

Problem representation on the fringe of institutions

Joan-Josep Vallbé

Institute for Information Law (IViR)
Universiteit van Amsterdam
jjvallbe@uva.edu

Abstract. Measuring the interaction between agents and institutions poses a number of challenges. The paper provides data and results from a project in political science aimed at representing the problem spaces described by organization members in depicting decision situations with high levels of procedural uncertainty.

1 Introduction

Institutions and organizations channel the behavior of people as agents. Either studied as “rules of the game” (North, 1990; Greif and Laitin, 2004; Riker, 1982; Shepsle, 1985, 2010, 1979) or as an outer environment to which agents adapt to solve problems through (boundedly) rational decision-making (March and Simon, 1958; Simon, 1996, 1947), organizations and institutions provide stable patterns of events, expectations, and guidelines to human behavior, which ultimately reduce uncertainty and enable the deployment of strategic behavior for goal fulfillment. They are, in a nutshell, stable frames of reference. Yet, these frames of reference may fail or change over time. For instance, bureaucratic organizations that rely largely on detailed routines may have difficulties in facing crises featuring unexpected behavior of agents—i.e., situations with no fixed behavioral guidelines (Vallbé, 2015). Changes in the outer environment may also be produced by agents themselves, that is, by shifts in their strategic behavior or in preferences (Przeworski, 2005, 2004; Vallbé, 2014). In addition, change or replacement of frames of reference may in turn produce changes in agents’ preferences, cognitive attachment and behavior (Magre et al., 2013; Vallbé et al., 2015).

This paper presents an empirical strategy to measure and analyze the way organization members represent situations with no fixed behavioral guidelines—i.e., when institutions fail to provide a stable pattern of events.

2 No fixed behavioral guidelines

Due to human cognitive limitations, both organizations and people develop mechanisms to cope with the regular complexity of problems (March and Simon, 1958) or to deal with critical, emergent situations (Klein and Calderwood, 1991). These mechanisms include task partitioning, departmentalization, selective perception, partial exposure, and routines.

Let's focus on routines. Simon (1997) defined them as “embodiments of ‘once and for all’ decisions”. As decisions, routines are understood as an individual's response to an environmental stimulus, but a response that “has been developed and learned at some previous time as an appropriate response for a stimulus of this class” (March and Simon, 1958). Nowadays, routines are still defined in a similar fashion: “*patterned processes* in the treatment of organizational issues regardless of whether these patterns derive from explicit policies and procedures or from implicit values and norms” (Royer and Langley, 2008).

Focusing on these *processes* rather than on the ready-made *embodiments*, Royer and Langley (2008) defend a procedural view of routines in front of a content-based or substantive one, in the vein of Simon's distinction between procedural and substantive rationality (Simon, 1978). March and Simon (1958) viewed this kind of response as one end of a continuum that marked two different kinds of response from the individual facing an organizational stimulus. The opposite end was represented by typical problem-solving activities that involve search and discovery of alternatives.

The observation of everyday situations and the accounts of actual decision making situations (Klein and Calderwood, 1991) shows that organization members usually do face situations that lie between these two extreme situations, in which cases they may search first in their “repertory of response programs” (March and Simon, 1958, 140) and then, if these programs are not adequate, they search for other alternatives, favoring a *procedural* view of routines. In fact, empirical evidence shows that a sizable proportion of organizational behavior is based on following rules rather than on calculating consequences of that behavior (Cyert et al., 1959; March and Olsen, 1998). Routine behavior, then, occurs when choice is simplified by the development of a script of a fixed response—which increases *predictability*. Routines are not only devices for solving immediate problems, but, most importantly, they are devices for learning and for turning inexperienced professionals into experts—mechanisms for making better decisions both at the individual and aggregate level (Levitt and March, 1988; March, 1990). Although normally these routines are distributed among organizational members and stored in “organizational memory”, sometimes this memory is lost (Levitt and March, 1988), contributing to uncertainty. In those cases, organizations cease to function as providers of stable patterns of events for their members, thus hampering organizational learning. When designing technology to help dealing with this problem—such as decision-support systems (Casanovas et al., 2006; Casellas, 2011)—outlining how decision-makers represent the environment of problematic issues becomes a major task.

3 The case of lower courts

Administrations of justice are typically large bureaucratic organizations filled with different kinds of professionals. In Civil Law legal systems (like most continental European countries, Spain included), these are mostly civil servants who usually develop lifelong careers within it. In Spain, Courts of First Instance and Magistrate [*Juzgados de Primera Instancia e Instrucción*], or simply lower courts, constitute the entry point into Spain's judicial system for most civil and criminal cases, and this is where where freshly appointed junior judges start their careers. Lower courts handle most civil cases

and decide on minor criminal offenses, but are also responsible for starting preliminary proceedings in any type of criminal offense. One of the main organizational principles governing the Spanish legal and judicial system is the *on-call service*. Regularly lower courts remain on call for full eight-days periods, during which the court office is responsible for handling all incoming cases reported by the police, the public prosecution or by citizens at large. When an offense takes place within the boundaries of a judicial district, the judge on call will be in charge of supervising all enquiries related to the facts of the case. The different activities the judge has to endure while on call may entail paying attention over a number of parallel issues (raised by the police, lawyers, prosecutors, citizens), and obviously the need for quick decisions seriously handicaps (or impedes) reviewing jurisprudence or precedents. At the best of times inexperienced judges have to rely on uncertain consultation with peers or senior judges, although crises may easily break in the wee hours, thus making consultation with peers quite unfeasible.

Thus when on-call, the decision-making processes are very likely to take place in a context of ambiguity, in sharp contrast with other routine and rule-based decisions that bind most legal proceedings in ordinary judicial decision-making. It is not surprising that the on-call period is perceived by most Spanish junior judges as a stressful period that continually challenges their training as judges (Ayuso et al., 2003). My general hypothesis is that the kind of problems junior judges encounter when on-call are of practical nature, to which neither junior judges nor the institution can provide systematic responses, representing the major source of uncertainty in lower-court judicial decision-making. This phenomenon represents both a professional and an organizational challenge to the whole Spanish judicial system. Professionally, the skills required to sort out the kind of problems during the on-call service would contrast sharply with the exclusively legal-theoretical education of Spanish judges. At the organizational level, the phenomenon would represent a suspension of the organization's role as a stable pattern of events, for ready-made solutions to most of these practical problems are not provided with at an organizational level (in contrast to, say, civil procedures), but judges must work them out by themselves. Hence the uncertainty.

Elsewhere (Vallbé, 2015) a full empirical test of the hypothesis is carried out. Here our interest is just to explore the nature of the problems raised during on-call situations and the role of the organizational environment in them. In particular, we explore *whether the problems faced when on call are not related to theoretical doubts but mainly practical doubts, i.e., they are identifiable demands from the outer environment whose solution is not specifically contained in the legal knowledge they acquire either in the law degree or preparing the entrance examination.*

4 Data and methods

4.1 Text as data

In order to carry out our analyses, we use three different textual corpora. These are composed by the set of responses to the open-ended questions contained in a survey carried out among Spanish junior judges. In particular, these questions were:

- A question about the main types of problems regarding *civil* issues during the first appointment. [111 responses out of 118 interviewed judges]

- A question about the main types of problems regarding *criminal* issues during the first appointment. [109 responses out of 118 interviewed judges]
- A question about the main types of problems during *on-call* periods. [110 out of 118 interviewed judges]

Our analysis involves exploring the content of these corpora in a systematic way, treating text as data. We assume that the significant differences among documents can be reduced to differences in the use of language in the documents. Specifically, if two documents refer to different topics (e.g. they refer to different kinds of problems), these differences can appear in both the types of words they contain, and the frequency of these words. We apply statistical methods to these textual data in order to account for these differences in meaningful ways. We adopt the Bag of Words Model, that represents documents as vectors in a common vector space, with numerous applications so far (Salton et al., 1975; Baeza-Yates and Ribeiro-Neto, 1999; Jakulin and Buntine, 2004). This way, each document d can then be represented as a vector \vec{d} of m weighted terms (Huang, 2008). Content similarity between two documents is understood as the two document vectors having similar vector representations (Manning et al., 2008).¹

To test our hypothesis, we carry out two different analyses. First, we need to retrieve the issues and problems that judges refer to in their responses. Second, we need to explore whether the kinds of problems they identify as specific of on-call situations are of a different kind than those raised during ordinary civil and criminal procedures. We need to *represent* them.

4.2 Probabilistic topic models

In order to extract the issues referred to by judges in their responses in a systematic way, we use Latent Dirichlet Allocation (LDA), a generative probabilistic topic model for the identification and representation of the content of collections of documents (Blei et al., 2003; Blei, 2012). Once topics have been identified and interpreted, they will be used as outcomes to test the practical nature of the problems raised during on-call situations.

Topic models address the need to explore the semantic structure underlying a collection of documents by finding semantically similar documents and “exploring the collection through the underlying topics that run through it” (Blei and Lafferty, 2009). Latent Dirichlet Allocation (LDA) is a *generative* Bayesian mixture model for discrete data (Grün and Hornik, 2011).² The idea behind LDA is that documents have a certain number of topics, where a topic is conceptualized as a probability distribution over a set of terms (a vocabulary V). In particular, LDA is based on the assumption that a collection of documents is associated to a number of topics, and that each document will present a certain proportion of these topics. Then each document can be seen as representing a particular combination of those topics. Since the only observed data are words

¹ For each corpus, we select the terms that have a mean term-frequency/inverse-document-frequency (tf-idf) score (Manning et al., 2008) score over the documents just over the median (Grün and Hornik, 2011), keeping only the relevant terms.

² The following explanation on LDA draws mainly on Blei and Lafferty (2009); Blei (2012); Buntine and Jakulin (2006); Buntine (2009); Mimno (2012).

within documents, LDA is usually formally described as a *hidden variable model* (Blei and Lafferty, 2009), in which the topic structure of the document collection and of each document is unknown and must be discovered using posterior probabilistic inference given the observed documents. The main goal of LDA, then, is to reverse the process depicted in the data generation model: given the set of observed terms in our collection of documents, we infer the underlying topic structure (Blei, 2012) through the posterior distribution of the latent variables, as if the data were generated by the model. This way we obtain probabilities at the different levels of analysis: topic distribution for documents, and term distribution for each topic. In this particular case, the evaluation of the performance (i.e. model perplexity (Blei et al., 2003)) of successive LDA models including from 2 to 100 topics recommended setting the number of topics to 10.³

5 Results

Figure 1 presents the distribution of topics and their probabilities per document for a sample of 20 documents. We see that the model favors the assignment of significant probabilities of each document in one or just a small number of topics. Notice that while some documents have a maximum of 3 topics with significant proportions, (e.g., document # 12), most of them have just one topic associated to a high probability.

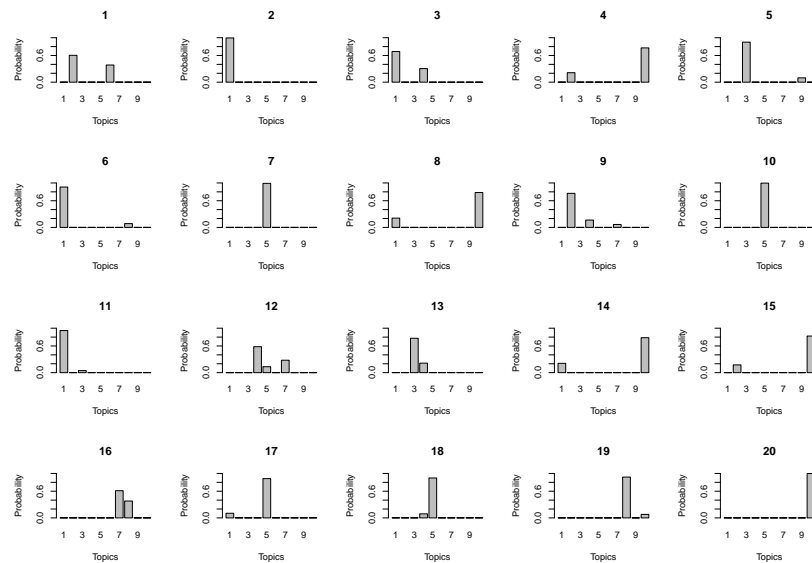


Fig. 1: Distribution of topics per document for a sample of 20 documents in the collection.

³ The decision is based on the marginal decrease of perplexity in the successive n -topic models. See details in Vallb  (2015).

5.1 Topics and their meaning

Let's now turn to the contents of the topics. It is worth noting here that our first interest is to identify different domains within the whole set of problems pointed out by judges. Table 1 lists the twenty more relevant terms of each of the 10 topics generated by our topic model. There are terms that appear in more than one topic, although the probability these terms are assigned to each topic by the model differ in each case. Therefore, both the presence and the order in which terms appear ranked within topics are relevant to their interpretation.

Table 1: List of the most relevant terms in each topic, with their assigned probability.

Topic 1		Topic 2		Topic 3		Topic 4		Topic 5	
term	<i>p</i>	term	<i>p</i>	term	<i>p</i>	term	<i>p</i>	term	<i>p</i>
1 médico	0.07	tema	0.04	duda	0.05	fiscal	0.08	asunto	0.02
2 fiscal	0.04	derecho	0.03	fiscal	0.03	delito	0.04	año	0.02
3 violencia	0.04	complicado	0.02	levantamiento	0.02	instrucción	0.03	fuerza	0.01
4 doméstico	0.03	servidumbre	0.02	violencia	0.01	causa	0.02	tema	0.01
5 internamiento	0.03	asunto	0.02	prisión	0.01	sumario	0.01	propiedad	0.01
6 protección	0.02	escuela	0.01	ley	0.01	procedimiento	0.01	sentencia	0.01
7 tema	0.02	sentencia	0.01	medida	0.01	responsabilidad	0.01	cuestión	0.01
8 urgente	0.02	tierra	0.01	social	0.01	prisión	0.01	letrado	0.01
9 nunca	0.02	real	0.01	dónde	0.01	complicado	0.01	duda	0.01
10 asunto	0.01	estudiar	0.01	dificultad	0.01	droga	0.01	quiebra	0.01
Topic 6		Topic 7		Topic 8		Topic 9		Topic 10	
term	<i>p</i>	term	<i>p</i>	term	<i>p</i>	term	<i>p</i>	term	<i>p</i>
1 audiencia	0.03	tema	0.03	ejecución	0.06	prisión	0.05	tío	0.04
2 tema	0.03	perito	0.03	ley	0.03	ley	0.02	tema	0.03
3 fiscal	0.02	asunto	0.02	sentencia	0.02	médico	0.02	prisión	0.02
4 sentencia	0.02	niño	0.02	procedimiento	0.02	cuestión	0.02	duda	0.02
5 procedimiento	0.02	realmente	0.02	embargo	0.02	año	0.02	medida	0.01
6 nacional	0.01	protección	0.02	pleito	0.02	causa	0.02	fiscal	0.01
7 libertad	0.01	medida	0.02	resolución	0.02	fiscal	0.02	niño	0.01
8 prisión	0.01	malos	0.02	pericial	0.02	violencia	0.01	acordar	0.01
9 asunto	0.01	tratos	0.02	año	0.01	protección	0.01	asunto	0.01
10 escuela	0.01	alejamiento	0.01	antiguo	0.01	asunto	0.01	concreto	0.01

Of these sets of terms, roughly three different groups of topics can be identified based upon their content. The first set of topics (2, 5 and 8) is composed mainly by legal technical terms of both civil and criminal procedures. On the other hand, topics 1, 3, 7, 9 and 10 refer to behavioral problems. Topics 4 and 6 are hard to classify in either of the previous two sets.

Technical topics If we take a look at the terms with highest probability to belong on topics 2, 5, and 8 we shall observe that most of them refer to technical, theoretical legal

typologies. For instance, the highly ranked presence of terms such as *servidumbre* [right of way], *obra* [construction], and *derecho* and *real* (for *derechos reales*) [tax on the conveyance of land or property, *jus in rem*] indicates that Topic 2 gathers terms belonging to problems on technical issues of civil procedures, especially regarding ancient civil figure such as the *jus in rem*.

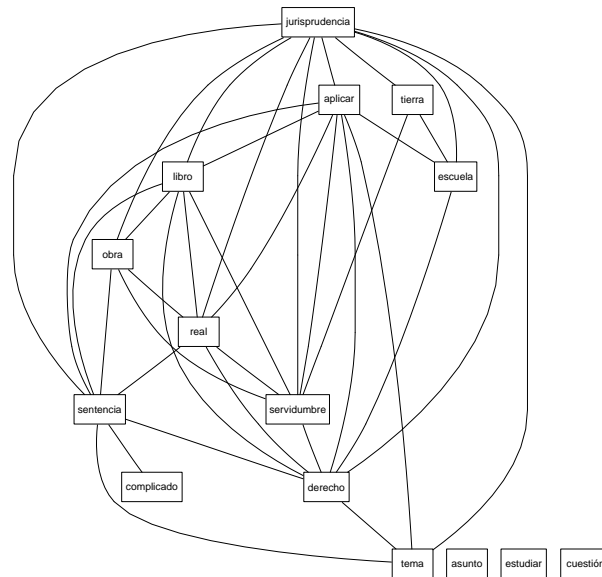


Fig. 2: Representation of the correlation network of the 15 most relevant terms of Topic 2.

A better assessment of these results is achieved when we plot the correlation between the terms within each topic, as in Figure 2.⁴ These plots reflect the way judges express their main problems, so the presence or absence of particular terms, on one hand, and the association among terms that do significantly appear in a topic, on the other, may be interpreted as a quite reliable representation of this particular portion of discourse related to their problems *as they describe them*. At the same time, terms (nodes) with a higher degree in the network (number of edges) may be interpreted as gravitational centers of the topic from which a data-driven interpretation of the results of the topic models can be made. This representation will be able to stress out, first, those aspects of particular situations which seem problematic by judges, and second, whether these situations are related to doubts regarding particular actors and their ac-

⁴ Due to the sparsity of the original matrix, we use a threshold of $r = .25$ in order to increase the density of the correlation networks.

tions (behavioral interpretation) or to doubts regarding theoretical interpretation. Just as an example of *technical* topic, in Figure 2 we see that technical terms such as *servidumbre* [right of way] and *derecho* and *real* have a prominent position in the correlation network. Moreover, the term *sentencia* [sentence, judicial decision] is correlated to this civil vocabulary, to other procedural legal terms such as *procedimiento* [procedure] and *ejecución* [enforcement of a judicial decision], and also to the key term *jurisprudencia* [precedent]. Problems with the enforcement of judicial decisions, in fact, seem to be the central point of all topics relating to civil procedures, although the topic model seems to separate different kinds of procedures.

Behavioral topics Topics 1, 3, 7, 9 and 10 seem to belong to an entirely different semantic context than the ones just interpreted, mainly because the terms and their pattern of relationships refer to behavioral problems, usually raised during on-call situations. For instance, Topic 1 contains terms that relate to three different but typical on-call situations. On one hand, terms such as *internamiento* [confinement] and *urgente* [urgent, urgently] do point to those cases involving the request of confining mentally disturbed people into mental institutions, a usual critical situation dealt with by the on-call service in lower courts. On the other hand, terms such as *muerte* [death] and *levantamiento* [removal, for corpse removal] point to on-call situations related to the death of a person and the subsequent removal of the corpse. Finally, words such as *violencia* and *doméstico*, and *protección* (as in *medida de protección*, injunction of protection) refer to cases of domestic violence (*violencia doméstica*).

Figure 3 shows the correlation between the first 15 terms in Topic 1 in the DocumentTerm matrix used to generate the topic model. If we observe it as an undirected network of terms (correlation goes both ways), it reveals that, although both the ME and the public prosecutor are in principle relevant to the topic (they both have high probabilities attached), the medical examiner has a relevant role as a node in the description of the on-call problems made by judges. As expected, other terms that present a high number of edges in the correlation network are *violencia*, *doméstico*, *protección*, *internamiento*, *muerte* and *prisión*.

Hard cases Topics 4 and 6 are difficult to interpret. These topics seem to be devoted to legal procedural issues, specifically on difficulties encountered during the preliminary proceedings (*instrucción del sumario*) that junior judges must initiate in their courts. These difficulties are most related to the constitution of a jury (*formación del jurado*), and lawsuits (*causa*), involving drugs (*droga*), fraud (*estafa*), and issues over liability (*responsabilidad*).

5.2 The behavioral/technical divide

The figures above look like incomplete maps of which we only see some relevant milestones, terms that may be seen as representing the main features of junior judges' problem spaces regarding a number of situations of their professional activity (be it the removal of a corpse or starting preliminary proceedings for fraud). Moreover, the results from the topic model have helped us classify the topics in two broad categories:

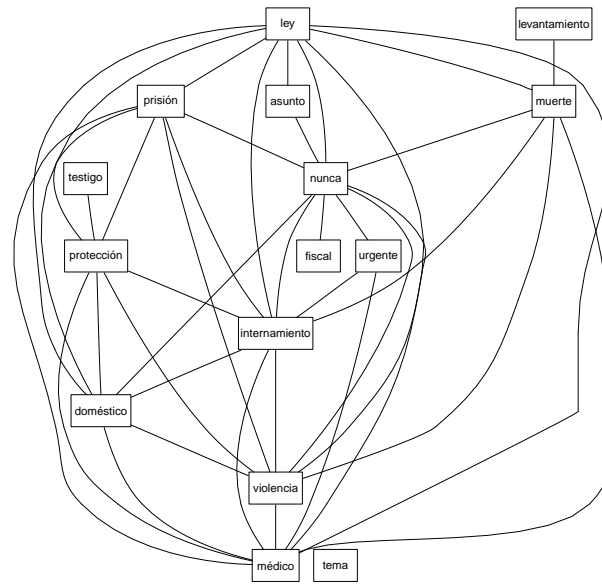


Fig. 3: Correlations among terms within Topic 1, which refer to three different kinds of problems during on-call situations.

some topics are clearly legal technical, others refer more to behavioral situations. A few seem to lie in between, in a neutral category. The theoretical classification between technical and behavioral topics receives support in the data. First, results of performing a Principal Components Analysis on the posterior probability attached to each document within each topic yield a reasonable classification of topics according to our theoretical expectations. Figure 4 represents the classification of the topics on the component that explains most of the variation in the data.

On the other hand, we fit an unordered multinomial logit model of the Technical-Neutral-Behavioral classification on the type of problem depicted by junior judges, with the Technical category as reference in both models. Our predictor for the type of problem is whether junior judges are explaining problems encountered during civil procedures, criminal procedures, or on-call situations. Table 2 shows the predicted probabilities of each type of document to belong to each category. The difference between civil and on-call problems is sharp regardless the classification model. In particular, problems on civil procedures have a 80% chance to be of technical nature, while on-call problems present a change between 70 and 80% of being behavioral, and only a 7% chance of being technical.

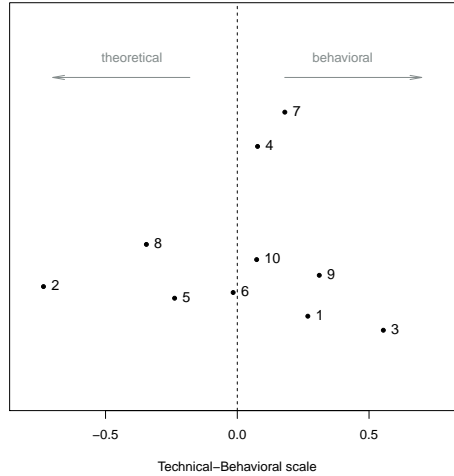


Fig. 4: Classification of the topics along the first component of the Principal Components Analysis on the posterior probability of each document on each topic.

Table 2: Predicted probabilities for the documents within each type of corpus of being technical, neutral, and behavioral, according to our theoretical model.

	<i>Classification model</i>		
	Technical	Neutral	Behavioral
Civil	0.81	0.04	0.15
Criminal	0.12	0.31	0.57
On call	0.07	0.12	0.81

6 Conclusions

Legal and other bureaucratic organizations prove inefficient in storing and routinize informal knowledge or knowledge produced while managing improvisation-prone situations. The work presented in this paper provides a framework for the analysis and representation of the interaction between agents (decision-makers) and their institutional environment through problem spaces. In addition, the methodology presented can help in the identification of knowledge-intensive areas when outlining the elements that constitute decision-makers' problem spaces in several decision situations. In particular, our use of various text analysis techniques could also help knowledge acquisition processes for the construction of knowledge-based systems aimed at improving decision-making processes.

Bibliography

- Ayuso, M., Bécue-Bertaut, M., Álvarez, R., Valencia, O., Álvarez, M., and Santolino, M. (2003). Análisis estadístico de las encuestas a los jueces en su primer destino. Internal Report SEC-2001-2581-C02-01/02, Consejo General del Poder Judicial [General Council of the Judiciary], Escuela Judicial (Barcelona).
- Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. ACM Press, New York.
- Blei, D. (2012). Probabilistic topic models. *Communications of the ACM*, 55:77–84.
- Blei, D. and Lafferty, J. D. (2009). Topic models. In Srivastava, A. N. and Sahami, M., editors, *Text Mining. Classification, Clustering, and Applications*, pages 71–94. CRC Press, Boca Raton, FL.
- Blei, D., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022.
- Buntine, W. (2009). Estimating likelihoods for topic models. NICTA Technical Report NICTA-SML-09-001, NICTA, Canberra (AU).
- Buntine, W. and Jakulin, A. (2006). Discrete component analysis. In *Subspace, Latent Structure and Feature Selection Techniques*. Springer-Verlag.
- Casanovas, P., Casellas, N., Vallbé, J.-J., Poblet, M., Benjamins, V. R., Blázquez, M., Peña-Ortiz, R., and Contreras, J. (2006). Semantic Web: A Legal Case Study. In Davies, J., Studer, R., and Warren, P., editors, *Semantic Web Technologies: Trends and Research in Ontology-based Systems*, chapter 12, pages 259–280. John Wiley & Sons, Chichester: UK.
- Casellas, N. (2011). *Legal Ontology Engineering. Methodologies, modelling trends, and the Ontology of Professional Judicial Knowledge*. Springer.
- Cyert, R. M., Feigenbaum, E. A., and March, J. G. (1959). Models in a behavioral theory of the firm. *Behavioral Science*, 4(2):81–95.
- Greif, A. and Laitin, D. D. (2004). A theory of endogenous institutional change. *American Political Science Review*, 98(4):633–652.
- Grün, B. and Hornik, K. (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, 40(13):1–30.
- Huang, A. (2008). Similarity measures for text document clustering. In Holland, J., Nicholas, A., and Brignoli, D., editors, *New Zealand Computer Science Research Student Conference (NZCSRSC)*, Christchurch (NZ). NZCSRSC.
- Jakulin, A. and Buntine, W. (2004). Applying discrete pca in data analysis. In Meek, C. and Halpern, J., editors, *Proceedings of the Twentieth Conference Conference on Uncertainty in Artificial Intelligence*, Arlington, Virginia. AUAI Press.
- Klein, G. A. and Calderwood, R. (1991). Decision models: Some lessons from the field. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(5):1018–1026.
- Levitt, B. and March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14:319–340.
- Magre, J., Vallbé, J.-J., and Tomàs, M. (2013). Moving to suburbia? Effects of residential mobility on community engagement. In *2013 Conference of the European Political Science Association*, Barcelona.

- Manning, C. D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press, Cambridge (UK).
- March, J. G. (1990). *Decisions and Organizations*. Basil Blackwell, Oxford, UK; Cambridge, Ma.
- March, J. G. and Olsen, J. P. (1998). The institutional dynamics of international political orders. *International Organization*, 52(4):943–969.
- March, J. G. and Simon, H. A. (1958). *Organizations*. John Wiley, New York, NY.
- Mimno, D. (2012). Computational historiography: Data mining in a century of classics journals. *Journal on Computing and Cultural Heritage*, 5(3):1.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press, New York, NY.
- Przeworski, A. (2004). Institutions matter? *Government and Opposition*, 39(2):527–540.
- Przeworski, A. (2005). Democracy as equilibrium. *Public Choice*, 123:253–273.
- Riker, W. H. (1982). The two-party system and Duverger's law: An essay on the history of political science. *The American Political Science Review*, 76(4):753–766.
- Royer, I. and Langley, A. (2008). Linking rationality, politics, and routines in organizational decision making. In Hodgkinson, G. P. and Starbuck, W. H., editors, *The Oxford Handbook of Organizational Decision Making*, pages 250–270. Oxford University Press, New York, NY.
- Salton, G., Wong, A., and Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the Association for Computing Machinery*, 18(11):613–620.
- Shepsle, K. (1985). Prospects for formal models of legislatures. *Legislative Studies Quarterly*, 10(1):5–19.
- Shepsle, K. A. (1979). Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science*, 23(1):27–59.
- Shepsle, K. A. (2010). *Analyzing Politics: Rationality, Behavior, and Institutions*. W & W Norton & Company, New York, 2nd edition.
- Simon, H. A. (1947). *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*. Macmillan, New York, NY, 1st edition.
- Simon, H. A. (1978). On how to decide what to do. *The Bell Journal of Economics*, 9(2):494–507.
- Simon, H. A. (1996). *The Sciences of the Artificial*. The MIT Press, Cambridge (MA), 3rd edition.
- Simon, H. A. (1997). Allen Newell, 1927-1992. *National Academy of Sciences Biographical Memoirs*, 71.
- Vallbé, J.-J. (2014). Measuring the judicial power of regions: A judicial regional authority index. In *ECPR 2014 Joint Sessions*, Salamanca (Spain). ECPR.
- Vallbé, J.-J. (2015). *Frameworks for Modeling Cognition and Decisions in Institutional Environments. A Data-driven Approach*. Number 21 in Law, Governance and Technology. Springer.
- Vallbé, J.-J., Magre, J., and Tomàs, M. (2015). Being metropolitan: the effects of individual and contextual factors on shaping metropolitan identity. *Journal of Urban Affairs*.