Talkin' bout Digitalization

A Comparative Analysis of National Discourses on the Digital Future of Work

Matteo Marenco¹

Timo Seidl²

20 April, 2021

Abstract

New forms of work intermediation - the gig economy - and the growing use of advanced digital technologies - the new knowledge economy - are changing the nature of work. The digitalization of work, however, is shaped by how countries respond to it. But how countries respond to digitalization, we argue, depends on how digitalization is perceived in the first place. Using text-as-data methods on a novel corpus of translated newspaper and policy documents from eight European countries as well as qualitative evidence from interviews and secondary sources, we show that there are clear country effects in how digitalization is framed and fought over. Drawing on discursive-institutionalist and coalitional approaches, we argue that institutional differences explain these discursive differences by structuring interpretative struggles in favor of the social coalitions that support them. Actors, however, can also challenge these institutions by using the discursive agency to change these underlying support coalitions.

¹ Scuola Normale Superiore

² Centre for European Integration Research, University of Vienna

We would like to thank Dorothee Bohle, Guglielmo Meardi, and the participants of the ECPRpanel on the Politics, Governance and Regulation of Big Transformations for helpful comments on earlier versions of this paper.

Talkin' about Digitalization

Introduction

From manufacturing goods to building relationships, from daily business to everyday life, from credit scores to social scores: digital technologies are rapidly transforming the way we live our lives and run our economies. But digitalization is no force of nature. It does not just sweep over the world, leveling everything in its path. Rather, it is molded and channeled by the uneven landscape of history and politics. Take the ride-hailing platform Uber as an example.

Uber has aggressively rolled out its app in dozens of countries, probing the bounds of existing labor and taxi regulations. Different countries, however, have “responded in wildly different ways to [Uber] – from welcome embrace and accommodating regulatory adjustments to outright rejection and legal bans” (Thelen 2018, 938). And not only that. The advent of Uber also provoked very different discursive “flashpoints” (Thelen 2018, 939) around which political struggles later revolved (e.g., taxation, competition, employment rights). These flashpoints, while resulting from institutional differences (e.g., different welfare systems), shaped the terms of public discourse on Uber. They led to the mobilization of different actors and the formation of different coalitions (e.g., taxpayers and regulators vs. Uber, Uber and consumers vs. drivers), which help explain the different regulatory responses.

This teaches us two lessons: first, platform companies like Uber are not only regulated differently; they are also talked and thought about differently (e.g., as a taxation vs. a labor law problem). Second, to understand political responses to digitalization, we need to understand how institutions influence the terms on which political struggles are fought. For country-specific characteristics “do not simply channel the same conflict in distinctive ways” but often “translate seemingly common trends into wholly different problems in divergent national contexts” (Thelen 2018, 939).

An institutional theory of digitalization, therefore, not only needs to account for “diverse responses to common challenges” (Steinmo and Thelen 1992, 6) but also for how seemingly common challenges are “refracted into divergent struggles over particular national practices” (Locke and Thelen 1995, 338). This refraction, we contend, happens in discourse, that is the process through which agents articulate

and communicate the substantive content of ideas (Schmidt 2008). It is in discourse where actors voice their concerns about digitalization and express what they think are its defining features and central challenges (Kjær and Pedersen 2001; Schmidt 2008; Wueest 2013). And it is in discourse that global trends are translated into local problems – problems which subsequently influence the scope and nature of political solutions (Hay and Rosamond 2002).

To be clear, our argument is not that discourse is all that matters. To the contrary, discourse itself is shaped by the politico-economic status quo and the social coalitions that underpin it. But digitalization also provides actors with an opportunity to build new coalitions and challenge this status quo (Emmenegger 2021). In this article, we put forward a discursive-institutionalist analysis of the varying ways in which digitalization is framed and fought about in eight European countries. We use various text-as-data approaches (Grimmer and Stewart 2013) on a novel dataset of text corpora to show how actors talk about digitalization (sentiment analysis) and what they talk about (keyword extraction, topic modeling). In addition, we use more qualitative evidence from interviews and fieldwork as well as secondary sources to contextualize these findings and show how discursive agency manifests itself. Our datasets consists of two types of corpora: articles in quality newspapers to analyze ‘communicative discourse’ between political actors and the public; and policy documents such as government white papers or manifestos by trade unions and business groups to analyze ‘coordinative discourses’ among policy actors (Schmidt 2008).

While digitalization is a multi-faceted process, we limit our analysis to one of its most important dimensions: the digital transformation of work [Neufeind.2018]. This refers to both the emergence of new forms of both online and offline work intermediation through digital platforms (‘gig economy’) (de Stefano 2016; Prassl 2018); and the growing use of advanced digital technologies like artificial intelligence to further rationalize the production process (‘new knowledge economy’) (Brynjolfsson and McAfee 2014; Schwab 2016). The former poses the problem of how to adapt labor regulations and social security systems to novel – and often highly precarious – forms of work; the later raises the question of how to prepare workers for – and compensate them in – a world in which machines can perform more and more tasks.

Our results suggest that digitalization is indeed talked and thought about in very different ways in different countries, and that these differences reflect underlying politico-economic differences (Wueest 2013). In particular we show that while socioeconomic factors matter, institutions are the primary source of national differences, or “country effects” (Lloyd and Payne 2019) in how digitalization is discussed. Based on our findings, we develop a typology of digitalization discourses that captures how collaborative (vs conflictual) as well as how opportunity-focused (vs threat-focused) national discourses are. In addition, we show that while institutions shape how actors perceive novel phenomena, actors can also use their discursive agency to challenge or reinvent these institutions.

Our paper makes three contributions. First, we provide a comprehensive mapping of digitalization discourses in a diverse set of eight European countries, showing that digitalization is far from being the same problem everywhere. Second, we theorize and empirically demonstrate how institutions – and their underlying support coalitions – mediate the digital transformation of societies and economies by shaping how digitalization is framed and fought over. Third, we show how novel text-as-data methods can be used to answer discursive-institutionalist questions, even and perhaps especially when they are comparative in nature.

We proceed as follows. We will first present our research questions and elaborate on our argument. We then outline our theoretical framework. Next, we explain our dataset and methodological approach, and discuss our empirical results. We conclude with a brief summary of our results and a discussion of the contributions and shortcomings of our paper.

Research Questions & Argument

Technological change is often portrayed as an unstoppable, tsunami-like force that simply rushes over societies. Economic history, however, is full of examples of ‘successful’ resistance to or differential adaption of new technologies (Frey 2019). In fact, whether new technologies are outrightly banned or eagerly introduced – or anything in between – “depends on who stands to gain from them and the societal distribution of political power”. Technological change, in other words, is a profoundly political

and politically contested process whose outcome depends not just on the technologies themselves but on how countries react them . And these reactions can differ quite dramatically.

But how to explain this variation? While much work on digitalization is rather devoid of a comparative perspective, an emerging ‘varieties-of-digitalization’ literature has started to pay attention to ‘country effects’ in the way that digitalization is viewed and responded to (Lloyd and Payne 2019; Thelen 2018). In this paper, we build on this literature – as well as existing country case studies – to draw attention to three things: first, to the systematic differences in the way that digitalization is thought and talked about in different countries; second, to the importance of institutions in structuring the way digitalization is understood; and third, the limited although important role of discursive agency in challenging and changing existing institutions. We address three research questions. First, are there relevant and theoretically meaningful differences in the way countries talk about digitalization? Finding descriptive differences in discursive tone or content is the minimal condition for a discursive-institutionalist account of country effects. Put differently, mapping discursive diversity is a precondition for explaining it.

Second, to the extent that there are such ‘country effects,’ how do we explain them? Why do countries perceive digitalization differently, and what is the relative importance of socioeconomic versus institutional factors?

Third, to what extent do different discourse merely reflect politico-economic differences – economic or institutional – and to what extent do they leave room for discursive agency? In other words, is discourse mainly a locus of continuity where structurally or institutionally powerful actors express their interests? Or is it a locus of change in which actors can challenge existing ways of thinking about and doing things? The example of Uber has shown that different discursive flashpoints organize debates around platform work and thus shape the coalitions in favor of specific regulatory responses (Thelen 2018). But do these discursive flashpoints merely reflect underlying institutional differences, or is there room for genuine discursive agency?

Our discursive-institutionalist conjecture is that discourses on digitalization i) differ systematically across countries (the descriptive part) because ii) institutional differences give different meaning and

relevance to it (the institutional part); and iii) because actors can use discourse to build new coalitions or drive a wedge between existing ones, thereby challenging and changing existing institutions (the agency part).

This addresses the two sides of the structure-agency debate in institutional theory. On the one hand, institutions are said to structure the ways in which novel problems are perceived, conceived, and ultimately assimilated. On the other hand, such accounts are criticized for being overly deterministic and for downplaying the opportunities that novel problems present to actors that are interested in changing (or shoring up) institutional arrangements. These two sides, however, are not incompatible. As the next section fleshes out, they can be reconciled by understanding (discursive) agency as the “coalitional work” that goes into creating, maintaining and changing institutions, while acknowledging the difficulties involved in challenging an institution’s underlying support coalition (Capoccia 2016; Emmenegger 2021; Hall 2016).

Discourse matters because it offers a window into the operation of both institutional structure and agency: it is in discourse that cultural categories are institutionalized and novel problems assimilated to existing ways of thinking (Capoccia 2016, 1104). But it is also in discourse that actors contest the meaning and legitimacy of an institution (Schmidt 2010, 15). A discursive-institutionalist analysis thus allows for more “rounded accounts of agency within institutional settings” (Bell 2012, 718) – accounts that take seriously both the ‘partial bite’ of institutions (Capoccia 2016) and the possibility of discretionary spaces.

Theoretical Framework

Institutionalism evolved as a research program focusing on how institutions structure actors’ behavior, thus explaining “policy continuities over time within countries and policy variation across countries” (Steinmo and Thelen 1992, 10). More recently, institutionalists have turned their attention to the role of actors in changing institutions by exploiting the inevitable incompleteness and ambiguity of institutions (Streeck and Thelen 2005, 14). What an institution “really means” is never quite clear but is negotiated in a continuous “interpretative struggle” over its meaning, scope, and legitimacy (Streeck

and Thelen 2005, 14–15). Actors can thus, in principle, challenge and change the institutions in and under which they operate.

This growing focus on agency, however, raises the “paradox of plasticity” (Hall 2016, 39): if agency and institutional change are ubiquitous, what remains of the very stickiness and ‘patterning’ effects that inspired institutionalist analysis in the first place? The “coalitional turn” (Emmenegger 2021, 2) in institutional analysis is meant to address this problem (Hall and Thelen 2009). It conceptualizes institutions “as the product of social coalitions” (Hall 2016, 39) and agency in terms of “the creation and maintenance of social coalitions that stabilize or challenge institutions” (Emmenegger 2021, 2). The advantage of such a coalitional understanding of institutions is that it allows to “strike a middle ground between characterizing institutions as absolutely constraining or as malleable tools in the hands of powerful actors” (Capoccia 2016, 1100). While institutions are “creatures of coalitions” (Hall 2016, 42) and therefore rely on the continuing support of actors, they also sustain the coalitions that support them (Hall 2016) – for example, by providing incumbent coalitions with material benefits (Pierson 2000) or control over the timing of institutional reform (Capoccia 2016).¹

Another way in which institutions ‘bite’ – and the one on which we focus here – is through the “institutionalization of cultural categories” (Capoccia 2016, 1102–1107). Institutions not only play a regulative role. They also play a constitutive role (Emmenegger 2021, 7) by shaping how “different groups come to have particular beliefs and conceptions of the world, their roles in it and their interests” (Whitley 2007, 549). “Interpretative feedback effects” (Capoccia 2016, 1102) tilt the interpretative struggles over institutions in favor of their support coalitions. Reformers, meanwhile, need to do a lot of “‘definitional work’ to change established categories and ways of thinking (Capoccia 2016, 1106). Novel problems are therefore often viewed in light of existing national models and, as a result, problem definitions and policy solutions follow ideational path dependencies (Cox 2004; Kjær and Pedersen 2001).

Discursive institutionalists have theorized this constitutive effect of institutions through the concept of background ideational abilities, which allow actors to create and maintain institutions

¹At the same time, collective action problems make it difficult for challenger coalitions to unseat incumbent coalitions (Emmenegger 2021).

(Schmidt 2008, 314). But they have also argued that institutions are not as ‘sticky’ and actors not as ‘unthinking’ as they are often made out to be. This is because actors also have foreground ideational abilities which they can use “to think, speak, and act outside their institutions even as they are inside them, to deliberate about institutional rules even as they use them, and to persuade one another to change those institutions or to maintain them” (Schmidt 2008, 314).

Conceptualizing actors as endowed with both background and foreground ideational abilities addresses another paradox of institutionalist theory: the “paradox of embedded agency”; it explains how actors, “whose belief and actions are strongly shaped by existing institutions [can] brake with these very same institutions and innovate” (Emmenegger 2021, 6). Together, coalitional and discursive institutionalist perspectives allow us to understand two things at the same time: i) how the common trend of workplace digitalization is not in fact “translated into common pressures in all national economies” but – through the constitutive power of institutions – “refracted into divergent struggles over particular national practices” (Locke and Thelen 1995, 338); and ii) how actors can use their discursive agency over digitalization to challenge coalitions, build new ones, and thereby change these national practices.

On the one hand, actors frame novel problems in ways that make defection from existing institutions less likely. For example, they stress certain aspects (e.g., taxation) of digitalization more than others (e.g., employment status), thereby rallying the supporters of a tax-based welfare system behind certain types of regulations (Thelen 2018).² By asserting the continuing relevance and legitimacy of established ways of doing and thinking about things, these frames shore up the coalitions that support the institutional status quo (Capoccia 2016, 1105). On the other hand, actors can frame a novel problem as an opportunity for change – or as necessitating it. By pointing to the inadequacy of the institutional status quo, these frames can be used to build new coalitions or drive a wedge between existing ones.

In sum, by taking seriously the coalitional underpinnings of institutions as well as the effect of (institutional) context on actors’ background ideational abilities and the possibility of agency that exists

²Following Entman (1993), p. 53, we understand framing as “selecting and highlighting some features of reality while omitting others” and discourse, following Schmidt (2010), as the interactive exchange of such frames.

because of actors' foreground ideational abilities, this approach strikes "a middle ground between characterizing institutions as absolutely constraining or as malleable tools in the hands of powerful actors" (Capoccia 2016, 1100). Institutions are created, maintained, and challenged by social coalitions, and a key part of their coalitional work is the definitional work of framing novel problems in ways that either reassert existing practices – or break with them.

Empirical Strategy & Data

Testing the twin arguments of institutional structuring and discursive agency requires combining two empirical strategies. First, we need ways to systematically assess whether and how discourses on digitalization differ across countries. Second, we need ways to connect these findings to the actual political strategies of social actors – their efforts in defending and defecting from existing institutions, and in keeping together or breaking apart the coalitions that sustain these institutions. The first task is accomplished using methods of quantitative text analysis; the second one through triangulating these findings with case studies based on fieldwork and secondary sources.

For the quantitative text analysis, we compiled two novel text corpora for eight European countries: France, Germany, Ireland, Italy, Poland, Spain, Sweden, and the United Kingdom. The rationale for our case selection was to investigate discursive diversity among countries that share some basic commonalities. Each of our eight countries, with the partial exception of Poland, is an advanced capitalist democracy, and each was, for the period under study, a medium-sized to large member of the European Union. Despite these similarities, the countries under study also differ in important respects, both with regard to the structure of their economy and with regard to their institutional makeup. This constrained diversity provides us with the appropriate setting to study how structural and institutional diversity is reflected in discourse.

The first corpus we collected consists of 6235 newspaper articles that directly relate to the digitalization of work and were published between 2013 and 2019. Articles were collected from major national newspapers, either directly or via Factiva. We prioritized quality over quantity (cf. Nicholls and Culpepper 2020), selecting only articles that contained an explicit reference to the gig economy,

workplace automation and robotization (including artificial intelligence), or the fourth industrial revolution or Industry 4.0 in their headline or lead paragraph (for details, see online appendix). This corpus represents what Schmidt calls the communicative discourse between policymakers and the public and is broadly reflective of the overall public discourse on digitalization in a country.

The second corpus consists of 2337 (sometimes split) documents published between 2012 and 2019 by major social partners, namely unions, governments or ministries, and employer organizations. Following this actor-based logic, documents were collected from the web, based on relevant search strings and available lists of the main stakeholders in each country (e.g., the national member organizations of the European Trade Union Confederation). Longer documents (over 40.000 characters) were split chapter-wise or similarly to have more and more meaningful units of analysis (for details, see online appendix). This second corpus represents what Schmidt calls the coordinative discourse among policymakers and is broadly reflective of the elite discourse on digitalization.

Substantively, both corpora cover the digitalization of work, which we understand as a two-pronged process. The first prong refers to new platform business models that enable and control novel interactions between buyers and sellers of labor; the second prong refers to the ways in which digital general-purpose technologies increasingly affect the ways in which business organize production in both the service and the industrial sector ('Industry 4.0'). We thus distinguish between new forms of online and offline work intermediation through digital platforms – what is often referred to as the 'gig economy' (de Stefano 2016; Prassl 2018); and the growing use of advanced digital technologies that promise to transform the production process more generally – what is variously referred to as the second machine age, the fourth industrial revolution, or the (new) knowledge economy (Brynjolfsson and McAfee 2014; Schwab 2016).

These two dimensions of digitalization come with different challenges. The former intensifies existing trends of workplace casualization (de Stefano 2016). It poses the problem of how to adapt labor regulations and social security systems to minimize the costs while maximizing the benefits of new, digitally intermediated forms of work. The latter intensifies existing trends of industrial and service-sector automation and routine-biased technological change (Frey 2019). This raises the

question of how to make sure that workers have the necessary skills to remain in the labor market or are provided with alternative sources of income.

Both corpora were automatically translated into English, using Google’s Translate API. Other studies have shown that this works as well as gold-standard translations for both dictionary-based and unsupervised methods of text analysis (Proksch et al. 2019; Vries, Schoonvelde, and Schumacher 2018). Building on these studies, we use the translated corpora as input for various methods of automated text analysis.³ While these methods inflict a certain degree of ‘semantic violence’ on texts, they can also be extremely useful when carefully validated and interpreted, which we discuss below (Benoit 2020). In addition, it is crucial to align theory and method. In this paper, we assume that the frequency of certain words is substantively meaningful (e.g., because it tells us something about dominant framings), so we can use bag-of-words approaches like topic models (cf. Nicholls and Culpepper 2020). If, however, we assumed words derive their meaning from ‘the company they keep’ and are interested in how their meaning changes over time or across contexts, word embeddings would be a more fitting approach.

In this paper, we use three text-as-data methods that match our theoretical interest in the tone and content of discourse. First, we are interested in how digitalization is talked about in a given country. For that purpose, we use sentiment analysis to identify the relative number of positive versus negative terms in different documents. We use the Lexicoder Sentiment Dictionary, which was specifically designed for political text (Young and Soroka 2012), taking into account negators (positive negatives and negative positives) as well. We cross-validate this approach by using four additional dictionaries (for details, see online appendix).

Second, we are interested in what is said in a given country. To investigate this, we use various keyword extraction techniques as well as structural topic modeling..⁴ Keyword extraction techniques make it possible to identify words or n-grams that capture essential content. We use four of such

³For translation, preprocessing, NLP annotation, and the analyses, we mainly relied on the R packages *googleLanguageR* (Edmondson 2020), *quanteda* (Benoit et al. 2018), *spacyr* (Benoit and Matsuo 2020), *stm* (Roberts, Stewart, and Tingley 2018) and *udpipe* (Wijffels 2019) as well as the tidyverse packages (Wickham et al. 2019).

⁴We also use a version of word vectors. For details and results, see online appendix

techniques: the RAKE algorithm, the textrank algorithm and sentence-based as well as skipgram-based co-occurrence metrics (for details, see online appendix).

Topic modeling allows us to identify common themes in our corpus. Topics are modeled as probability distributions over words and documents as probability distributions over topics (Blei 2012). Put differently, topics are understood as combinations of words (or n-grams) and documents as mixtures of topics. Structural topic modeling extends this basic topic model by allowing observed covariates to influence the proportion of a document devoted to a topic (the topic prevalence), which allows us to estimate the conditional expectation of a topic occurring given that a document is from a certain country or part of the communicative discourse. We chose a topic model with $k = 60$ topics, some of which we subsequently combined into “frame packages” (Nicholls and Culpepper 2020, 11). Online appendix provides a detailed discussion of our choice of k , preprocessing steps, interpretation, and validation, as well as a complete list of topics with terms, labels, and assigned frame package.

We rely on qualitative evidence to complement our quantitative findings. In particular, we relied on in-depth analysis of relevant governments’ and social partners’ policy documents, academic as well as journalistic sources. For four countries in our sample – France, Spain, Italy, and Germany – we complemented this document analysis with semi-structured interviews with members of employment Ministries, social partners concerned with the digitalization of work, as well as platform activists and experts (see appendix). Where we did not collect interview data, we rely on the aforementioned secondary literature only. The purpose of such qualitative data is two-fold: first, to let the quantitative discourse analysis ‘come to life,’ i.e., to show how discursive differences relate to what actually happens ‘on the ground’; and second, to zero in on the relative importance of institutional structure and discursive agency which we can only superficially analyze quantitatively.

Results & Discussion

Do discourses on the digital future of work differ in tone and content across countries, and if so, how do we explain these differences? Here, we first look at the tone of discourse and tentatively derive a typology of digitalization discourses. We then flesh out this typology by looking at the content of

discourse. In both cases, as highlighted above, we combine quantitative evidence – from the sentiment analysis in the case of discursive tone and from keyword extraction methods and topic models in the case of content – with qualitative data from secondary sources and interviews.

The Tone of Discourse

The relative number of positive versus negative words is a good proxy for a country's view of digitalization: is it perceived as something bad, threatening, destructive? Or as something good, promising, useful? Figure 1 plots discursive sentiment across different countries, for different dictionaries as well as separately for coordinative and communicative discourses (with values averaged across dictionaries).⁵ A first finding is that Sweden has by far the most positive discourse. Importantly, this cannot be explained by the country's economic position alone. The UK or Ireland also have service-dominated economies with many companies close to the digital frontier. But they perceive digitalization more negatively. We argue that what explains these differences is not the structure of Sweden's economy but the nature of its institutions.

Specifically, Sweden has an institutionalized tradition of proactively and inclusively adapting to technological change. This makes digitalization a lot less scary (Katzenstein 1985). Sweden's embracing view of digitalization is epitomized by the opening quote of a recent report by its digitalization commission: *Tempora mutantur nos et mutamur in illis* – the times change, and we change with them. "In Sweden," as the Swedish minister for employment and integration, Ylva Johansson, put it, "if you ask a union leader, 'Are you afraid of new technology?' they will answer, 'No, I'm afraid of old technology. (...) The jobs disappear, and then we train people for new jobs. We won't protect jobs. But we will protect workers'" (Goodman 2017). Institutionally, this is reflected, among other things, in the country's job security councils, which are jointly run by employers and trade unions and which give intensive support and retraining to people when they are laid off. As we will further see below, Sweden's ability

⁵In online appendix , we provide wordclouds with the most commonly matched positive of negative words to illustrate what a positive or negative tone means and to further validate this approach. We thank an anonymous reviewer for this suggestion.

to collaboratively negotiate technological change fundamentally changes the country view of and reaction to digitalization. By contrast, Poland stands out by having a particularly negative coordinative discourse, precisely because it lacks an established tradition of social and policy dialogue. As a result of this difficulty of building cross-class coalitions, there is a conspicuous lack of forward-looking plans and institutionalized cooperation in shaping digital change.

The UK and Ireland also lack such a system of institutionalized cooperation between winners and losers of technological change. This results in a particularly negative communicative discourse and a government-dominated coordinative discourse, which pushes Union into voicing their concerns directly to a receptive public. The Irish government, for example, is very sanguine about digitalization, having long espoused an active policy of catering to tech companies and casting Ireland as a digital frontrunner (Brazys and Regan 2017). By contrast, the Irish public appears more worried about digitalization's potential downsides and inequities, as evidenced by the much more negative discourse on the digital future of work. Our findings dovetail with a recent study, which finds that at least 25 percent of the Irish population finds itself excluded from an increasingly digital society because of socio-economic reasons, and almost half describe themselves as lacking digital skills (Accenture 2020). Similarly, the British government frames digitalization very positively, but its unilateral celebration of digitalization as a vindication of and boost to the British way does not chime well with a more skeptical public.

Figure 1: Discursive sentiment across countries and discourse types with 95 % bootstrapped confidence intervals (R = 10000)

There are large class and regional disparities in Britain when it comes to digitalization, and less trust in the country's ability to find collaborative solutions. Unions, for example, are irritated by the government unilateral approach, which did not involve them in drafting the UK's digital industrial strategy, and which one-sidedly stressed the benefits of the flexible 'British way' in its response to the Taylor Review of Modern Working Practices (UK Government 2018a). This response struck many as

rather tone-deaf at a time when digitalization threatened to exacerbate the downsides of the flexible British labor market. Again, it is institutions – in this case a lightly regulated labor market and a unilateral (as opposed to corporatist) style of policymaking – result in large discrepancies between a relatively cheerful discourse by the government and a much more negative discourse among the public.

Importantly, without having a seat on the table and without being listened to, Unions resort to more dramatizing discursive strategies. They also shift their focus from the coordinative to the communicative discourse and address the public more directly, which explains the quite positive coordinative and the quite negative communicative discourse (cf. Schmidt 2008, 312–313). . Using an aggressive style of campaigning that highlighted the conflicts and inequities in the British labor market, the Independent Workers Union of Great Britain (IWGB), for example, successfully raised awareness for the working conditions not just of gig workers but of precarious workers in general. As the IWGB's vice-president Mags Dewhurst puts it, “this isn't just about the gig economy. All business is going more and more digital, leaner and leaner. (...) Next, it will be banking and retail. These bad practices have to be stopped now”(Roberts 2018). Thus, the IWGB has used the spotlight that digital platforms have put on the inequities of the British labor market to challenge the support coalition behind the flexible ‘British way,’ for example by pointing to the wide-spread non-enforcement of labor rights, which is not limited to platform companies (Roberts 2018).

In sum, the tone of public discourse clearly differs across countries. As the above discussion has hinted at and as the analysis of discursive content will flesh out, discourses differ along two dimensions. The first dimension relates to how collaborative policymaking is. Specifically, corporatist countries seem to be better at mitigating fears of technological disruption. This is because the losers of technological change can plausibly assume that they will either be compensated in the digital future or prepared for what it brings, allowing for broader coalitions in favor of technology-embracing policies (Katzenstein 1985; Ornstorn 2012). The stakes are higher in countries with a more conflictual style of policymaking, where each side has to fear that the other gets its way and one ends up getting the short end of the stick, which inclines groups to fight for their short-term interest. The second dimension has

to do with whether previous experiences with technological change and a country's politico-economic makeup incline countries to perceive digitalization as an opportunity or as a threat. Is my country likely to benefit from digitalization, can it be an opportunity to overcome existing problems or will it make them worse? While the first dimension reflects purely institutional differences, the second dimension has both economic and institutional elements (e.g., the distance to the digital frontier, the nature of the labor market).

Table 1 brings these two dimensions together, distinguishing between four ideal typical discourses. If the policymaking style is collaborative and digitalization is viewed as an opportunity, a proactive discourse dominates in which digitalization is embraced and the emphasis is on investments that help workers and companies survive and thrive in a digital economy. This makes it much easier to build broad coalitions that support future-oriented institutions such as (re-)training programs. If the policymaking style is at least somewhat collaborative but digitalization is viewed as a threat, the focus is on how to passively compensate individuals and business that stand to lose from digitalization. If the policymaking style is conflictual and digitalization is viewed as an opportunity, the discourse will be unilateral in the sense of different actors emphasizing their own and often conflicting interpretations of how digitalization can benefit society. If the policymaking style is conflictual and digitalization is viewed as a threat, a Luddite discourse will emerge in which actors try to stop or reverse digitalization to the extent that it hurts them. This will lead to narrow, conflicting coalitions, which makes it hard to build future-oriented institutions. The following paragraphs will add empirical flesh to these conceptual bones by zeroing in on the content of discourse, and by showing how actors can use their discursive agency to either confirm the discursive status quo or shift discourses from one quadrant to another.

Table 1: Typology of Digitalization Discourses

		<i>Policymaking Style</i>	
		Collaborative	Conflictual
<i>Digitalization Perception</i>	Opportunity	Proactive Discourse	Unilateral Discourse
	Threat	Compensatory Discourse	Luddite Discourse

The Content of Discourse

Not only do digitalization discourses differ in tone, they also importantly diverge in content. Figures 2 depicts the most important keywords identified by three keyword extraction algorithms.⁶ It depicts – with size indicating importance – the unigrams and bigrams that we can think of as most ‘essential’ to the communicative (black) and coordinative (grey) discourses in the respective countries. They capture the issues around which national discourses revolve. Figure 3 depicts the estimated effect of the country covariate on the prevalence of different topics (or aggregation of topics). It depicts the conditional expectation of a bundled set of topics or ‘frame package’ to occur in a document given that this document is from a given country. It thus helps us answer the question which topics are important in some countries relative to others.

So far, we have encountered Sweden as a country with a proactive discourse on digitalization. But while some aspects of digitalization play into Sweden’s hands, others challenge core pillars of its social model. On the one hand, the country’s corporatist institutions and “ideology of social partnership” give it an ability “to live with change” instead of resisting it (Katzenstein 1985, 32, 211; Ornston 2012). This “asymmetric embrace of markets” (Pontusson 2011, 107) – a perfect example of ideational path-dependency (Cox 2004) – credibly ensures flexible-yet-inclusive adjustments. This stabilizes coalitions for forward-looking adjustments, which, as we have seen, manifests itself in a remarkably positive discourse on digitalization. On the other hand, platform companies like Uber challenge the country’s tax-based welfare state and collective bargaining tradition by systematically avoiding taxes and labor market regulations.

This brings actors to the scene that want to defend the Swedish social model and therefore put the issue of taxation on top of the agenda. In fact, tax evasion in the gig economy is a far more important topic in Sweden than in any other country, and mainly so but not exclusively among social partners. Qualitatively, we see this mirrored in the fact that the “prime concern” of Unionen, Sweden’s biggest trade union, is not the employment status of gig workers but that labor platforms follow Swedish

⁶For details on these methods, and for similar results using the textrank algorithm, see online appendix .

taxation rules and collective agreements (Söderqvist 2017; Thelen 2018). In what is strong evidence for ideational path-dependence, Unionen has developed a “Nordic approach to regulating labour platforms” that wants to make “collective agreements easier to integrate and more compatible with the platform firms’ software, in essence developing digitalized versions of existing collective agreements” (Söderqvist 2018, 301). In short, rather than banning digital business models, Unionen wants to make it easier for digital platforms “to be good platforms” (Söderqvist 2017). Thus, despite the challenges it poses, even Swedish unions see digitalization more as an opportunity – as something that requires a particular response but can ultimately strengthen the Swedish model as long as one is “proactive in the implementation of the new technology” (Wallin 2017).⁷

⁷This positive, proactive attitude is echoed in Sweden’s corporatist Scandinavian neighbors. In Norway, both unions and employers view automation as fundamental to maintaining the country’s labor market and welfare model (Lloyd and Payne 2019, p. 217) while the Danish union 3F was one of the first to sign a collective agreement with the cleaning service digital platform Hilfr.

In Germany, the newest wave of digitalization has brought “deep and transformational changes” (Diessner, Durazzi, and Hope 2020, 10) to the country’s dominant manufacturing sector, which has sparked an intense debate about its future. This debate, however, was and is not about abandoning manufacturing but about defending it through focusing on the “digital transformation of products and production within the traditional industrial core” (Thelen 2019, 295). We see this discursive shift reflected in the centrality of digital manufacturing in the German discourse, which comprises topics such as industry 4.0, smart factories, or cloud computing. What explains this “ongoing active adaption” (Thelen 2019, 302) to more knowledge-intensive types of manufacturing is the sectorial organization of German producer groups and the dominance of its manufacturing sector therein, which put digital manufacturing high on the agenda pushed for higher investments in education and R&D to support this adaption (Thelen 2019, 300–304).

It is striking to see that across social groups, the challenges posed by digitalization have not been met with protective defensiveness but with a rather proactive willingness to meet them head-on. Central political actors have used their discursive agency to frame digitalization not as something that should be cushioned by compensatory policies but as something that can and should be actively shaped. The SPD-led labor ministry, for example, started a big debate on the future of ‘good work’ with its influential white paper ‘Work 4.0,’ which served as a discursive focal point. Rather than being overtaken by events, the ministry actively initiated a debate on what ‘good work’ could look like in the digital age. Moreover, it systematically involved social partners in coming up with answers. This helped shift the debate into issues that social partners could live with or find common ground on, such as flexible working time, training, or data protection. We see this reflected in the keywords on data protection, work time, and further training as well as in the importance of the skills and education frame package.⁸

Contrary to their conservative reputation, German unions, too, have “opted to go on the offensive and adopt a strategy aimed at securing active participation in shaping change, as opposed to rejecting

⁸We also see this reflected in a topic that is specifically on work councils, collective agreements, data protection, and flexible working hours.

it and then fighting over the consequences” (Haipeter 2020, 242). They have also concentrated on the above-mentioned issues. This allowed them to have a say in shaping digitalization rather than just slowing it down or compensating its losers. They have chiefly done so through work councils and workplace-level agreements, which is quite remarkable given that work councils were not always so proactive about technological change (Haipeter 2020, 247). Germany’s corporatist institutions thus channel the German debate in ways that are related to but cannot be derived from the particular challenges the country’s manufacturing-heavy and export-oriented growth model faces. At the same time, political actors use the corporatist institutional environment to push issue onto the agenda that allow them to have a say in how digital technologies are used, while often doing so in ways that break with institutional path-dependencies.

Institutions also shape the terms on which digitalization is discussed in the UK, where there is a wide-spread perception – further bolstered by combative unions – that digital platforms threaten to put an additional strain on already highly flexible labor markets. We can see this in the many references to gig work, self-employment, zero-hour contracts, or minimum wages. We also see it in the importance of topics about the legal and economic status of platform workers, which, *nota bene*, is particularly prevalent in communicative discourses. In a climate dominated by fears of a further erosion of worker’s rights and the unequal effects of automation on the labor market, even the conservative government had to concede that certain minimum standard needed to be upheld. But it also emphasized the importance of defending “the UK’s position as one of the best places in the world to do business” (UK Government 2018b) by continuing the business-friendly tradition of the ‘British way.’ This has led to a fairly conflictual discourse in which concerns over the regulation of labor platforms (e.g., over the regulation of Uber in London) are layered on top of a wider conflict over the nature and merits of this ‘British way.’

Tellingly, topics that are about the threat of automation – either as destroying jobs or increasing inequality – are by far the most prominent in the UK and Ireland. At the same time, topics that are about investments in knowledge-based capital, mainly skills but also universities, research, and digital infrastructures, are much less important. These topics figure more prominently in corporatist Sweden

and Germany, and they are overwhelmingly an issue in coordinative discourses. In Germany, in addition to the unions' efforts to put further training center stage, education has been a central concern for the ministry for labor and employment. In fact, the ministry has even toyed with rebranding itself as the ministry for labor and qualification and introducing a right to lifelong learning (interviewee DE-GOV₁). In Sweden, Unionen sees digitalization as an important opportunity to mobilize people around the issue of skill formation: "For a trade union movement that has spent decades trying to convince legislators to invest in lifelong learning, the challenge of digitalization has finally pushed the issue higher up the agenda" (Söderqvist 2018, 298). Such supply-side investments also found "ready allies" (Thelen 2019, 305) among Sweden's most influential business groups (cf. Ornston 2012).

This underscores the point that corporatist institutions create more future-oriented discourses by allowing actors to collaboratively focus on positive-sum responses to digitalization while mitigating fears of automation (Hicks and Kenworthy 1998). In the UK and Ireland, such corporatist institutions are much weaker. Especially in Ireland, but also in the UK, there is a debate about digital industrial policy, i.e., government measures to support the competitiveness of domestic companies in the digital age. But this debate is precisely not about collaborative investments in knowledge-based capital but about how the government can support domestic businesses through venture funds, taxation, or other developmental tools (for Ireland, see Brazys and Regan 2017). Overall, we thus find a rather polarized discourse, where governmental actors try to defend the status quo – shoring up the underlying coalition by making smaller concessions – while other actors use digitalization to expose broader problems and push for broader changes, trying to build new coalitions with a potentially responsive public.

Figure 3: Estimated effect of country variable on the prevalence of topic clusters. Color gradients indicate the ratio of the estimated effect of the discourse type variable, with red colors representing a higher conditional expectation to find the respective topic packages in coordinative (as opposed to communicative) discourses, and blue colors representing the opposite. Triangles indicate that the effect of the coordinate discourse was stronger than the effect of the communicative discourse, i.e. that the ratio of the two was greater than 1.

Unsurprisingly, public initiatives are a major topic in France as well, where the state traditionally plays an important role in steering economic transformations (cf. Trumbull 2004). The national plan on Artificial Intelligence by the Hollande administration is one example of this, Macron’s presidency’s

ambition to transform France into a ‘start-up nation’ another. Relatedly, France has an intense debate on gig work. The 2016 Loi Travail introduced specific legislation aimed at protecting platform workers. There is, however, also a lot of conflict over what the gig economy is. The very term ‘collaborative economy,’ which is mainly used by the government, is itself contested as it one-sidedly highlights the positive dimension of gig work. This linguistic difference mirrors broader conflicts over digitalization. Unions, platform activists, and the National Digital Council, for example, have criticized the Loi Travail for relying on the social responsibility of digital platforms thereby ruling out labor law solutions (interviewees FR-EXP2/TU1, FR-GOV1, FR-PA1). We see this conflict reflected in the intense debates about the status of platform workers. Thus, not unlike in the UK, the government’s ambition to make France a model digital nation clashes with broader concerns about the distributional inequities of the digital economy (interviewee FR-EXP1).

The discourse in Poland is even more contentious, due to highly combative struggles between Uber and taxi groups which have come to dominate the country’s debate on gig work. Different actors – even within the government – hold widely different views on Uber, with some viewing it as an innovative company and arguing for liberalization while others depicting it as a law-breaker and questioning not only its legality but also its legitimacy (Serafin 2019). Due to the lack of a cohesive strategy by social partners and unions in particular, these conflicts have often taken a Luddite turn culminating in physical attacks on Uber cars and drivers. The other big issue is the impact of artificial intelligence – the by far most important keyword – on the Polish economy. On the one hand, Poland is a potential digital challenger with a relatively large ICT talent pool, relatively high-quality digital infrastructure, and relatively mild legacy technology lock-ins (McKinsey 2018). On the other hand, Poland strongly relies on sectors at risk of being further automated, from resource extraction to transport to ancillary industries (Skóra 2018). Despite a number of initiatives to speed up the adoption of new technologies, there is no inclusive plan for how to respond to digitalization, with unions focusing on more tangible bread-and-butter issues (Skóra 2018, 465) and many employers not knowing how to incorporate digital technologies into their business model or production processes (Prokop 2018). As we have seen, this

institutional inability to pull together manifests itself in a much more negative coordinative discourse among social partners, and in the prevalence of threatening and contentious keywords and topics.

Similar to Poland, the Spanish discourse is also one of ambivalence about the effects of the digital transformation, which is seen as both a threat to a country long plagued by economic difficulties and inequalities and as an “opportunity for change” (Accenture and Mobile World Capital Barcelona 2017, 7). Importantly, however, both the Spanish government and social partners have taken on a more active role in trying to make sure that the latter view becomes reality. Quantitatively, this is reflected in the important role that topics around government support for digitalization play in the Spanish discourse. Qualitatively, it shows itself in a plethora of initiatives such as the Digital Agenda for Spain (2013), the Agenda for Strengthening the Industrial Sector in Spain (2014), the Connected Industry 4.0 Initiative (2015), and the National AI Strategy (2019). Employers have been supportive of such policy initiatives while stressing the need for enhanced labor market flexibility to harness the potential of digitalization as well as identifying skill shortages as the major obstacle to a successful digital transformation (Confederación Española de Organizaciones Empresariales 2018). Conversely, unions explicitly recognize the potential benefits of digital technologies but warn against their implications for an already strongly segmented labor market (Comisiones Obreras de Industria 2016).

While governmental initiatives point to a generally positive but somewhat unilateral discourse, Spanish unions in particular have also pushed for more social dialogue around digitalization (interviewee ESP_EXP1). For example, the Connected Industry 4.0 resulted from a state-led open consultation process in which unions took part actively while also criticizing it as too ‘soft’ and not going far enough. Moreover, COO and UGT – the industrial union federations – and the Alliance for the Competitiveness of the Spanish Industry signed the Declaration of social partners in 2016, urging the development of a state pact for industry. Likewise, social partners have promoted the ‘Manifesto for the Leadership of Digital Transformation in the Spanish Economy through Talent Development,’ containing proposals on training and education to succeed in the digital transformation. These deliberate attempts to seek social partner collaboration and government support indicate a push to move the discourse from a more unilateral into a proactive discourse.

Similar to the Spanish discourse, the Italian discourse focuses a lot on the various ways in which Italian companies can compete or be more competitive in a global digital economy. The 2016 ‘Calenda Plan’ on Industry 4.0, which foresaw significant financial incentives for firms that decided to invest in digital technologies, is one example. And we also see this reflected in the importance of digital industrial policy topics in Italian discourse. Similar to Germany, there is a large debate on Italian companies can adapt to the digital transformation of manufacturing. Unlike in Germany, however, there is no systematic social partner dialogue about how this transformation can be shaped in a proactive and inclusive manner, initiatives like the CGIL’s Progetto Lavoro 4.0 notwithstanding. Reflecting this lack of a corporatist tradition, the focus is on how the government can help Italian companies survive digitalization, the Calenda Plan being a prominent example.⁹

Conclusion

There has been much debate as to how platform work and the spread of advanced digital technologies will change the nature of work. In this paper, we have taken a step back and looked at these debates themselves. We have argued that shaping these debates is an essential part of the coalitional work (Emmenegger 2021) that goes into maintaining or changing institutions. And have found that discourses on the digital future of work are indeed quite different across countries, both in sentiment and content, as dominant social coalitions attempt to frame digitalization in ways that shore up existing institutional

⁹One thing that our analysis does only partly pick up on, and this is true for both Italy and Spain, is the contested debates on the status of gig workers. We do find that Topic 38, on the collective representation of workers, is most prominent in these countries, but topics on the status of platform workers are not as prominent as in the UK. In both Italy and Spain - especially in the latter case - much regulation came from court rulings. But political actors have actively pushed the issue of platform work on the agenda, which is only partly reflected in our analysis. In Spain, Podemos has presented digital platforms as predatory capitalists that flourish at the expense of ‘left-behind’ workers (interviewee ESP-EXP1). The current government, formed by the Socialist Party and Unidas Podemos alliance, aims at legislating on the matter and Minister of Labor Yolanda Díaz, labor, has started a dialogue with the riders’ collective RidersxDerechos. In Italy, the Five Star Movement has been the most active political formation on gig work regulation (interviewee IT-EXP1). When serving as Minister of Economic Development, Labour and Social Policies, Luigi di Maio called food-delivery riders a ‘symbol of an abandoned generation left with no protection.’ After attempts to regulate platform work via social concertation, in November 2019, the so-called ‘Riders decree’ classified riders as self-employed who nonetheless have access to protections typical of dependent work.

arrangements. After all, how digitalization is construed in discourse shapes which political responses are seen as cognitively plausible and normatively desirable. Different problem definitions create different policy imperatives – things that need to be done in response to digitalization – and therefore narrow or broaden the policy solution space (cf. Hay and Rosamond 2002). For example, if platform work is viewed primarily as an issue of tax avoidance, what is to be done about it will fundamentally be different than if it is viewed as an issue of worker exploitation (cf. Thelen 2018).

But digitalization discourses are more than institutional reflexes. Actors can reframe debates in order to challenge or reform existing institutional arrangements. An example for this is when German unions departed from their institutional path and took on a more active role in shaping technological change. They did so by shifting the discourse to issues in which they had a good chance of having a say in, such as data protection or flexible working times. By taking such a discursive and comparative perspective on digitalization, we have advanced the literature on digitalization in three main ways. First, we have demonstrated empirically that digitalization discourses differ systematically. Mapping such discursive differences helps us better understand how institutions shape the way novel problems are assimilated, but also how actors can use their discursive agency to depart from institutional paths. Second, we have taken up the criticism of paired comparisons (Locke and Thelen 1995) and given it a discursive-institutionalist spin. We have argued that different responses to digitalization cannot be understood as different responses to the same problem. Rather, we stressed the importance of discursive problem construction, and this is discursive work is a key part of the coalitional work that goes into maintaining or changing institutions. It is through such discursive processes that global challenges are translated into a series of local problems, which are what countries subsequently ‘respond to.’ Third, we have demonstrated how novel methods of quantitative text analysis can be used and combined in a comparative framework to capture sentiment and content of national discourses. This opens up new avenues for comparative research, be it with the intention to generalize findings across contexts or, conversely, to use comparative findings as a starting point for more detailed case studies of how discursive features translate into concrete policy decisions.

Despite these contributions, our study also suffers from two main shortcomings, which future research should remedy. First, it remains agnostic when and under what conditions actors are able to escape their institutional constraints and move discourses in a new direction. Understanding these conditions requires both individual case studies that inquire into the motivations of such discursive entrepreneurs, and more explanatory analyses of the institutional or discursive determinants of such agency (on this, see Emmenegger 2021). Second, while our study is a successful proof of concept and therefore broader in its focus, this comes at the price of thinner descriptions and somewhat sweeping explanations of discursive differences and changes. In certain respects, therefore, a more detailed look at, say, the employment status of gig workers or the contestation of delivery platforms would have yielded richer results. These shortcomings notwithstanding, our study has shed new light on old questions about the role of ideas and discourse in comparative political economy and opened up what we see as promising avenues for comparative political research.

Appendix: Interview Table

Table 2: Interview Table

Country	Interview Code	Interview Category	Interviewee Role	Interview Date	Place of Interview
FR	GOV ₁	Policymaker	National Digital Council	28.02.2020	Paris
FR	PA ₁	Platform activist	President and Coordinator Active member, Collective Autonome de Livreurs de Paris (CLAP)	21.02.2020	Paris
FR	EXP ₁	Expert	Labour Sociologist (Université Paris Dauphine-PSL)	11.02.2020	Paris
FR	EXP ₂ /TU ₁	Expert/Trade Unionist	Economist, Insitut de Recherches Economiques et Sociales (IRES)	17.02.2020	Paris
IT	EXP ₁	Expert	Labour lawyer, IE Law School	31.03.2020	Online
ESP	EXP ₁	Expert	Labour sociologist, Malaga University	10.04.2020	Written Interview
DE	GOV ₁	Policymaker	Affiliate of a Federal German Ministry	01.10.2019	Berlin

References

- Accenture. 2020. "Bridging the gap: Ireland's digital divide." https://www.accenture.com/_acnmedia/PDF-128/Accenture-RO-Bridging-The-Gap.pdf#zoom=40.
- Accenture, and Mobile World Capital Barcelona. 2017. "Digital economic opportunity in Spain: September 2017 how digitalization may boost the Spanish economy." https://www.accenture.com/tooo1o1o1tooooooz__w__/es-es/_acnmedia/pdf-59/accenture-strategy-digital-transformation.pdf.
- Bell, Stephen. 2012. "Where are the institutions? The limits of Vivien Schmidt's constructivism." *British Journal of Political Science* 42(03): 714–719.
- Benoit, Kenneth et al. 2018. "Quanteda: An R package for the quantitative analysis of textual data." *Journal of Open Source Software* 3(30): 774.
- Benoit, Kenneth. 2020. "Text as data: An overview." In *The SAGE handbook of research methods in political science and international relations*, eds. Luigi Curini and Robert J. Franzese.
- Benoit, Kenneth, and Akitaka Matsuo. 2020. "Spacyr: Wrapper to the 'spaCy' 'NLP' library." <https://CRAN.R-project.org/package=spacyr>.
- Blei, David M. 2012. "Probabilistic topic models." *Communications of the ACM* 55(4): 77.
- Brazys, Samuel, and Aidan Regan. 2017. "The politics of capitalist diversity in Europe: Explaining Ireland's divergent recovery from the euro crisis." *Perspectives on Politics* 15(02): 411–427.
- Brynjolfsson, Erik, and Andrew McAfee. 2014. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York, NY: Norton.
- Capoccia, Giovanni. 2016. "When do institutions 'bite'? Historical institutionalism and the politics of institutional change." *Comparative Political Studies* 49(8): 1095–1127.
- Comisiones Obreras de Industria. 2016. "La digitalización de la industria." <https://industria.ccoo.es/76ad8479cd94fdf84192c5843d7e8c26000060.pdf>.
- Confederación Española de Organizaciones Empresariales. 2018. "Plan digital 2025. La digitalización de la sociedad española." https://contenidos.ceoe.es/CEOE/var/pool/pdf/publications_docs-file-496-plan-digital-2025-la-digitalizacion-de-la-sociedad-espanolav2.pdf.

- Cox, Robert. 2004. "The path-dependency of an idea: Why scandinavian welfare states remain distinct." *Social Policy & Administration* 38(2): 204–219.
- de Stefano, Valerio. 2016. "Crowdsourcing, the gig-economy, and the law." *Comparative Labor Law & Policy Journal* 37(461-470).
- Diessner, Sebastian, Niccolo Durazzi, and David Hope. 2020. "Liberalization and skill-biased technological change in the knowledge economy: The critical case of german manufacturing." *Max Weber Programme Working Paper*.
- Edmondson, Mark. 2020. "googleLanguageR: Call google's 'natural language' API, 'cloud translation' API, 'cloud speech' API and 'cloud text-to-speech' API." <https://CRAN.R-project.org/package=googleLanguageR>.
- Emmenegger, Patrick. 2021. "Agency in historical institutionalism coalitional work in the creation, maintenance, and change of institutions." *Theory and Society*.
- Entman, Robert M. 1993. "Framing: Toward clarification of a fractured paradigm." *Journal of Communication* 43(4): 51–58.
- Frey, Carl Benedikt. 2019. *The technology trap: Capital, labor, and power in the age of automation*.
- Goodman, Peter S. 2017. "The robots are coming, and sweden is fine." <https://www.nytimes.com/2017/12/27/business/the-robots-are-coming-and-sweden-is-fine.html>.
- Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political Analysis* 21(03): 267–297.
- Haipeter, Thomas. 2020. "Digitalisation, unions and participation: The german case of 'industry 4.0'." *Industrial Relations Journal* 51(3): 242–260.
- Hall, Peter A. 2016. "Politics as a process structured in space and time." In *The oxford handbook of historical institutionalism*, eds. Karl-Orfeo Fioretos, Tulia G. Falletti, and Adam D. Sheingate. Oxford: Oxford University Press, p. 31–50.
- Hall, Peter A., and Kathleen Thelen. 2009. "Institutional change in varieties of capitalism." *Socio-Economic Review* 7(1): 7–34.

- Hay, Colin, and Ben Rosamond. 2002. "Globalization, european integration and the discursive construction of economic imperatives." *Journal of European Public Policy* 9(2): 147–167.
- Hicks, Alexander, and Lane Kenworthy. 1998. "Cooperation and political economic performance in affluent democratic capitalism." *American Journal of Sociology* 103(6): 1631–1672.
- Katzenstein, Peter J. 1985. *Small states in world markets: Industrial policy in europe*. Ithaca, NY: Cornell University Press. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=972827>.
- Kjær, Peter, and Ove K. Pedersen. 2001. "Translating liberalization: Neoliberalism in the danish negotiated economy." In *The rise of neoliberalism and institutional analysis*, eds. John L. Campbell and Ove K. Pedersen. Princeton: Princeton University Press, p. 219–248.
- Lloyd, Caroline, and Jonathan Payne. 2019. "Rethinking country effects: Robotics, AI and work futures in norway and the UK." *New Technology, Work and Employment* 34(3): 208–225.
- Locke, Richard M., and Kathleen Thelen. 1995. "Apples and oranges revisited: Contextualized comparisons and the study of comparative labor politics." *Politics & Society* 23(3): 337–367.
- McKinsey. 2018. "The rise of digital challengers: How digitization can become the next growth engine for central and eastern europe: Perspective on poland." https://digitalchallengers.mckinsey.com/files/Rise_of_Digital_Challengers_Perspective_Poland.pdf.
- Nicholls, Tom, and Pepper D. Culpepper. 2020. "Computational identification of media frames: Strengths, weaknesses, and opportunities." *Political Communication*: 1–23.
- Ornston, Darius. 2012. *When small states make big leaps: Institutional innovation and high-tech competition in western europe*. Ithaca: Cornell University Press.
- Pierson, Paul. 2000. "Increasing returns, path dependence, and the study of politics." *American Political Science Review* 94(02): 251–267.
- Pontusson, Jonas. 2011. "Once again a model: Nordic social democracy in a globalized world." In *What's left of the left*, eds. James E. Cronin, George Ross, and James Shoch. Durham: Duke University Press, p. 89–115.

- Prassl, Jeremias. 2018. *Humans as a service: The promise and perils of work in the gig economy*. Oxford: Oxford Univ. Press.
- Prokop, Jakub. 2018. "Mentality 4.0. How poland is (not)embracing the digital revolution - mentality 4.0. How poland is (not)embracing the digital revolution." <https://www.eecpoland.eu/2018/en/news/mentality-4-0-how-poland-is-not-embracing-the-digital-revolution,334665.html>.
- Proksch, Sven-Oliver et al. 2019. "Multilingual sentiment analysis: A new approach to measuring conflict in legislative speeches." *Legislative Studies Quarterly* 44(1): 97–131.
- Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2018. "Stm: R package for structural topic models." *Journal of Statistical Software* 10(2).
- Roberts, Yvonne. 2018. "The tiny union beating the gig economy giants." <https://www.theguardian.com/politics/2018/jul/01/union-beating-gig-economy-giants-iwgb-zero-hours-workers>.
- Schmidt, Vivien A. 2010. "Analyzing ideas and tracing discursive interactions in institutional change: From historical institutionalism to discursive institutionalism."
- Schmidt, Vivien A. 2008. "Discursive institutionalism: The explanatory power of ideas and discourse." *Annual Review of Political Science* 11(1): 303–326.
- Schwab, Klaus. 2016. *The fourth industrial revolution*. Geneva: World Economic Forum.
- Serafin, Marcin. 2019. "Contesting the digital economy: Struggles over uber in poland." In *The contested moralities of markets*, eds. Simone Schiller-Merkens and Philipp W. Balsiger. Bingley: Emerald Publishing, p. 187–201.
- Skóra, Maria. 2018. "Poland: Developing a smart digital agenda." In *Work in the digital age*, eds. Max Neufeind, Jacqueline O'Reilly, and Florian Ranft. London: Rowman & Littlefield International, p. 461–470.
- Söderqvist, Fredrik. 2017. "How to unite unions, platforms and government through algorithms?" <https://www.youtube.com/watch?v=We3LV-xD9ok>.
- Söderqvist, Fredrik. 2018. "Will history lead the way in the age of robots and platforms?" In *Work in the digital age*, eds. Max Neufeind, Jacqueline O'Reilly, and Florian Ranft. London: Rowman & Littlefield International, p. 295–304.

- Steinmo, Sven, and Kathleen Thelen. 1992. "Historical institutionalism in comparative politics." In *Structuring politics*, eds. Sven Steinmo and Kathleen Thelen. Cambridge: Cambridge Univ. Press, p. 1–32.
- Streeck, Wolfgang, and Kathleen Thelen. 2005. "Introduction: Institutional change in advanced political economies." In *Beyond continuity*, eds. Wolfgang Streeck and Kathleen Thelen. Oxford: Oxford Univ. Press, p. 1–39.
- Thelen, Kathleen. 2018. "Regulating uber: Comparative perspectives on work and welfare in the 'gig' economy." *Perspectives on Politics* 16(4): 938–953.
- Thelen, Kathleen. 2019. "Transitions to the knowledge economy in germany, sweden, and the netherlands." *Comparative Politics* 51(2): 295–315.
- Trumbull, Gunnar. 2004. *Silicon and the state: French innovation policy in the internet age*. Washington/D.C: Brookings Institution Press.
- UK Government. 2018a. "Good work: A response to the taylor review of modern working practices." https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/679767/180206_BEIS_Good_Work_Report__Accessible_A4_.pdf.
- UK Government. 2018b. "Millions to benefit from enhanced rights as government responds to taylor review of modern working practices." <https://www.gov.uk/government/news/millions-to-benefit-from-enhanced-rights-as-government-responds-to-taylor-review-of-modern-working-practices>.
- Vries, Erik de, Martijn Schoonvelde, and Gijs Schumacher. 2018. "No longer lost in translation: Evidence that google translate works for comparative bag-of-words text applications." *Political Analysis* 5: 1–14.
- Wallin, Gunhild. 2017. "A nordic model for fair platform economies." www.nordiclabourjournal.org/i-fokus/in-focus-2017/sharing-economy-2017/article.2017-05-22.1940247186.
- Whitley, Richard. 2007. "Varieties of institutionalism and their problems: Some comments on john campbell's institutional change and globalization." *Socio-Economic Review* 5(3): 548–556.
- Wickham, Hadley et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4(43): 1686.

- Wijffels, Jan. 2019. "Textrank." <https://cran.r-project.org/web/packages/textrank/textrank.pdf>.
- Wueest, Bruno. 2013. "Varieties of capitalist debates: How institutions shape public conflicts on economic liberalization in the united kingdom, germany and france." *Comparative European Politics* 11(6): 752–772.
- Young, Lori, and Stuart Soroka. 2012. "Affective news: The automated coding of sentiment in political texts." *Political Communication* 29(2): 205–231.

Appendix

A Corpus Collection

A.1 Communicative Discourse

The communicative discourse consists of 6235 newspaper articles collected from 40 newspapers via Factiva and NexisUni (see Table Table 3 for an overview). We have articles from most important quality newspapers across the political spectrum. We could not obtain articles from Frankfurter Allgemeine Zeitung and Libération in particular, but we do have other center-right and center-left newspapers from Germany and France. To be part of our sample, articles had to contain at least one reference to the

gig economy **OR** (automation/robotization/artificial intelligence) **AND** (job **OR** jobs **OR** work)
OR fourth industrial revolution **OR** industry 4.0.

The specific search strings for the non-english-speaking countries were as follows:

France: [gig economy **OR** uberisation] **OR** [[travail or emploi] **AND** [robotisation **OR** automatisation **OR** numérisation **OR** Intelligence Artificielle]] **OR** [quatrième* réolutio* industriell* **OR** Industrie 4.0]

Germany: gig economy **OR** [[arbeit or arbeitspl*] **AND** [automatisierung **OR** robot* **OR** künstliche* intelligenz]] **OR** [viert* industr* revolution **OR** Industrie 4.0]

Italy: gig economy **OR** [lavoro **AND** [intelligenza artificiale **OR** automazione **OR** robotizzazione]] **OR** [Industria 4.0 **OR** Impresa 4.0 **OR** Quarta rivoluzione industriale]

Poland: gig economy **OR** [[prac* **OR** zatrudnieni*] **AND** [automatyzacj* **OR** robotyzacj* **OR** sztucz* inteligencj*]] **OR** [przemysł 4.0 **OR** platform* przemysłu przyszłości **OR** Czwart* rewolucj* technologic*]

Spain: gig economy **OR** [trabajo **AND** [automatización **OR** robotización **OR** inteligencia artificial]] **OR** [Industria 4.0 **OR** cuarta revolución industrial]

Sweden: [gig-ekonomi **OR** gig-jobb **OR** gig economy] **OR** [[jobb or arbete] **AND** [automation **OR** automatisering **OR** robot* **OR** artificiell intelligens]] **OR** [Industry 4.0 **OR** industri 4.0 **OR** fjärde industriella revolutionen]

To make it more likely that articles were really about these topics and did not just contain a single mention somewhere in an otherwise unrelated text, these search strings had to appear in the headline or the lead paragraph of articles. The choice of search strings was meant to ensure that the discourse on the digital future of work is represented in its entirety. In countries where gig economy is uncommon and did not yield many results, we included equivalent language-specific terms such as *ubérisation* in France. For each country, we tried different combinations of search strings and manually checked whether they work, i.e.: do they result in a number of articles comparable, given the country's size, to other countries? Do they return articles that cover what we are interested in?

Another potential problem is that the term *industry 4.0* is relatively specific (but not exclusive) to some countries (Germany, Italy). We therefore also included the term *fourth industrial revolution* (in addition to robotization and automation), which should ensure that a higher proportion of industry reference really is due to the fact that these countries talk more about industrial issues – and not due to the fact that there is a terminological bias in our sample. A final problem is that the term *gig economy* is used more broadly in the UK (and Ireland), and not just referring to digital labor platforms. We do want to capture when gig work and digital gig work are discussed together, but we do not want articles that just talk about gig work in ways that articles in other countries would talk about precarious work. To remedy this, we removed articles that did not contain at least one reference to specifically digital aspects of the gig economy (such as Uber, digital platform, etc.). We also skimmed over all articles and manually removed those that were still not really related to our topic of interest. In a last step, we removed very long (>30000 characters) and very short (<200 characters) articles. All in all, our goals was to make the articles included in our sample as relevant as possible. This means much less noise in our data than had we simply tried to maximize the number of articles based on relatively general search strings.

Table 3: Newspaper Corpus Overview

Newspaper	Newspaper Leaning	Number of Articles
France		
L'Humanité	(center-)left	93
Les Echos	(center-)right	389
Le Figaro	(center-)right	119
La Croix	(center-)right	30
Le Monde	center-left	138
L'Opinion	centrist	85
Capital Finance	centrist	12
Germany		
Süddeutsche Zeitung	(center-)left	233
Der Tagesspiegel	(center-)left	135
Deutschlandfunk	(center-)left	76
DIE ZEIT	(center-)left	45
taz - die tageszeitung	(center-)left	33
Die Welt	(center-)right	239
Handelsblatt	centrist	466
Ireland		
The Irish Times	(center-)left	103
Irish Independent	(center-)right	42
The Irish Examiner	centrist	24
Italy		
La Repubblica	(center-)left	264
Il Sole 24 Ore	(center-)right	400

Table 3: Newspaper Corpus Overview (*continued*)

Newspaper	Newspaper Leaning	Number of Articles
Corriere della Sera	centrist	656
La Stampa	centrist	103
Poland		
Rzeczpospolita	(center-)right	290
Gazeta Wyborcza & Wyborcza.pl	centrist	109
Gazeta.pl	centrist	44
Dziennik Gazeta Prawna	centrist	24
Wyborcza.biz	centrist	18
Spain		
El País	(center-)left	330
El Periódico	(center-)left	93
La Vanguardia	(center-)right	194
ABC	(center-)right	93
El Mundo	centrist	81
Sweden		
Aftonbladet	(center-)left	23
Svenska Dagbladet	(center-)right	100
Dagens Nyheter	centrist	190
Business Post	centrist	80
United Kingdom		
Guardian/Observer	(center-)left	264
Telegraph/Sunday Telegraph	(center-)right	245
Financial Times	centrist	191
Independent/Independent on Sunday	centrist	181

A.2 Coordinative Discourse

Our coordinative discourse consists of 2337 documents (see Table 4 for an overview). To enter our corpus, documents had to directly deal with either digitalization of production processes (automation, robotization, Industry 4.0) or new, digitally-enabled forms of work (platform economy, gig work). While the bulk of documents are standalone, some of them are the result of splitting apart longer documents with several sub-chapters. The documents were published between 2012 and 2019 (see Figure 4 for an overview).

Documents were manually collected from the websites of policy actors. We focused on the major social partners, namely governments, employer organizations, and trade unions. The government category includes relevant ministries such as those charged with innovation, work, technology, or social policy. Generally, we selected employer organizations that are part of Business Europe. But since our goal is to obtain documents from the most representative policy actors in a country, we also used – when available – qualitative data from Diresoc Project, whose country reports provide important information about the importance of specific social partners. When an organization did not belong to Business Europe, but was mentioned by Diresoc country reports, we included it. We also gave centrality to representativeness when selecting trade unions.

We chose unions that are part of the European Trade Union Confederation (ETUC). Following the same logic as in the case of employer organizations, we triangulated this criterion with Diresoc country reports. In addition, if actors were considered relevant based on country-specific knowledge but were not part of our sample so far, we still included them as in the case of the German IGMetall or the British IWGB. Due to linguistic barriers, we asked to native speakers to collect policy documents on Poland and Sweden. We asked them to follow the guidelines described above.

Table 4: Policy Documents Corpus Overview (Note that documents are of unequal length).

Actor	Actor Type	Number of Documents
France		
CFDT	Union	7

Table 4: Policy Documents Corpus Overview (Note that documents are of unequal length). *(continued)*

Actor	Actor Type	Number of Documents
CGT	Union	20
CNN	Government	123
COE	Government	22
CPME	Employer	6
FO	Union	69
Government	Government	15
MEDEF	Employer	4
MEF-MRP	Government	27
MEFI-MBCR	Government	54
MT	Government	8
Socialpartners	Union	16
U2P	Employer	11
Germany		
BDA	Employer	12
BDI	Employer	5
BMAS	Government	66
BMBF	Government	11
BMWi	Government	24
BR	Government	3
BVMW	Employer	1
DGB	Union	62
IGBCE	Union	28
IGMetall	Union	19
INQA	Government	1

Table 4: Policy Documents Corpus Overview (Note that documents are of unequal length). *(continued)*

Actor	Actor Type	Number of Documents
N3etwerk	Employer	1
VDMA	Employer	1
verdi	Union	18
Ireland		
DBEI	Government	34
DCENR	Government	9
IBEC	Employer	19
ICTU	Union	2
Italy		
CGIL	Union	42
CGILCISLUIL	Union	1
CISL	Union	8
Confindustria	Employer	157
FDV	Union	4
INAPP	Government	9
MISE	Government	7
MLPS	Government	2
Poland		
BCC	Employer	5
GovernmentFPPP	Government	1
GovernmentMC	Government	8
GovernmentMinisters	Government	2
GovernmentMPT	Government	7
GovernmentMRPPS	Government	1

Table 4: Policy Documents Corpus Overview (Note that documents are of unequal length). (*continued*)

Actor	Actor Type	Number of Documents
GovernmentV4	Government	1
Lewiatan	Employer	13
NSZZSolidarnosc	Union	4
OPZZ	Union	4
PracodawcyRP	Employer	6
Spain		
CC	Employer	6
CCOO	Union	50
CEOE	Employer	76
Cepyme	Employer	348
ConsEcoSoc	Government	12
FiM	Union	39
MEE	Government	13
MEETA	Government	52
MESS	Government	46
MIET	Government	46
MIET-MHAP	Government	7
MITC	Government	12
MTMSS	Government	23
UGT	Union	45
Sweden		
Gov	Government	297
LO	Union	15
SACO	Union	56

Table 4: Policy Documents Corpus Overview (Note that documents are of unequal length). *(continued)*

Actor	Actor Type	Number of Documents
SvensktNar	Employer	9
TCO	Union	38
United Kingdom		
CBI	Employer	11
Commons	Government	6
DDCMS	Government	11
FSB	Employer	24
Gov	Government	32
IWGB	Union	4
Taylor	Government	14
TUC	Union	65

Figure 4: Number of Documents per Year (Note that documents are of unequal length).

B. Details on Methods

B.1 Sentiment Analysis

In the paper, we use the Lexicoder Sentiment Dictionary developed by Young and Soroka (2012). While the dictionary was specifically designed for political text and is therefore be appropriate for our corpora, we cross-validated our findings with other dictionaries given the validity issues associated with dictionary-based analysis (Boukes et al. 2020; van Atteveldt, van der Velden, and Boukes 2021). For these dictionaries, we complemented the dictionary-based approach to sentiment analysis with natural language process in order to allow us to account for negators, amplifiers and deamplifiers. We used standard negators, amplifiers and deamplifiers obtained from the lexicon package (Rinker 2018). Specifically, we used the following four widely used and relatively general dictionaries to cross-validate our findings:

- the AFINN dictionary developed by (Nielsen 2011);
- the Bing sentiment lexicon developed by Hu and Liu (2004);
- the NRC Word-Emotion Association Lexicon developed by Mohammad and Turney (2010); and
- the syuzhet dictionary (and accompanying R package) developed by Matthew Jockers and the Nebraska Literary Lab (Jockers 2017).

In case of the NRC dictionary, we did not use their emotion-based dictionaries (such as those for anger, fear, or hope) but only those for positive and negative terms. The *afinn* and *syuzhet* dictionaries have a higher resolution than *bing* and *nrc* by scoring words not just as negative (-1) and positive (+1) but allowing more gradation. We transformed these higher-resolution scales to a binary positive-negative score for both technical reasons and to provide more robust estimates. After all, whether a word is positive or negative is a much more straightforward question than whether a positive word is quite (+3), very (+4) or extremely (+5) positive, especially across different contexts and for machine-translated documents.

The strong similarities across dictionaries make us confident that the differences we measure are real differences and not driven by the particularities of any one of these dictionaries (see Figure 4). This is particularly true for the larger differences, i.e., the bottom and top of the distribution, on which we focus in the paper (we don't, for example, substantively discuss the small differences between Germany, France and Spain).

Figure 5 shows the 50 most frequently matched positive and negative words for the different dictionaries. This is meant to both illustrate what positive and negative tone mean and to validate the dictionaries. As is evident from the plot, the most commonly matched words are what one would expect to appear in positive and negative framings of the digital future of work.

Figure 5: Discursive sentiment across countries, dictionaries and discourse types with 95 % bootstrapped confidence intervals (R = 10000)

Figure 6: Wordclouds with the most commonly matched negative and positive words

B.2 Keyword Extraction Techniques

RAKE The RAKE – or rapid automatic keyword extraction – algorithm starts from the idea that keywords usually contain several informative words but rarely punctuation or stopwords (Rose et al. 2010). It thus first tokenizes a given text, using spaces and punctuation to break it at word delimiters. It then creates sequences of contiguous words, so called candidate keywords. For example, the short text “robots and artificial intelligence will change the nature of work and our work life balance” will be tokenized as follows:

[“robots,” “and,” “artificial,” “intelligence,” “will,” “change,” “the,” “nature,” “of,” “work,” “and,” “our,” “work,” “life,” “balance”]

Reading from left to right, the algorithm then creates candidate keywords every time a common stopword is encountered, like this

[“robots,” “artificial intelligence,” “change,” “nature,” “work,” “work life balance”]

Next, a score is created using the following formula:

$$wordscore = \frac{degree(word)}{frequency(word)}$$

Frequency refers to the number of times a word appears in the list of candidate keywords.

$frequency(robots) = 1,$

$frequency(artificial) = 1,$

$frequency(intelligence) = 1,$

$frequency(change) = 1,$

$frequency(nature) = 1,$

$frequency(work) = 2,$

$frequency(life) = 1,$

$frequency(balance) = 1$

Degree refers to how frequently a word co-occurs with other candidate keywords in a given text. This is equivalent to the number of times a word occurs in the candidate keywords that contain this word (a higher word degree can therefore also indicate that it appears in longer candidate keywords):

$$\text{frequency}(\text{robots}) = 1,$$

$$\text{frequency}(\text{artificial}) = 2,$$

$$\text{frequency}(\text{intelligence}) = 2,$$

$$\text{frequency}(\text{change}) = 1,$$

$$\text{frequency}(\text{nature}) = 1,$$

$$\text{frequency}(\text{work}) = 4,$$

$$\text{frequency}(\text{life}) = 3,$$

$$\text{frequency}(\text{balance}) = 3$$

As we saw, the word score is proportional to the degree of word and inversely proportional to its frequency. RAKE thus favors words that occur not too frequently but often in combination with other keywords, especially long ones. To calculate the candidate keyword score, we add the word scores of its constituent parts and take the highest-scoring percent (e.g. the highest-scoring 33%) as our keywords to be extracted. For our example, this would look as follows:

$$\text{score}(\text{robots}) = 1/1 = 1,$$

$$\text{score}(\text{artificialintelligence}) = (\text{word}_s\text{core}(\text{artificial}) = 2/1 = 2 + \text{word}_s\text{core}(\text{intelligence}) = 2/1 = 2) = 4$$

$$\text{score}(\text{change}) = 1/1 = 1,$$

$$\text{score}(\text{nature}) = 1/1 = 1,$$

$$\text{score}(\text{work}) = 4/2 = 2,$$

$$\text{score}(\text{worklifebalance}) = (\text{word}_s\text{core}(\text{work}) = 4/2 = 2 + \text{word}_s\text{core}(\text{life}) = 3/1 = 3 + \text{word}_s\text{core}(\text{balance}) = 3/1 = 3) = 8$$

This means that “work life balance” and “artificial intelligence” were selected as our most central keywords. Rake was implemented using the UDPipe R package (Straka and Straková, 2017).

Co-occurrences We also constructed bigrams based on their number of co-occurrences, either within a skipgram of size four or within a sentence. In the former case (reported in the paper), we counted bigrams when words followed one another directly or when we skipped up to three words in between. In the latter case (reported here), we counted bigrams when they occurred within the same sentence. The results are largely similar to those based on skipgrams, although produce somewhat more sensible results since the sentence restrictions is less sensible as the skipgram-context restrictions (as the latter is narrower).

Textrank Due to space constraints, we did not include Textrank-based keywords in the paper. However, as Figure shows, the words with the highest PageRank score support our general findings. Textrank is a graph-based ranking model that identifies keywords by constructing a word network based on whether two words follow one another (or co-occur in a window of N words) (Mihalcea and Tarau 2004). If they do, an edge is created between them, the weight of which depends on how often they follow each other in a given text. Next, the PageRank algorithm is applied to this network to rank words in their order of importance (this is what is reported in Figure). Relevant words that follow one another can then be combined to obtain keywords. Textrank was implemented using the R textrank (Wijffels 2019).

Figure 7: Keywords with the highest PageRank score

B.3 Word Vectors

In addition to the analysis presented in the paper, we also use word vectors to identify semantic similarities among words and to assess whether they systematically differ across countries for key terms of interest. Word vectors are based on the idea that words that appear near each other have similar meanings. Or, in John Rupert Firth famous phrase, we ‘know a word by the company it keeps’ (Spirling and Rodriguez 2019). Words are mapped to a multidimensional vector that represents its ‘meaning,’ with the different dimensions representing different semantic or syntactic connotations of a word. Somewhat unorthodoxly, we used simple word counts and matrix factorization to calculate word vectors (Moody 2017).

We used a moving slide window to create skipgrams of length eight and then calculated the pointwise mutual information (PMI) metric for all word pairs (see, Silge 2017). This gives us information about which words occur together (within a moving window of size 8) more often than expected based on how often they occur on their own. PMI is simply the logarithm of the normalized skipgram

probability which in turn is the result of dividing the frequency of two words co-occurring in a skipgram by the frequencies of their individual occurrence. This is a well-understood metric for how frequently two words occur jointly rather than independently (Moody 2017). We then cast a sparse matrix where each row represents word 1 and each column word 2 and each value is the PMI calculated above. From there, we reduce the dimensionality of that matrix using singular value decomposition, specifying the number of principal components to 300 – a value common in the word embeddings literature. Using a tidy approach, we now have a tibble of word vectors which we can use to find synonyms (or analogues, etc.).

Here, we first provide additional examples for this synonym task which demonstrate that our approach reliably finds highly plausible synonyms. As Table 5 shows, our approach works surprisingly well given that it is much less complex and computationally demanding than using neural networks, and given the relatively small size of the data. We can then do the same for terms that are of more theoretical interest. Table 6 depicts a number of such words. Basically, what we do here is to zoom in on words of interest by looking at nearby vectors in a vector space, that is, by looking at synonyms. We see, for example, that in Italy and Germany terms like ‘4.0’ or factory are closely associated with future and work, reflecting the importance for both countries of getting digital manufacturing right. In line with what we found in the paper, in German discourse the future is viewed in more positive terms, with work-life-balance issues taking up a prominent place (time, family, unpaid). The French discourse, to give a second illustrative example, reflects the political commitment to making France a digital frontrunner, and the concomitant entrepreneurial role ascribed to the state in general and the National Digital Council (Conseil National du Numérique) in particular. Accordingly, words like project, program and plan are associated with the future, and words like ambition, national, or public are associated with digital. A final example are the UK and Ireland where work is closely associated with insecurity, reflecting the often-precarious nature of these countries’ labor markets.

Table 5: Synonyms for capital, gates, google, and public

country	capital	gates	google	public
France	capital, venture, funds,	gates, musk, elon, bill,	google, apple, facebook,	public, innovation,
	fund, investment, ventures,	stephen, hawking, boss, ne,	amazon, search, microsoft,	procurement, private,
	partners, private, equity,	pas, tesla, bezos, ou, late,	2012, maps, engine, twitter,	services, data, service,
	angels, risk, seed, fcpr,	etats, spacex	gmail, hotel, youtube, beat,	action, digital, economy,
	stage, scr		2010	support, policies, social, purchasing, participation
Germany	capital, venture,	gates, founder, boss, bill,	google, facebook, apple,	public, administration,
	investment, fund, wealth,	tesla, microsoft, day, musk,	amazon, microsoft, giants,	service, transport, services,
	investors, start, ups,	elon, fair, cyberattacks,	valley, alphabet, search,	scientific, insurance,
	money, profits, short,	thousands, shareholders,	corporations, alexa, silicon,	private, defense, wlan,
	investments, taxes, income,	hanover, march	uber, pichai, ibm	million, social, financial,
	invested			sector, debate
Ireland	capital, allowances,	gates, bill, founder, musk,	google, facebook, amazon,	public, sector, services,
	venture, relief, accelerated,	elon, microsoft, apple,	translate, search, twitter,	consultation, private,
	tax, credit, assets,	david, william, attendees,	linkedin, ebay, youtube,	healthcare, service,
	investors, funding,	german, steve, stephen,	recognition, marketplaces,	administration, bodies,
	efficient, start, expenditure,	japanese, phrase	siri, card, offline, marshall	community, security,
	ventures, investment			office, i.i, costs, national
Italy	capital, human,	gates, bill, elon, musk,	google, facebook, amazon,	public, private,
	investments, venture,	founder, stephen, hawking,	microsoft, giants,	administration,
	equity, investment, foreign,	robots, tesla, robot, techno,	intelligence, artificial,	investment, financial,
	funds, tangible, fixed,	physicist, fear, income,	apple, bezos, glass, silicon,	national, research, funds,
	assets, intangible, loans,	including	valley, jeff, ceo, cars	intervention, finance,
	banks, depreciation			policies, support, debt, social, investments
Poland	capital, funds, german,	gates, bill, stephen, elon,	google, duplex, assistant,	public, administration,
	foreign, venture, investors,	musk, fears, hawking,	pichai, facebook, sundar,	sector, private, entities,
	deloitte, investment, fund,	previously, shared, race,	giant, tesla, amazon, apple,	entity, information,
	human, vc, startup,	expressed, warned, late,	ai, conversation, alphabet,	consultations, disclosure,
	financed, ventures,	contact, famous	wave, presentation	procurement, official, inter,
	investor			isp, governmental, website

Table 5: Synonyms for capital, gates, google, and public (*continued*)

country	capital	gates	google	public
Spain	capital, venture, barcelona,	gates, founder, bill,	google, amazon, microsoft,	public, private,
	ventures, madrid, human,	microsoft, rowe,	apple, facebook, samsung,	administration,
	club, corporate, 25ok, 1m,	philanthropist, ii, iii,	huawei, artificial, alphabet,	administrations, bodies,
	seed, 5ook, 2m, fund, 5m	farming, richard, robots,	tecnologia, twitter, hp,	services, system,
		price, page, firm, vertical	intel, intelligence, netflix	participation, institutions, collaboration, information, investment, agencies, entities, policies
Sweden	capital, venture, human,	gates, founder, bill, musk,	google, microsoft,	public, sector, private,
	financing, almi, inland,	elon, steven, recently, tesla,	facebook, amazon, ibm,	procurement, business,
	loan, invest, guarantees,	silicon, valley, hawking,	ericsson, google's, apple,	service, authorities,
	gains, subtotal, sme,	ingvar, stephen, physicist,	behshad, behzadi, news,	publications, actors,
	productivity, private,	henry	giants, search, volvo,	activities, transport,
	investments		investing	administration, customers, publicly, services
United Kingdom	capital, venture, labour,	gates, bill, elon, musk,	google, facebook, google's,	public, sector, local,
	funds, investment, equity,	stephen, hawking, founder,	deepmind, microsoft,	services, private, policy,
	tech, investors, fund,	mark, zuckerberg, warned,	search, google's, siri,	libraries, radio, content,
	markets, firm, deals,	billionaires,	amazon, apple, alphabet,	investment, finances,
	wealth, closed, trading	www.independent.co.uk, tax, ceo, pizza	schmidt, deep, owned, bought	england, health, report, test

Table 6: Word vector similarities for different keywords across countries

country	digital	future	work
France	digital, economy, ambition,	future, industry, project, factory,	digital, time, economy, life,
	innovation, technology, companies,	tomorrow, program, education,	employees, workplace, employment,
	public, council, society, literacy,	european, world, labor, industrial,	skills, information, mission, home,
	france, national, french, appendices,	investments, impact, launched, plan	workers, labor, autonomy,
	sectors		technologies
Germany	digital, industry, digitization, 4.o,	future, 4.o, industry, factory, digital,	digital, world, employees, family,
	change, world, data, transformation,	world, people, digitization, change,	future, 4.o, time, life, industry,
	employees, capitalism, platforms,	development, machines, social,	intensity, people, digitization,
	technology, future, virtual, business	german, market, germany	unpaid, hours, forms

Table 6: Word vector similarities for different keywords across countries (*continued*)

country	digital	future	work
Ireland	digital, transformation, economy, internet, ai, market, technologies, single, related, business, workers, technology, innovation, infrastructure, ecosystem	future, expert, study, legal, market, systems, technologies, ability, world, 1, current, ibec, europe, trends, 27	insecure, digital, life, contracts, hours, smarter, development, one's, plan, scope, appendix, flexible, forms, insight, placements
Italy	digital, innovation, 4.o, industry, hubs, transformation, hub, technological, technologies, economy, dii, platforms, industrial, business, manufacturing	future, automation, factory, robots, revolution, interventions, people, 4.o, frey, osborne, 1, oxford, industry, press, articoli	digital, people, organization, workers, employment, production, hours, 4.o, training, labor, paper, forms, tasks, agile, world
Poland	digital, transformation, competences, innovation, economy, competence, technologies, solutions, priority, kazimierz, talents, industrial, identified, business, i.e	future, foundation, report, professions, platform, job, industry, labor, en, act, automation, field, jobs, 4.o, offers	intelligence, employees, people, artificial, automation, services, repetitive, data, routine, physical, employee, performing, actions, poland, labor
Spain	digital, industry, transformation, spain, digitization, 4.o, business, economy, companies, technologies, sector, ametic, plan, society, technology	future, author, lab, accelerator, revolution, market, bic, seedrocket, current, challenges, ventures, industrial, robots, initland, life	workers, life, hours, health, organization, jobs, people, environment, home, time, hygiene, labor, psychosocial, assessment, business
Sweden	digital, agenda, platforms, competence, services, digitization, skills, service, market, regional, economy, commission, 2015, agendas, development	future, society, commission, digitalisation, digitization, challenges, report, intelligence, space, employment, skills, artificial, world, researcher, competence	environment, digital, labor, performed, tasks, economy, market, time, tax, employment, responsibility, 24, business, platforms, workplace
United Kingdom	digital, skills, businesses, uk, technology, business, connectivity, firms, dcms, smes, infrastructure, world, future, voice, plans	future, digital, automation, http, skills, jobs, shaping, report, 5g, world, economy, infrastructure, digitalisation, tech, intelligence	people, jobs, insecure, hours, world, economy, time, life, job, balance, future, technology, employment, activities, gig

B.4 Topic Modelling

Preprocessing For pre-processing, we used annotated part-of-speech tags to select nouns, adjectives, and verbs. We discarded punctuation and stopwords as well as semantically less meaningful parts of speech like determiners or names entities like dates. This seems justifiable, given that we are not interested in linguistic style or subtle word use but rather in the broad thematic contours of discourse. We also discarded location-specific information like capital cities or languages to avoid linguistically (as opposed to substantively induced) country-effects.

We also constructed a list of frequent n-grams such as artificial intelligence, machine learning, virtual reality, big data, tax evasion, further training etc. This list, which contains 82 n-grams, was manually compiled based on the most frequent collocations identified in the text (with log-frequency biased mutual dependency used as the ordering metric). We see no reason not to include such information as there are obvious theoretical reasons to prefer such n-grams to their separate constitutive unigrams.

We lowercased but did not stem our document feature matrix as the difference between singular and plural forms can be meaningful while the difference between uppercased and lowercased words is most likely not – at least in the types of policy and newspaper documents we are looking at. For example, it can make a difference whether a text speaks of robots in the plural – as in the abstract threat that robots pose ('the robots will take our jobs') – and a robot in the singular, which is more likely to be described as something useful or positive ('the robot does x').

We also removed remaining word trash such as html tags, common untranslated words (e.g. della), as well as country-specific information using the named entity information. This latter removal is meant to ensure that differences in topic prevalence are, as much as possible, the result of substantive differences and not of local vernaculars or parochial word usage.

Lastly, we removed words that appeared in more than 50% of documents as such words do not contain much information (e.g. digital). We also removed words that appeared in less than 0.5% of documents. While this is a somewhat arbitrary (although commonly used) standard, qualitative

inspection revealed but proved to be a useful threshold that removed many very specific and rare terms while still retaining un-common but not unimportant words.

Number of Topics We chose a topic model with $k = 60$ topics. While our decision was assisted by several metrics, the choice was ultimately a theoretical one, based on two criteria. First, given that we are primarily interested in comparing the content of discourse, we want our topics to be broad enough to be at least potentially relevant in different countries, but narrow enough to capture the themes we are interested in (e.g., automation vs compensation). On the one hand, we want obtain topics that are broad enough to be at least potentially relevant in different countries and for different actors. If we chose $k=350$, for example, we might get many topics that are about particular events in a country (e.g. the introduction of a new technology at a particular company), and will most likely not be discussed in other countries. On the other hand, we want enough topics to allow for meaningful differences to emerge. If we only had, say, 5 topics, these will be too coarse to say anything interesting about cross-country or other differences. Second, while topics are often considered “as an operationalization of policy frames” (Gilardi, Shipan, and Wüest 2020, 3), topic models easily uncover more sensible topics than there could possibly be frames (Nicholls and Culpepper 2020, 8). We therefore opted for using a higher number of topics than we would have had we assumed that topics directly capture real-world frames. We then aggregated these topics into “frame packages” (Nicholls and Culpepper 2020, 11) that capture discursive foci that are neither too broad nor too narrow and are therefore theoretically interesting. For example, topics on industry 4.0, the technological transformation of the production process, cloud computing, or smart factories were combined into the frame package ‘digital manufacturing,’ which covers debates on how digitalization is changing industrial production and manufacturing.

Based on our intuition about which k should yield topics that are both broad and interesting, we wanted k to be somewhere between 25 and 75. We therefore run topic models with k s between 10 and 100 so as to allow various quantitative metrics to guide our decision. Figure 8 plots four metrics – semantic coherence, exclusivity, residuals, and held-out-likelihood – for models with different k s. Semantic coherence is a metric that measures how often the most frequent words in a topic actually co-occur

in a document. While semantic coherence has been shown to correlate well with human judgments of topic quality, it has been shown to increase when topics are dominated by very common words (Roberts, Stewart, and Tingley 2018). Exclusivity, by contrast, penalizes models with few dominant top words. It measures the share of top words which are distinct to a given topic, thus creating something of a trade-off with semantic coherence.

The residuals capture overdispersion of the variance of the multinomial in stm's data generating process (Roberts, Stewart, and Tingley 2018). Higher values indicate overdispersed residuals, implying that the latent topics cannot account for the overdispersion and more topics may be needed to use up the extra variance. Held-out likelihood estimates the probability that words appear in a document when these words have been removed before the estimation. It is a measure of predictive performance, with higher values indicating better performance.

Hence, we want semantic coherence, exclusivity, and held-out likelihood to be as high and the residuals to be as low as possible. Topics numbers higher than the number of plausible individual frames on digitalization - of which there can only be so many - are justifiable as we can combine different topics to frame packages (Nicholls and Culpepper 2020). In the end, we settled for 60 topics, which the quantitative metrics supported and which also made sense qualitatively. It has to be said, though, that topics were very similar and stable with 10 or even 15 fewer or more topics.

Figure 8: Model diagnostics for different numbers of topics (10 to 100)

Validation

While topic models are powerful tools “for discovering and exploiting the hidden thematic structure in large archives of text” (Blei 2012, 82) they are no magic bullet and need to be carefully interpreted and validated. Validation requirements are somewhat lower if the main goal is not to measure some pre-specified latent trait but, as is the case here, to explore and usefully summarize texts in order to facilitate comparisons (cf. Ying, Montgomery, and Stewart 2019, 1–2). It is nonetheless crucial to at least make this process transparent and to ideally also establish “semantic validity” (Quinn et al. 2010, 216; Ying, Montgomery, and Stewart 2019, 6) by showing that topics are coherent and make sense to external annotators.

Table 7 shows a complete list of topics and, when applicable, the frame package to which we assigned this topic. We only assigned topics to one of the 6 frame packages presented in the paper, which does not mean that the others could not have been assigned to different frame packages or that other topic-frame-package combinations would not have been possible. The labeling of topics as well as their assignment to frame packages was an interpretative process, based on the most *F*requent and *EX*clusive Words (FREX) and theoretical reasoning. For this reason, we make the topic terms transparent, and we externally verified the semantic validity of our topics. That is, we checked whether they make sense to human readers and whether they can be semantically distinguished from other topics by asking two external researchers to complete two tasks (Chang et al. 2009; Ying, Montgomery, and Stewart 2019). First, we gave them the list below but instead of only the correct category label, we gave the correct label plus two other randomly chosen labels. The task was to pick the correct label. Second, we gave them a list with the top 4 FREX terms plus the highest FREX term from another randomly selected topic. Their task was to correctly identify the ‘incorrect’ word. In the first task, they, on average, picked the right topic label 85 % of the time; in the second task, they chose the non-fitting word 73,3 % of the time. These results compare well with results from similar tasks (Ying, Montgomery, and Stewart 2019) and make us confident that our topic model produces semantically valid topics.

Table 7: Complete List of Topics and Topic Categories

Topic	Top FREX Words	Topic Label	Aggregated Topic Category
1	artists, admission, story, father, artist, love, art, scene, journalist, film, writer, woman, stories, watch, novel, book, friend, exhibition, journalists, famous	Art	/
2	ambition, inclusion, opening, actors, mediation, open, public, literacy, communities, citizens, practices, citizen, empowerment, administrations, encourage, governance, uses, stakeholders, recommends, administrative	Digital Government	Digital Industrial Policy
3	exposure, prevention, occupational, accidents, preventive, safety, psychosocial, risks, accident, chemical, healthy, exposed, dangerous, n°, substances, protective, agents, diseases, fax, aspects	Occupational Safety	Status Platform Workers
4	industrial, industry, revolution, manufacturing, automotive, production, factories, transformation, smart, industries, advanced, leading, energy, robotics, chains, electronics, competitiveness, strategy, 3d_printing, manufacturers	Industry 4.0	Digital Manufacturing
5	revenue, waste, circular_economy, recycling, materials, environmental, packaging, raw_materials, recovery, circular, reuse, emissions, material, water, disposal, collection, energy, carbon, sustainability, renewable	Sustainability	/
6	patients, patient, healthcare, medical, care, hospital, health, medicine, doctors, doctor, hospitals, benefiting, clinical, drug, treatment, diseases, disease, treatments, drugs, cancer	Health Care	/
7	artificial_intelligence, ai, machine_learning, algorithms, intelligence, algorithm, artificial, computer, language, scientists, ethical, computers, ethics, deep, humans, machine, recognition, neural, images, human	Artificial Intelligence	Automation & Compensation
8	differences, table, digitization, index, percentage, degree, proportion, significant, size, family, smes, workload, intensity, difference, higher, scope, affected, expectations, compatibility, variables	Statistics	/
9	trades, uberization, salaried, collaborative_economy, social_dialogue, collective, subordination, mobilization, impacts, cooperative, status, telework, branch, anticipate, profession, generalization, autonomy, disconnect, subject, evolve	Platform Work	Status Platform Workers
10	taxi, ordering, accounting, license, permit, traffic, center, centers, amendment, vehicle, section, authority, holders, shall, vehicles, equipment, regulations, reservations, licenses, permits	Taxi Regulations	/
11	cloud, data, big_data, computing, customer, applications, digital_transformation, solutions, analytics, digital, technologies, blockchain, iot, devices, customers, systems, virtual, technology, processing, things	Cloud Computing	Digital Manufacturing

Table 7: Complete List of Topics and Topic Categories (*continued*)

Topic	Top FREX Words	Topic Label	Aggregated Topic Category
12	entrepreneurs, academics, investors, financing, entrepreneurship, loans, capital, credit, bank, start_ups, startups, loan, entrepreneur, growth, entrepreneurial, funds, venture_capital, academic, investor, startup	Entrepreneurship	/
13	transformation, technological, productive, strategies, competitive, textile, processes, value_chain, sector, sectors, efficiency, process, components, enablers, technologies, digitization, integration, competences, elaboration, structured	Transformation of the Production Process	Digital Manufacturing
14	determination, data_protection, flexibility, flexible, collective_agreements, collective_bargaining, co, working, interests, councils, seize, federal, social_partners, employers, unions, shaping, participation, options, working_conditions, statutory	Work Councils	/
15	year, turnover, quarter, sales, first, months, revenues, largest, last, half, annual, group, third, grow, forecast, total, forecasts, figures, compared, trend	Economic Statistics	/
16	digital_skills, skills, skill, employability, learning, graduates, literacy, training, cognitive, apprenticeship, solving, career, transversal, learn, mastery, apprenticeships, digital_technologies, careers, qualifications, soft	Digital Skills	Digital Investments
17	cars, car, electric, driving, driverless, roads, autonomous, bus, truck, vehicles, trucks, road, vehicle, parking, cities, drones, city, wheel, aviation, flight	Autonomous Vehicles	Automation & Compensation
18	theme, path, capable, born, excellence, explains, innovation, starting, chain, widespread, dedicated, destined, underlines, universities, shapes, choices, center, leap, thanks, hi	Other	/
19	robot, robots, robotic, robotics, humans, arm, arms, humanoid, programmed, hands, brain, beings, science_fiction, vacuum, human, 1950s, body, colleague, movements, moves	Robots	Automation & Compensation
20	profiles, degrees, artistic, occupation, animation, technicians, sound, competences, profile, audiovisual, graphic, content, programming, design, graduated, designer, visual, professionals, training, techniques	Creative Skills	Digital Investments
21	labour, independent, platform, knowledge_intensive, platforms, freelancers, clients, professionals, firms, earnings, freelance, transaction, might, tasks, firm, wage, skilled, temporary, consultants, organising	Independent Platform Work	Status Platform Workers
22	policies, productive, social_dialogue, administrations, competitiveness, reforms, employment, labor, essential, digitization, fabric, adequate, priority, youth, country, commitment, necessary, deficit, regulatory, reform	Competitiveness Reforms	Digital Industrial Policy

Table 7: Complete List of Topics and Topic Categories (*continued*)

Topic	Top FREX Words	Topic Label	Aggregated Topic Category
23	talent, recruitment, managers, candidates, hr, executives, hiring, leadership, recruit, candidate, recruiting, surveyed, teams, respondents, leaders, believe, talents, professionals, management, selection	Human Resources	/
24	higher_education, college, education, university, courses, educational, students, student, vocational, colleges, universities, continuing_education, secondary, doctoral, institutions, programs, school, exam, lifelong_learning, grants	Higher Education	Digital Investments
25	uber, taxi_drivers, drivers, delivery, passengers, restaurants, protest, deliveries, riders, eat, driver, strike, taxis, couriers, app, bicycle, restaurant, passenger, rings, corporations	Gig Work	Status Platform Workers
26	graph, television, video, advertising, music, content, games, audiovisual, publishing, books, consumption, internet, tv, newspapers, turnover, game, billing, film, radio, distribution	Multimedia	/
27	think, thing, say, want, lot, going, look, talk, get, things, much, ca, feel, money, got, bad, tell, ask, happy, go	Other	/
28	jobs, automation, occupations, automated, workforce, skilled, productivity, replaced, wages, wave, inequality, job_losses, likely, rise, risk, advances, economist, roles, routine, economists	Automation	Automation & Compensation
29	sharing_economy, platforms, consumer, transactions, sharing, rental, collaborative_economy, users, rent, legal, platform, trader, law, liability, disputes, housing, dispute, rules, legislation, user	Sharing Economy	/
30	amortization, hyper, investments, incentives, super, depreciation, tax_credit, interventions, machinery, plan, continuation, incentive, maneuver, budget, decree, relief, subsidized, purchase, extension, credit	Investment Incentives	Digital Industrial Policy
31	century, humanity, lives, era, book, societies, capitalism, history, 21st_century, planet, 19th, intellectual, live, philosopher, revolutions, revolution, imagine, man, invention, let	Digital Future	/
32	agenda, member, regional, issues, reports, representatives, adopted, agendas, members, report, proposals, commission, consultation, county, forum, follow, dialogue, meetings, appointed, meeting	EU Agenda	/
33	labor_market, unemployment, proportion, unemployed, part_time, restructuring, educated, unions, structural, salary, increased, decreased, groups, figure, differs, wage, longer, union, extent, fixed	Labor Market Inequality	Automation & Compensation

Table 7: Complete List of Topics and Topic Categories (*continued*)

Topic	Top FREX Words	Topic Label	Aggregated Topic Category
34	smes, promote, projects, initiatives, tourism, actions, promotion, promoting, collaboration, facilitate, initiative, aimed, awareness, entities, implementation, aid, support, objectives, programs, lines	SME Support	Digital Industrial Policy
35	tech, said, chief_executive, voice, chief, centre, firms, might, officer, banks, firm, banking, bank, organisations, founder, seeing, centres, lawyers, biggest, office	Business News	/
36	broadband, infrastructure, connectivity, network, networks, telecommunications, energy, mobile, coverage, smart, operators, connected, deployment, fiber, high_speed, fast, connections, electricity, wireless, buildings	Digital Infrastructure	Digital Investments
37	municipalities, municipal, respondents, extent, individuals, assignments, survey, home, estimated, municipality, libraries, reduction, mail, costs, effect, net, estimate, expected, websites, differences	Municipalities	/
38	union, bargaining, relations, representation, autonomy, citizenship, inequalities, social, institutional, organizational, contractual, democratic, collective, participation, decent, reflection, organization, labor, forms, relationships	Industrial Relations	/
39	incidents, security, attacks, trust, cyber, privacy, expressed, page, answer, cybersecurity, suffered, continuity, attack, communications, micro, files, incident, consulted, damage, survey	Cyber Attacks & Data breaches	/
40	further_training, qualification, operational, employees, councils, mechanical_engineering, department, continuing_education, processes, requirements, learning, networking, topics, agile, participation, map, departments, digitization, involved, employee	Further Training	Digital Investments
41	programme, deliver, businesses, organisations, programmes, ensure, enable, enterprise, ensuring, delivering, adoption, local, wider, approach, address, supporting, engagement, digitalisation, range, vital	Business Support	Digital Industrial Policy
42	free, personal_data, users, user, search, neutrality, web, site, exploitation, sites, online, loyalty, advertising, consent, collection, audience, engine, freedom, blogs, terminals	Data Protection & Internet Regulation	/
43	professions, robotization, disappear, automation, routine, qualified, specialists, jobs, replace, tasks, threatened, original, replaced, job, labor_market, repetitive, profession, machines, experts, author	Automation	Automation & Compensation

Table 7: Complete List of Topics and Topic Categories (*continued*)

Topic	Top FREX Words	Topic Label	Aggregated Topic Category
44	gig_economy, minimum_wage, tribunal, gig, riders, rights, ruling, holiday_pay, protections, insecure, couriers, contractors, contracts, pay, sick, status, workers, worker, entitled, courier	Gig Worker Rights	Status Platform Workers
45	networked, topic, networking, medium_sized, boss, manufacturer, federal, wants, digitalization, mechanical_engineering, sees, politics, says, location, standards, board, corporations, association, enormous, head	Digital Manufacturing	Digital Manufacturing
46	plan, client, clients, brand, team, store, customer, best, marketing, product, ideas, purchase, experience, mind, channels, sell, brands, channel, commerce, moment	Business Strategy	/
47	exports, scenarios, crisis, elaborations, imports, diversification, ranking, economies, manufacturing, weight, goods, trade, foreign, export, added, specialization, geographical, relative, recorded, dynamics	Exports	/
48	president, edition, conference, yesterday, director, meeting, explained, event, organized, vice, held, stressed, deputy, attended, presentation, speakers, dedicated, headquarters, fair, head	High-level meetings	/
49	women, teachers, school, teaching, gender, schools, female, girls, men, teacher, students, children, science, male, parents, mathematics, classroom, boys, stereotypes, educational	Women & STEM	Digital Investments
50	research, efforts, r&d, universities, innovation, researchers, funding, colleges, entrepreneurship, collaboration, grants, climate, environments, scientific, institutes, strategic, institutions, societal, evaluations, grant	Research & Development	Digital Investments
51	factory, sensors, plant, assembly, maintenance, plants, components, factories, machine, warehouse, machines, additive_manufacturing, glasses, production, logistics, printing, parts, mechanical, manufacturers, augmented_reality	Smart Factory	Digital Manufacturing
52	self_employment, work_environment, self-employed, employer, false, investigation, hired, responsibility, client, contractor, employee, persons, work-, staffing, phenomenon, assignments, self-, employed, safety, contractors	(False) Self-Employment	/
53	election, politics, politicians, political, vote, party, left, elections, presidential, voters, reform, governments, minister, immigration, wing, chairman, liberal, reforms, anti, campaign	Elections	/
54	r&d, aid, clusters, expenditure, mission, creators, high_speed, favor, patent, fund, equity, venture_capital, heart, funding, patents, innovative, deployment, incubators, funds, seed	Industrial Policy	Digital Industrial Policy

Table 7: Complete List of Topics and Topic Categories (*continued*)

Topic	Top FREX Words	Topic Label	Aggregated Topic Category
55	competition, market, markets, economy, prices, price, global, consumers, profits, players, competitors, value, currency, traditional, margins, giants, monopoly, goods, consumption, profitable	Competition	/
56	vat, taxation, article, taxable, decree, invoice, obligations, paragraph, directive, invoicing, obligation, discipline, purposes, profits, tax, art, entities, electronic, establishment, compliance	Taxation	Taxation
57	approval, tax_system, approved, tax, fees, tax_evasion, crime, investigation, fraud, taxes, applicant, abuse, revocation, grounds, nutritional, chapter, deductions, conduct, error, section	Tax Evasion	Taxation
58	competence, digitalisation, efforts, society, opportunities, development, increased, small, needed, possibilities, contribute, welfare, goals, county, industries, important, parts, actors, businesses, regional	Digital Competences	Digital Investments
59	income, pension, insurance, retirement, contributions, basic, taxes, social_security, dividends, universal, compensation, unemployment, paid, self_employed, wages, unconditional, pay, salary, allowance, welfare	Compensation	Automation & Compensation
60	disappears, choices, productivity_gains, unit, evolutions, salaried, dematerialization, developments, poses, progress, emancipation, deregulation, divide, stagnation, tion, intervention, transforms, struggles, isolation, forms	Other	/

References

- Blei, David M. 2012. “Probabilistic topic models.” *Communications of the ACM* 55(4): 77.
- Boukes, Mark et al. 2020. “What’s the tone? Easy doesn’t do it: Analyzing performance and agreement between off-the-shelf sentiment analysis tools.” *Communication Methods and Measures* 14(2): 83–104.
- Chang, Jonathan et al. 2009. “Reading tea leaves: How humans interpret topic models.” *Advances in Neural Information Processing System*.
- Gilardi, Fabrizio, Charles R. Shipan, and Bruno Wüest. 2020. “Policy diffusion: The issue–definition stage: American journal of political science.” *American Journal of Political Science*.

- Hu, Mingqing, and Bing Liu. 2004. "Mining and summarizing customer reviews." *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*: 168–177.
- Jockers, Matthew. 2017. "Package 'syuzhet'." <http://mirrors.nics.utk.edu/cran/web/packages/syuzhet/syuzhet.pdf>.
- Mihalcea, Rada, and Paul Tarau. 2004. "TextRank: Bringing order into texts."
- Mohammad, Saif M., and Peter D. Turney. 2010. "Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon." *CAAGET '10: Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*: 26–34.
- Moody, Christ. 2017. "Stop using word2vec." <https://multithreaded.stitchfix.com/blog/2017/10/18/stop-using-word2vec/>.
- Nicholls, Tom, and Pepper D. Culpepper. 2020. "Computational identification of media frames: Strengths, weaknesses, and opportunities." *Political Communication*: 1–23.
- Nielsen, Finn Årup. 2011. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs." http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6006/pdf/imm6006.pdf.
- Quinn, Kevin M. et al. 2010. "How to analyze political attention with minimal assumptions and costs." *American Journal of Political Science* 54(1): 209–228.
- Rinker, Tyler W. 2018. *lexicon: Lexicon data*. Buffalo, New York. <http://github.com/trinker/lexicon>.
- Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2018. "Stm: R package for structural topic models." *Journal of Statistical Software* 10(2).
- Rose, Stuart et al. 2010. "Automatic keyword extraction from individual documents." In *Text mining*, eds. Michael W. Berry and Jacob Kogan. Chichester, UK: John Wiley & Sons, Ltd, p. 1–20.
- Silge, Julia. 2017. "Tidy word vectors, take 2!" <https://juliasilge.com/blog/word-vectors-take-two/>.
- Spirling, Arthur, and Pedro L. Rodriguez. 2019. "Word embeddings what works, what doesn't, and how to tell the difference for applied research."
- van Atteveldt, Wouter, Mariken A. C. G. van der Velden, and Mark Boukes. 2021. "The validity of sentiment analysis: comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms." *Communication Methods and Measures*: 1–20.

Wijffels, Jan. 2019. "Textrank." <https://cran.r-project.org/web/packages/textrank/textrank.pdf>.

Ying, Luwei, Jacob M. Montgomery, and Brandon M. Stewart. 2019. "Inferring concepts from topics: Towards procedures for validating topics as measures."

Young, Lori, and Stuart Soroka. 2012. "Affective news: The automated coding of sentiment in political texts." *Political Communication* 29(2): 205–231.