# Russia, the Federal Reserve, and the *New York Times*: Machine-Learning Models of How Elites Interpret Policy Announcements*

David Sylvan†        Jean-Louis Arcand‡        Ashley Thornton§

April 2, 2021

### Abstract

How do elites understand foreign and monetary policy? In the past, scholars have explored this question by modeling it as a psychological mapping process: elites focus on particular features of policies and map them onto particular psychologically relevant categories, such as adversary or danger, each of which usually has one or more action implications. The problem with understanding policy interpretation in this way is that it requires specifying multi-level rules by which particular words or phrases, in order to be mapped onto categories, are first mapped onto context, with the contexts themselves being mapped onto higher-level contexts, and so on ad infinitum. As an alternative, we reconceive of policy interpretation as a matter of nonhierarchical textual entailment, in which different combinations of syntactic and semantic information will be associated with stylized accounts of what is, or might, or should not be happening and why. Specifically, we use machine learning techniques to generate newspaper interpretations of two streams of policy announcements: one, from the White House and State Department about Russia, and the other, from the Federal Reserve, about monetary policy. Preliminary results suggest that our approach is good, in a non a priori way, at capturing the hypothesized issue domain specificity of foreign policy as opposed to monetary policy.

Keywords: Foreign policy, political economy, methodology

## 1   Introduction

This paper, and the project underlying it, starts from the observation that state agencies regularly issue announcements about the policies which they are carrying out, and that those announcements have significant political and economic consequences even before the policies can be partly or completely implemented. These consequences are quite varied, depending, but only to some degree, on the type of announcement and, more generally, the issue domain (a point to which we return below). Thus, investors may react to certain announcements about

†Department of International Relations / Political Science, Graduate Institute of International and Development Studies, Geneva, Switzerland. Email: david.sylvan@graduateinstitute.ch

‡Department of International Economics, Graduate Institute of International and Development Studies, Geneva, Switzerland. Email: jean-louis.arcand@graduateinstitute.ch

§Intrepid Project, Graduate Institute of International and Development Studies/IDIAP, Switzerland. Email: ashley.thornton@graduateinstitute.ch

interest rates or monetary policy stance (such as quantitative easing), or for that matter to announcements about the escalation of a military standoff, by buying or selling stocks, bonds, or other financial instruments; diplomats may shift or reinforce their negotiating positions in reaction to an announcement about an impending vote in the Security Council; lobbyists may concert their efforts to overturn a given policy in reaction to any number of different types of announcements; and of course, members of the public may decide to vote for or against a given political party on the basis of what is, or is not, being said by spokesmen for the governing coalition. All of these specific phenomena have been studied for many years.

What those studies miss, however, is the precondition necessary for each of these phenomena to occur, namely, that for state agencies' announcements to be reacted to, they must first be interpreted and that those interpretations are carried out, more often than not, by individuals specialized in following those agencies. For example, stock brokers have neither the time nor perhaps the technical command of macroeconomic theory to read through a 20-page memorandum in which the Federal Reserve's Open Market Committee (FOMC) details what went on at its monthly meeting some six weeks earlier. Instead, they have to rely on wire service and analysts' reports or claims by experts about what the Fed's announcement means: what the Fed is doing now, or planning on doing in the future, or will not be doing, what its priorities are, and so forth. Similarly, most diplomats do not necessarily read through every word of the White House press secretary's daily briefing; instead, they read the *New York Times* articles on that subject as well, of course, as memos by colleagues charged with keeping tabs on the White House (the latter, in fact, usually rely on the press as well). Our concern here is with these necessary and ubiquitous acts of interpretation by various specialists, particularly journalists assigned to the coverage of particular state agencies.

In our view, it is impossible to imagine the modern state without both announcements and interpretations of those announcements. On the one hand, the range of activities dealt with by the state is far greater than a century ago, with many of those activities being carried out by multiple agencies and, to boot, being highly technical in nature. On the other hand, those same activities are aimed directly at numerous "attentive" publics: from investors and officials of foreign governments to legislators and potential voters. In this regard, journalists are privileged interlocutors, focused on intently by politicians and agency personnel, with the announcements being crafted for maximum impact and with those announcements themselves being the object of additional announcements (e.g., "off the record" briefings of journalists about press releases). Journalists, in this sense, are interpreters who, on a daily basis, select, gloss, and repackage state announcements. This paper is a progress report about a research project on that interpretation process.

Specifically, the paper is divided into several parts. We begin by discussing, in an abstract fashion, the concepts of announcement and interpretation, and argue for an alternative way of researching interpretation. Ever since Lippmann's seminal work (1922) on public opinion, the standard account of how journalists carry out interpretation is that they engage in mapping, taking announcements and rewriting them, so to speak, as a series of claims about what is, or what might be, or what will not be, happening, and why. This mapping is domain-specific, so that, for instance, an announcement by a central bank is assumed to be conditioned by fear of

inflation, whereas an announcement about a military ally is assumed to be directed implicitly at an actual or potential adversary. There are, however, several distinct problems with this account of journalistic interpretation, and thus we propose a non-mapping approach, based on textual entailment and machine learning (ML). That approach, we further note, implies a considerably different take on the notion of domain specificity.

We then turn to the details of our alternative, laying out the data construction of both interpretations and announcements, as well as the entailment-based method of assessing how the former are linked to the latter. This enables us to revisit the issue of domain specificity, a topic on which we put forward several proto-hypotheses. Although the data for evaluating both the validity of our proposed interpretation-announcement link, as well as the proto-hypotheses on domain specificity, are only partially constructed, preliminary analysis, particularly on the issue of domain specificity, is intriguing.

## 2    Interpretation: mapping vs. entailment

To start with, some terminology. A **policy announcement** is a written statement issued by a state agency describing its actions or decisions. These announcements can take the form of written communiques, press releases, statements read before the media, background briefings (sometimes in anonmous form), and answers to journalists' or legislators' questions (the questions and answers are often transcribed and subsequently released in written form). Many state agencies, in many countries, make regular policy announcements, often on a pre-scheduled basis, such as weekly, monthly, or quarterly. These announcements may be made out of a sense of obligation (indeed, in some cases, they may be legally mandated, as in the case of the Humphrey-Hawkins Full Employment Act in the United States, which instructed the Board of Governors of the Federal Reserve to transmit a Monetary Policy Report twice a year to the Congress) or out of a desire to send signals to different political and economic actors. While individual agencies in different states vary considerably in the frequency, formality, and extent of their announcements, when they do make such announcements,[1] they nonetheless do so on a regular basis for many years. As we will see below, this regularity makes it possible to construct corpora from streams of announcements and, at least for certain agencies, for those corpora to be sufficiently dense as to make ML techniques feasible.

An **interpretation** is a statement made about a policy announcement. Interpretations can be made by journalists, politicians, state officials, academics, analysts, experts, and "talking heads" of various sorts. Although interpretations may originally be made in oral form and thus can be listened to, many interpretations are written, as, for example, newspaper articles; indeed, many oral interpretations, such as legislative speeches, are at the least available as transcriptions and may, originally, have begun as written texts which were then read aloud.

---

[1]For example, just comparing France with the United States, in the latter, the White House has press briefings several times a week and the FOMC issues statements about its meetings 8 times a year; in the former, the spokesman for the prime minister presents a "compte-rendu" of the Conseil des ministres once a week, whereas when the Banque de France was responsible for French monetary policy, it only issued a "compte-rendu" once a year.

Given the significance of policy announcements, it is almost always the case that each such announcement generates a number of different interpretations: multiple news stories, political commentaries, analyst reports, or even, long after the announcements were made, academic studies. As we argued above, journalists' interpretations are of particular interest; and as we will see below, if individual announcement corpora can be aligned with associated interpretation corpora, such as streams of newspaper articles about the announcements, we thus will have naturally-occurring training data for ML models of interpretation.

One way to conceptualize the link between announcements and interpretations is to see the latter as a mapping of the former onto an already-structured domain. For example, if the chairman of the Federal Reserve gives a press conference and says that "It's just a lot of people who need to get back to work, and it's not going to happen overnight, it's going to take some time," and if a journalist glosses this as the Fed being "willing to allow inflation to move higher without reacting" (*New York Times*, 17 March 2021), then the journalist can be seen as mapping the chairman's language about people needing to get back to work onto a domain characterized by a presumed tradeoff between inflation and unemployment, an additional presumed Fed preference for low levels of inflation, a further presumed Fed default policy of raising interest rates to maintain those low levels, and finally, in that domain, a surprising influence (hence, "news"): that the chairman's words seemed to be signaling that the Fed is acting in a way opposite to its default policy.

This way of thinking about interpretation as mapping is reminiscent of classic work on belief systems as "an interrelated set of affect-laden cognitions concerning some aspects of the psychological world" (Abelson and Carroll 1965: 24)(see also Abelson 1971, 1973, 1979), such that particular events are mapped onto instances of generic states of affairs (e.g., North Korea attacks South Korea in June 1950 is an instance of Communist aggression against the Free World), with the default action implications of the latter (e.g., if there is Communist aggression against the Free World, then stand firm) being remapped onto specific policy positions in this particular situation (e.g., send troops to South Korea as an instance of standing firm). Unsurprisingly, the mapping approach to interpretation has been picked up in computer science (e.g., Carbonell 1978), political psychology (e.g., Jost, Federico and Napier 2009), foreign policy analysis (e.g., Taber 1992), and, of particular note for this discussion, studies of journalism (e.g., Gans 1979; Herman and Chomsky 1988; Gamson and Modigliani 1989; Carvalho 2007; Matthes 2009).

There are two interrelated problems with thinking about interpretation as mapping. The journalist is acting akin to a function that goes from the words of the announcement the rewrite of the news article, but what actions or perceptions by the journalist would that function entail? We might imagine that the journalist is using rules of some sort: for example, if a Fed chairman talks about pressure on wages, he is referring to inflation, whereas if he talks about the economy being weak, he is referring to unemployment. However, those rules will not be all-purpose, because the chairman could specify that in the current context, pressure on wages is likely to be short-term and limited. One would thus add a second set of contextual rules; but this does not solve anything, because if there are multiple possible contexts, then one needs meta-contextual rules, ad infinitum. Quite apart from the practical impossibility of storing and

searching among even a large number of rules, the more serious problem is that there is no evidence at all to suggest that many contexts had any rules specifying the conditions of their use: which financial journalist, prior to 2020, could possibly have had rules for how to interpret Fed action during a worldwide pandemic?

Note that the contextual rules problems that arise in thinking about interpretation as mapping revolve around the issue of domain specificity. Consider, for example, a statement by the Fed chairman about "tightening": normally, this would be mapped onto a monetary policy issue domain, concerned with interest rates and the money supply. But if the comment occurs when talking about bank supervision, or exchange rate policy, or financial sanctions against officials of foreign governments, then the issue domain, and hence the meaning of "tightening," would be different. Indeed, even in the context of monetary policy, if the Fed were trying to induce banks to lend more freely, had made funds available for low-interest loans, and discovered that banks were not in fact lending as much as desired, it might "tighten" the conditions under which it made those funds available to banks. Thus, the practical impossibility of specifying a set of rules that both define all conceivable contexts for mapping terms and the conditions under which one (or more?) contexts apply is due to the fact that interpretation is domain-specific. Perhaps unsurprisingly, even studies which directly address the content and differences between domains (e.g., work on semantic frames: Fillmore, Johnson and Petruck 2003) tend to proceed in either an extremely abstract (e.g., in Framenet, Attaching, or Communication) or generic (e.g., conflict, economic consequences, human interest, or morality: Semetko and Valkenburg 2000) fashion, or else simply omit "most of the words [e.g., 'blatant'] that one confronts in naturally occurring text" (Pavlick et al. 2015: 408).[2]

In fairness, it should be pointed out that these problems, which have been known in the field of artificial intelligence for decades (e.g., Davis and Marcus 2015), have sparked numerous proposals, ranging from "case-based" or analogical approaches to approaches based on combining "common-sense" reasoning with neural networks (e.g., Hwang et al. 2020). Our approach, though closer to the second, is based critically on textual entailment, i.e., the semantic implications between texts (Dagan et al. 2006). Recall the above example about Fed "tightening." Here, what a journalist would presumably do is to use a) the everyday meaning of "tightening," b) sentences before and after about what will not be done, c) syntactic information about the modality of the verb tighten, its direct object, and its target, and d) prior information about what the Fed is currently doing about monetary policy, and, perhaps, about the Fed chairman's presumed preferences, in order to construct a summary of the who-what-when-why-type information expressed in the phrases or sentences about tightening. However, since the text of the chairman's statement contains large amounts of syntactic and semantic information, even when that statement is not trying to be vague or obfuscatory,[3] the number of combinatorial possibilities for the use of that information is simply too great to be specified a priori.

---

[2]Similar problems are apparent in the foreign policy literature on "substitutability" (e.g., Most and Starr 1984; Palmer and Bhandari 2000).

[3]The former chairman of the Federal Reserve Board, Alan Greenspan, famously said that "I guess I should warn you, if I turn out to be particularly clear, you've probably misunderstood what I said." Much the same can be said of diplomatic announcements: officials spend considerable amounts of time using coded language designed to be understood by insiders but not necessarily by others, as in the Treaty of Breda (10 CTS 231) ending the Second Anglo-Dutch War, which famously ceded the island of Manhattan without ever saying so.

Instead, to capture how journalists key off of multiple textual elements to interpret announcements, we propose to combine human coding to specify particular interpretations in news articles about announcements, then to use ML techniques, trained on coding-announcements pairs, to reproduce those pairs and generate others. Below, we will discuss this approach in some detail, but note for now that the issue of domain specificity is to a considerable degree obviated: the researcher need not establish in advance a taxonomy of interpretive issue domains, nor the features of any such domains. To be sure, a wildly heterogeneous corpus of news articles may make it more difficult both to code and to use ML techniques on that coding, but that is an empirical question, and in fact, the approach enables human coders' intuitions about domain specificity (say in sorting articles as between those on the Fed and those on relations with Russia) to be checked, an issue to which we return at the end of this paper.

As indicated above, the ML techniques will use both semantic and syntactic information. To explain why, consider two examples of journalistic interpretation, one about a Federal Reserve press release, the other about a State Department background briefing on Syria.

> "The F.O.M.C. meeting ended at 12:55 PM; there is no further announcement." (Federal Reserve press release, December 22, 1998)

> "The Federal Reserve voted today to hold interest rates steady, judging that three rate reductions in the fall had given the economy a big enough kick to keep it growing at a healthy pace into the new year." (*New York Times*, December 23, 1998)

In this example, we can see that the journalist is using information about what the F.O.M.C. can do (e.g., raise interest rates) as well as what it usually does (i.e., make no announcement when it is neither raising nor lowering rates); the journalist also uses information about the world (how the economy is doing) and what the Fed did over the last few months (three reductions in the fall).

> "We already, as you know, provide non-lethal aid. We do everything we can, working with the London 11 and a group of partners, some of whom provide other things to the opposition coalition. And we will continue, as Secretary Kerry said, as President Obama said, standing next to President Hollande to look at every option that is open to us to see what else we can do to be helpful." (State Department background briefing on Syria, February 14, 2014)

> "Diplomats here said the administration might consider stepping up an existing covert program to train and arm the moderate Syrian opposition or even weigh the threat of military force to compel the delivery of humanitarian aid. The senior official declined to say whether a policy shift was underway, saying options were always being reviewed." (*New York Times*, February 15, 2014)

In this example, we can see that the journalist is not only using semantic information (e.g., what the U.S. and its allies are doing in Syria; what "helpful" means) but also syntactic information (the temporal scope of "will continue ... to look"). It is precisely this type of information that we want to employ, and that standard "bag of words" methods will tend to miss.

# 3 An alternative approach

The two aspects of our alternative involve human construction of data and machine learning of the relation between announcements and interpretations. Originally, the research project was designed to explore the announcement-interpretation link for two issue areas (monetary policy and foreign policy toward Russia/the Soviet Union), three countries (the United States for both issue areas, France for foreign policy, and Canada for monetary policy), both left- and right-of-center newspapers (the *New York Times* and the *Wall Street Journal*; *Le Monde* and *Figaro*; and *La Presse* and the *Globe and Mail*, respectively), and streams of announcements and interpretations that stretched from 1967 to 2017. Unsurprisingly, issues of data construction turned out to be quite involved, and so far, monetary policy data have been created for the *New York Times*, with U.S.-Russia policy data currently being created for the same newspaper.

## 3.1 Data construction

For each case, data construction involves three facets: assembling a corpus of newspaper articles on the issue area; assembling a corpus of announcements interpreted in those articles; and constructing a corpus of the articles' interpretations.

### 3.1.1 Newspaper articles

To collect candidate *New York Times* articles for the Fed case, we used a combination of two bi-grams, Federal Reserve and monetary policy, in Lexis-Nexis to download articles that might make references to Fed announcements. (Articles prior to 1980 and thus not in Lexis-Nexis required a modification of this approach.) This search strategy was motivated by limiting the number of false positives, i.e., articles related to monetary policy but not containing a reference to a current monetary policy announcement.

Once the articles were downloaded, the next step in the workflow was to human-audit the articles and keep only those containing a direct reference to a Fed announcement. We used a calendar-based approach as most candidate true positive articles are found within 2 days of a scheduled announcement. We also looked for particular terms in the article such as the FOMC or the name of the Fed chairperson at the time, which was a high probability marker of an article containing a reference to a Fed policy announcement. A second and third human audit were subsequently carried out by an additional team member.

Once a list of true positive articles was assembled, the metadata for those articles were examined and opinion pieces, editorials, and (usually) extended articles in the Sunday magazine were eliminated. An additional check, done by skimming over the articles themselves on the *New York Times* website, was carried out to eliminate articles which were simply transcripts of interviews, speeches, or testimony. In the end, the corpus (51 years long) contains just under 5,500 articles. After this, the articles were transformed into uniformly formatted .txt files, which subsequently were re-read to construct a master list (itself placed in the dropbox) of the articles for that year, with the entry for each article containing pointers to the announcements mentioned in the article (see below).

We are currently using a two-pass search strategy for the U.S.-Russia case. We begin with

a general search with the keywords United States AND Russia AND policy, then merged the results with a specific keyword search: White House or State Department or President's name or Vice-President's name or Secretary of State or Defense Secretary or National Security Advisor or Assistant Secretary of State for Eurasian Affairs or any <key office bearer> AND Russia.[4] This two-part search ensures a reasonably extensive picture of both general foreign policy coverage and specific foreign policy coverage related to U.S. office-bearers. Following the search, a human audit is performed, along the lines of the Fed case, to see whether the articles actually touch on U.S. policy regarding Russia (for example, about what either the U.S. or other countries are doing in reaction to Russian state actions), as opposed to articles about other events, such as hockey games or power struggles in the Kremlin on which there is no U.S. announcement.

### 3.1.2 Announcements

To assemble corpora of announcements, we began by finding the URLs of websites pertaining to various state agencies. For the Federal Reserve, this is easy (one regional Fed website serves as a sort of repository for speeches and publications not found on the main Fed website); but for the White House and various cabinet departments, it is complicated by the fact that websites from prior presidential administrations are archived, and not easily searchable. By the same token, search engines for congressional hearings, for electronic news media (even though we are excluding video- and audio-only records), and for various nongovernmental organizations (who often invite officials to speak before them) are somewhat catch-as-catch-can. Not all statements and press events are themselves conserved on even official government websites, and so we also found ourselves supplementing coverage with collections available on embassy websites or later Fed publications. We have not yet begun to develop procedures for gathering announcements which have themselves not been digitized. It should also be noted that the sheer length of certain announcements (e.g., the transcript of a 3-hour congressional hearing) presents particular challenges for ML parsing.

Obviously, there are large numbers of announcements that are never written about, just as there are some articles for which the announcements cannot be found. The latter, though annoying, do not pose a problem, as we will discuss below; but the former are sufficiently extensive that it makes no sense to aim at collecting them all. Instead, we perform a two-pass process: first, when the master list of articles for a given year is constructed, a team member retrieves the announcements that seem to be referred to in the articles; and subsequently, when the articles are annotated to construct the interpretations, the annotators check on the accuracy of the pointers to those announcements and add to or revise them.

More generally, it should be noted that the second announcement corpus constructed for each case is explicitly linked to the article corpus. Such linkage is necessary to construct interpretation corpora (see below), but it is also a built-in relevancy criterion simply assumed in any nunber of claims about putative reactions: for example, stock market behavior following Fed interest rate increases, or deterrence models. (For example, Wall Street traders may not, for the most part, actually read a Fed press release: they may be reacting to an interview with an expert, or

---

[4]It should be noted that there are significant coverage differences between searching for an individual's full name and only his/her family name.

8

to each other's behavior. By the same token, generals or their superiors may not be attending to either the announcements or the latest moves of their adversaries, but instead executing pre-set scenarios triggered by an assessment of overall tension.) In one sense this eliminates the possibility of modeling the way in which journalists filter the information that streams across their desks, the vast majority of which, we assume, is considered by them as irrelevant. In another sense, though, the built-in relevancy means that the interpretation-construction task can be far more focused.

### 3.1.3 Interpretations

To construct a corpus of articles' interpretations, we annotate each article, looking for portions of text that seem to be about announcements. Each such portion is in turn tagged with minimal journalistic information: a) a pointer to the document containing the announcement, a textual sub-passage containing b) the journalist's identification of who (e.g., the Fed or the U.S.) is either acting or whose (possible) action is being discussed, another sub-passage containing c) the nature of the action (e.g., raising interest rates; imposing sanctions), and, d) in the case of policy toward Russia, a final textual sub-passage containing an explanation of why the act in question is relevant to U.S. policy toward Russia (Russia Link). In addition, there are optional fields in the textual portions: e) evidence (e.g., the rate of inflation) flagged by the journalist, from the announcement, as justifying the act; f) motivation for the act (what was intended to be accomplished); g) temporal scope (when the act would go into effect or how long it would last); and h) attribution (when an actor other than the Fed would give his/her opinion about what should/should not be done, or have been done; this field is rarely used in the US-Russia case, presumably because Executive Branch announcers serve at the pleasure of the president). Since in many cases, the journalist's wording may be obscure, or because she may be assuming knowledge that is not explicitly stated, annotators add their own characterization of both the action and the Russia Link; they may also add their own characterization of the other fields.

As can be seen, these interpretations are act-focused, not only on what was, is, or will be done, but also on what should or should not be done. Since a single announcement can refer to more than one act, it is not only possible but frequent for an article to contain multiple interpretations of the same announcement (with the interpretations perhaps overlapping significantly, depending on the length of the article), and indeed for that announcement to be interpreted in multiple articles, just as it is also possible for individual articles to interpret multiple announcements. It should be noted that these types of multiplicity have several features, whose significance will be discussed below: in the article, there may be references to more than one announcement by a single individual or agency; there may be references to announcements by two or more individuals or agencies; and the references may cover not only announcements made the day of, or the day before, the article's publication, but announcements made days, weeks, months, or even years before. As we will discuss below, this last type of lag is of particular interest, signaling far more than just the range of the journalist's memory or the extent of her files.

## 3.2 Machine learning

For a given case, the data construction process thus produces three corpora: newspaper articles, interpretations of those articles, and announcements referred to in the interpretations. The task of the ML is, using both syntactic and semantic information from the announcements, to generate the interpretations, with particular attention being paid to the annotators' characterization entries and not only to the portions of article text contained in certain fields. Several observations about the ML aspect of the project should be noted.

First, the link between announcements and interpretations is, fairly often, one-to-many. As noted above, a given announcement can be referred to more than once in a given article, as well as in multiple articles. This kind of multiplicity ought not to be understood as a sign of inexactitude (although the annotators were occasionally struck by what appeared to be journalistic sloppiness), or even of differences of opinion as to the nature of the act, or of motives or evidence. Rather, multiple interpretations of the same announcement differ most commonly in the amount of detail they provide, as well as in the explicitness and vividness of the wording.

Second, both the number of items (articles, interpretations, announcements) in any given corpus, as well as the length of those items (number of words) is generally quite small, at least by both ML and computational linguistics standards. Although, as noted above, the occasional hearing transcript may stretch on for scores of pages, in most cases, announcements rarely exceed 8 pages double-spaced. The fact that by social science data construction criteria the corpora are large, with thousands of items in each corpus, and at least the Fed case, with the items stretching for over half a century, ameliorates matters relatively little from an ML perspective. It is for this reason that it is important to develop robust semantic (and perhaps pragmatic) information, so that terms' referents can be clearly specified and the recent or typical expectations associated with those referents made explicit.

Third, it is important to understand the biases inherent in using newspaper articles as the basis for interpretations. Journalism involves "deciding what's news," to use the title of the Gans book (1979) mentioned above, which is to say that continuation of the status quo ante is not normally written about. This may be one reason why, when nothing new happens on U.S.-Russia relations, there are few articles about the announcements of routine activities. On the other hand, even when the Fed decides not to change interest rates, that is still news, albeit less written about. As we will see below, this points to a possible distinction in the two issue areas; for now, it means that journalists are faced with less of a filtering issue regarding Fed announcements than in the case of U.S.-Russia policy; from an ML perspective, the latter are more complex than the former.

## 4    Domain specificity revisited

To date, progress has been quite disparate. Data construction for the Fed case was finished in the autumn of last year, and ML efforts for that case are now underway. The U.S.-Russia case is at a considerably earlier stage, with data construction for the Obama years around half-done. This means, for now, that we do not have ML results to report. What we can do, however, is to

revisit the issue of domain specificity, a topic of particular interest in numerous scholarly fields, not least of them the study of politics.

We indicated above that there were significant problems with specifying a priori the features of, and distinctions between, issue domains. However, many of the same problems apply to standard "bag of words" approaches to identifying domains a posteriori, such as is commonly the case in topic modeling (Spirling 2012; Nowlin 2016; Greene and Cross 2017; Vayansky and Kumar 2020). Recall that topic modeling is based on covariance between words, with the analyst then assigning labels to particular covariance clusters. If the concern is simply to distinguish specialized vocabulary characteristic of a domain, this is not unreasonable: it is no more likely that terms such as "federal funds rate" will be used in Defense Department announcements than that terms such as "carrier strike group" be used in Federal Reserve announcements. However, unless there are equally specialized short verb phrases (cf. Sajid, Jan and Shah 2017), topic modeling is of much less use to uncover the activities characteristic (much less uniquely so) of domains, not to mention the way in which motives are ascribed to acts or individual actors described as pursuing certain act agendas. It is for this reason that our search procedure for articles on either the Fed or U.S.-Russia policy was quite minimalist, with the net being cast widely and criteria for inclusion fairly minimal.

Nonetheless, even before we have a trained ML model whose accuracy can be compared across cases, we can make several preliminary observations about the difference between our two cases. All the annotators, as well as those doing the newspaper searches, were struck by how different newspaper articles (and interpretations in them) were for the Fed as compared with U.S.-Russia policy. The latter articles were longer, and far more intricate, as was the way they dealt with announcements, and with the past. We can stylize these observations as being about the multiplicity of foreign policy articles as opposed to monetary policy ones. There are two potential issues here.

First, the Federal Reserve is, technically speaking, a collective body. Although there is a chairperson who usually is able to sway others to his or her views, those others in fact have a vote, which means that they can chime in with their opinion; for that matter, the Fed chairperson can try to influence the other members of the FOMC via his/her announcements. This is quite different from U.S.-Russia policy, which is under the formal control of the President. Dissenters typically lobby within the Executive Branch, and even leaks to the media are potentially dangerous for the leaker. On the other hand, the Fed is a single agency, with no other organizations under its control (technically, the Fed does not even give orders to regional Fed branches), a position which is at the opposite extreme from the Executive Branch. U.S. policy toward Russia, or indeed toward any other country, may and often does involve acts by numerous U.S. government agencies: the State Department, the Defense Department, perhaps the Treasury Department and (though this is less often announced), the CIA; by the same token, announcements can also come from multiple officials and spokespersons within different organizations (including the White House). We thus have a first problem for the journalist, which is that even though foreign policy is considerably more unitary from a decisional point of view than the Fed's monetary policy, it is carried out by many more actors.

A second issue has to do with time. When the Fed acts, it more often than not supersedes

its previous policy: for example, if it raises interest rates from 4 percent to 4.5 percent, the former rate is no longer in operation. This does not mean that the journalist might cease referring to the earlier rate, but then the story is likely to be retrospective, concerned with how policy evolved or why it took so long to do so. By contrast, because of the number of different agencies involved and the difficulty in getting them to act simultaneously, specific foreign policy acts are likely to emerge over a period of time, especially if the act was signaled as a possibility contingent on action by another country (the Fed also signals, but tries not to change policy several times in a short interval). This means, then, that the journalist has a second problem, which is to indicate how acts occurring over a period of time are indicative of a single policy.

Multiplicity problems are of course not only faced by journalists, but by authors of fiction as well. Novelists, playwrights, and screen writers need to establish a single narrative thread, even though there may be multiple characters in the work, existing simultaneously in different places. Some years back, Woloch (2003) argued that 19th century European fiction was able to establish a unity of narrative perspective through a series of looped-back secondary and tertiary characters. His work was picked up by numerous scholars (Sack 2012; Marazzato and Sparavigna 2014; Min and Juyong 2016; Grayson et al. 2016; Mourchid et al. 2018; Labatut and Bost 2019; Jones, Quinn and Koskinen 2020), who showed how various forms of character co-occurrence, from interacting in the same scene to mere mention or recollection (with non-interacting physical proximity an interesting middle ground) helped to establish networks centered around protagonists and/or escalating in temporal compression.

Although character interaction in newspaper articles is almost always more skeletal than fictional scenes, the above work on character co-occurrence suggests that it may be possible to see analogues to such occurrence in articles and to use them as indicators of how the two multiplicity problems are solved. We therefore suggest two hypotheses about differences between articles about the Fed and about U.S.-Russia policy. First, the number of distinct announcements (as a sort of stand-in for the number of agencies/characters) should be less in the former corpus than in the latter; and second, the temporal gap between the oldest cited announcement and the date of the article should be less in the former corpus than in the latter. The logic behind the first of these hypotheses is that journalists will establish the existence of a single policy by citing announcements from multiple officials and/or agencies; by the same token, officials will also establish the existence of a single policy by casting their temporal net more widely.

Although we have not yet annotated enough years of U.S.-Russia articles to do a broad comparison with the Fed ones, we are able, purely for exploratory purposes, to display data on the two hypotheses for a particular year, namely 2015. As the table shows, the average article in the U.S.-Russia corpus is in fact characterized by a somewhat greater number of distinct announcements (mean of 2.03 vs. 1.92), as well as by a considerably greater temporal gap between the date of the article and the oldest announcement interpreted in the article (mean of 101.79 days vs. 17.78 days). We would expect that the former difference will grow as we move back in time, when the number of articles about U.S.-Russia relations grows (around the time of the annexation of Crimea and of confrontations in Syria).

In short, although the project is as yet not far enough along to train and evaluate an ML model of the announcement-interpretation link, there appear to be discernible differences,

relevant to that link, between the two corpora we are assembling. One need not a priori specify interpretive domains in order to explore them.

## 5 Conclusion

Newspaper interpretation of policy announcements can be understood as a matter of textual entailment, rather than a rule-based mapping of statements from one structured domain to another. We have laid out a hybrid human coder/ML training methodology to explore that entailment, and have begun to assess it using two cases. Preliminary evidence to date suggests that the methodology is a feasible way of understanding the announcement-interpretation connection.

## References

**Abelson, Robert P.**, "The Ideology Machine," 1971. Annual Meeting of the American Political Science Association.

_ , "The Structure of Belief Systems," in Roger Schank and Kenneth Mark Colby, eds., *Computer Models of Thought and Language*, New York: W.H. Freeman & Co., 1973.

_ , "Differences Between Belief and Knowledge Systems," *Cognitive Science*, 1979, *3* (4), 355–366.

_ **and J. Douglas Carroll**, "Computer Simulation of Individual Belief Systems," *American Behavioral Scientist*, 1965, *8* (9), 24–30.

**Carbonell, Jaime G. Jr.**, "POLITICS: Automated Ideological Reasoning," *Cognitive Science*, 1978, *2* (1), 1–15.

**Carvalho, Anabela**, "Ideological Cultures and Media Discourses on Scientific Knowledge: Re-reading News on Climate Change," *Public Understanding of Science*, 2007, *16* (2), 223–243.

**Dagan, Ido, Oren Glickman, and Bernardo Magnini**, "The PASCAL Recognising Textual Entailment Challenge," in Quionero-Candela J., Dagan I., Magnini B., and d'Alché-Buc F., eds., *Machine Learning Challenges: Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Textual Entailment*, 2006, pp. 177–190.

**Davis, Ernest and Gary Marcus**, "Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence," *Communications of the ACM*, August 2015, *58* (9), 92–103.

**Fillmore, Charles J., Christopher R. Johnson, and Miriam R.L. Petruck**, "Background to FrameNet," *International Journal of Lexicography*, 2003, *16* (3), 235–250.

**Gamson, William A. and Andre Modigliani**, "Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach," *American Journal of Sociology*, July 1989, *95* (1), 1–37.

**Gans, Herbert J.**, *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*, New York: Random House, 1979.

**Grayson, Siobhán, Karen Wade, Gerardine Meaney, and Derek Greene**, "The Sense and Sensibility of Different Sliding Windows in Constructing Co-occurrence Networks from Literature," *2nd International Workshop on Computational History and Data-Driven Humanities (CHDDH)*, May 2016, pp. 65–77. Dublin, Ireland.

**Greene, Derek and James P. Cross**, "Exploring the Political Agenda of the European Parliament Using a Dynamic Topic Modeling Approach," *Political Analysis*, 2017, *25*, 77–94.

**Herman, Edward S. and Noam Chomsky**, *Manufacturing Consent: The Political Economy of the Mass Media*, New York: Pantheon, 1988.

**Hwang, Jena D., Chandra Bhagavatula, Ronan Le Bras, Jeff Da, Keisuke Sakaguchi, Antoine Bosselut, and Yejin Choi**, "COMET-ATOMIC 2020: On Symbolic and Neural Commonsense Knowledge Graphs," *arXiv preprint arXiv:2010.05953*, 2020.

**Jones, Pete, Eithne Quinn, and Johan Koskinen**, "Measuring Centrality in Film Narratives Using Dynamic Character Interaction Networks," *Social Networks*, October 2020, *63*, 21–37.

**Jost, John T., Christopher M. Federico, and Jaime L. Napier**, "Political Ideology: Its Structure, Functions, and Elective Affinities," *Annual Review of Psychology*, 2009, *60*, 307–337.

**Labatut, Vincent and Xavier Bost**, "Extraction and Analysis of Fictional Character Networks: A Survey," *ACM Computing Surveys*, 2019, *52* (5). Article 89.

**Lippmann, Walter**, *Public Opinion*, New York: Harcourt, Brace and Company, 1922.

**Marazzato, Roberto and Amelia Carolina Sparavigna**, "Extracting Networks of Characters and Places from Written Works with CHAPLIN," *arXiv:1402.4259*, 2014.

**Matthes, Jörg**, "What's in a Frame? A Content Analysis of Media Framing Studies in the World's Leading Communication Journals, 1990-2005," *Journalism & Mass Communication Quarterly*, 2009, *86* (2), 349–367.

**Min, Semi and Park Juyong**, "Mapping Out Narrative Structures and Dynamics Using Networks and Textual Information," *arXiv:1604.03029*, 2016.

**Most, Benjamin A. and Harvey Starr**, "International Relations Theory, Foreign Policy Substitutability, and Nice Laws," *World Politics*, April 1984, *36* (3), 383–406.

**Mourchid, Youssef, Benjamin Renoust, Hocine Cherifi, and Mohammed El Hassouni**, "Multilayer Network Model of Movie Script," *arXiv:1812.05718v1*, 2018.

**Nowlin, Matthew C.**, "Modeling Issue Definitions Using Quantitative Text Analysis," *Policy Studies Journal*, 2016, *44* (3), 309–331.

**Palmer, Glenn and Archana Bhandari**, "The Investigation of Substitutability in Foreign Policy," *Journal of Conflict Resolution*, February 2000, *44* (1), 3–10.

**Pavlick, Ellie, Travis Wolfe, Pushpendre Rastogi, Chris Callison-Burch, Mark Dredze, and Benjamin Van Durme**, "FrameNet+: Fast Paraphrastic Tripling of FrameNet," *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Short Papers)*, 2015, pp. 408–413. Beijing, China, July 26-31, 2015.

**Sack, Graham**, "Character Networks for Narrative Generation," in "INT@AIIDE" 2012.

**Sajid, Anamta, Sadaqat Jan, and Ibrar A. Shah**, "Automatic Topic Modeling for Single Document Short Texts," in "2017 International Conference on Frontiers of Information Technology (FIT)" 2017, pp. 70–75.

**Semetko, Holli A. and Patti M. Valkenburg**, "Framing European Politics: A Content Analysis of Press and Television News," *Journal of Communication*, June 2000, *50* (2), 93–109.

**Spirling, Arthur**, "U.S. Treaty Making with American Indians: Institutional Change and Relative Power, 17841911," *American Journal of Political Science*, January 2012, *56* (1), 84–97.

**Taber, Charles S.**, "POLI: An Expert System Model of U.S. Foreign Policy Belief Systems," *American Political Science Review*, December 1992, *86* (4), 888–904.

**Vayansky, Ike and Sathish A. P. Kumar**, "A Review of Topic Modeling Methods," *Information Systems*, 2020, *94*, 1–15.

**Woloch, Alex**, *The One vs. the Many: Minor Characters and the Space of the Protagonist in the Novel*, Princeton: Princeton University Press, 2003.

| | Min | Max | Median | Q1 | Q3 | Median absolute deviation | Mean | Standard deviation | 95% confidence interval | Skewness |
|---|---|---|---|---|---|---|---|---|---|---|
| **FOMC-2015: $N = 49$** | | | | | | | | | | |
| Number of reference documents in the article (Count) | 1 | 7 | 2 | 1 | 2 | 1.48 | 1.92 | 1.26 | 0.36 | 1.88 |
| Distance between earliest reference document and article (Number of days) | 0 | 365 | 2 | 1 | 15 | 1.48 | 17.78 | 53.62 | 15.40 | 5.79 |
| **US-RUSSIA (2015): $N = 126$** | | | | | | | | | | |
| Number of reference documents in the article (Count) | 1 | 7 | 2 | 1 | 2.75 | 1.48 | 2.03 | 1.30 | 0.23 | 1.62 |
| Distance between earliest reference document and article (Number of days) | 0 | 2287 | 2 | 1 | 13 | 1.48 | 101.79 | 310.92 | 54.82 | 4.44 |