomic and Monetary Policy Management; Financial Stability and Supervision, and Payment and Settlement System; and Leadership and Governance), research work, and networking and collaboration platforms for capability building in central banking knowledge.

The SEACEN Centre

The South East Asian Central Banks (SEACEN) Research and Training Centre