Tracking Hot Topics for the
Monitoring of Open-world Processes

Remo Pareschi[1], Marco Rossetti[2], Fabio Stella[2]

[1] Department of Bioscience and Territory, University of Molise, Pesche (IS), Italy
`remo.pareschi@unimol.it`
[2] Department of Informatics, Systems and Communication, University of Milano-Bicocca, Milan, Italy
`{rossetti,stella}@disco.unimib.it`

Abstract. We introduce the notion of open-world process to refer to processes that generally require flexible execution and are influenced by external factors. Among such processes we focus on those that fit also with the notion of "business process". We then introduce "hot topics" to capture contextual information flows that, by flowing into the context of execution of open-world business processes, may affect their definitions. Hot topics are discovered using unsupervised learning techniques based on Probabilistic Topic Modeling and by tracking variations of the information flows into topics over time. We illustrate an application of this methodology to the tracking of recent innovations in labor laws that affect a variety of open-world business processes, from labor sourcing to merge-and-acquisition. Finally, we define a number of future directions for research.

Keywords: open-world, business process management, probabilistic topic modeling, process monitoring

1   Introducing open-world processes

When talking about processes in computer science and in artificial intelligence we generally refer to computational objects that, once defined, have clear-cut and predictable behaviors. The background from which such processes arise may be itself characterized by formal clarity, as is the case of the goals of a robot that through the verification offered by logical reasoning are synthesized into a process corresponding to an executable plan. Conversely, it may be of an initially murkier kind, as is the case of sequences of logs of actions or tasks performed by people with diverse roles within an organization, which can be mined and combined into explicitly defined work processes. However, in either case, the result is an object with fixed formal and computational properties, corresponding to a repeatable sequence of actions that may allow diverse degrees of flexibility, which are however known and planned for in advance.

Compare this situation with the following quote from the book War, a best-selling in-depth report on the current Afghan war (written by the journalist Sebastian Junger, who witnessed first-hand all the dramatic episodes described there): "the war also

diverged from the textbooks because it was fought in such axle-breaking, helicopter-crashing, spirit-killing, mind-bending terrain that few military plans survive intact even for an hour." Without reaching such extremes, even in far more peaceful and less tragic states of affairs, there are lots of work and business processes that diverge substantially upon execution from the way they were planned, and that consequently redefine themselves dynamically according to circumstances. Such dynamic redefinitions may, more often than not, be quite radical, and hence go beyond the boundary of mere adaptation.

Clearly, these processes include all those that require real-time interactions between the agents involved and are heavily entangled with the physical world, for instance, in addition to military activities, those relating to geographical and geological exploration, logistics, energy planning. However, they include also processes where the role of the physical world is less immediate and the interaction among the participating agents is more asynchronous. This is the case with the execution of different aspects of corporate strategy (like market expansion, go-to-market plans, technology transfer, protection of intellectual property, merge-and-acquisition), the running of electoral and advertisement campaigns and financial placements.

We shall refer to processes of this kind as "open-world" by borrowing a terminology that was used in [1] to refer to the evolution of software development from a "closed-world" process to an "open-world" one. A more general way to look at these processes comes from observing that they are subject to flexible execution and are strongly influenced by external factors [2]. It should be noticed that in no way we are implying that "open-world" processes are to be considered better, or even just more relevant, than "closed-world" ones. As a matter of fact, closed-world processes are certainly easier to deal with both from the point of view of organizational models and of methods of computer support. On the other hand, open-world processes do exist. Furthermore, in a time when the traditional boundaries between organizations, institutions and countries are getting more and more friable, they are likely to increase. Hence there is room and need to increase also the support that can be derived for them from information technology.

We take the following properties to be characteristic of an open-world process:

1. First of all, it is, indeed, a process: namely it is defined as a number of steps that can be executed in a sequence in order to achieve a certain type of objective, with possible choice points subject to pre-conditions determined by its context of execution; taking a business-oriented definition, it can also be viewed as "a set of linked activities that take an input and transform it to create an output. Ideally, the transformation that occurs in the process should add value to the input and create an output that is more useful and effective to the recipient either upstream or downstream." [3].
2. However, the process definition is open to the possibility of continuous revision, refinement, and re-interpretation due to the interaction with external agents during its execution; similarly, roles and resources in the process may need to be revised, for instance in consequence of the encounter of hitherto unknown stakeholders, or

of the impossibility to access resources that turn out to be unavailable at execution time.

3. Open-world processes of this kind are most often mission-critical for their originating organizations or institutions, that thus generally create explicit or implicit roles for decision-makers of last resort, who have the powers to redefine the process and reset roles and resources; hence the need of identifying clearly such decision makers and of extracting the decisions that they have made in order to get a grasp on the status of definition of the process and on how much it has eventually stabilized.

Point 3 defines the scope of this article, and of the future developments that can follow from it.

Our approach hinges on the tracking of increased information flows, what we call hot topics, around the execution of open-world processes, thus detecting situations where the current process execution is stressed by external factors, and an evolutionary change of its definition is therefore likely to take place. Hot topic tracking derives from the application of Probabilistic Topic Modeling (PTM) [4], a framework for the unsupervised learning of topics in collections of textual contents that has proved robust and effective in a variety of contexts. We shall illustrate an application of our approach to a specific type of external factors. In fact, we shall show how the implementation of legislation, and in particular of labor law, may affect a number of business processes where employees are among the key stakeholders, either directly, as in the case of hiring and dismissal, or indirectly, as in the case of business transfers.

The remainder of this paper is therefore structured as follows:; in Section 2 we illustrate the general principles underlying our application of PTM to the tracking of hot topics; Section 3 is the core of the paper, where we apply hot topic tracking to the monitoring of the evolution of open-world processes in the context of implementation of labor law, a very dynamic and socially "hot" sector of civil law; Section 4 outlines directions for future work and concludes the article.

2   Topic Modeling for Hot Topic Tracking

Text mining approaches based on Probabilistic Topic Modeling (PTM) are recently gaining considerable value as they allow the discovery of high-level dependencies between contents of a document corpus. Probabilistic Topic Models are statistical methods capable to handle, through unsupervised learning techniques, large volumes of unstructured data. The main purpose of these algorithms is the analysis of words in natural language texts in order to discover themes represented by sorted lists of words. For instance, Figure 1 shows 4 out of 300 topics extracted from the TASA corpus [5]. It is easy to see that words in the four topics are related to each other and can be considered as consistent *themes*. Furthermore, PTM is also able to provide topic proportions for each document, which is very useful to understand which themes a document is about. PTM-based text mining approaches aim to get the best of both worlds, by providing richly structured representations of the knowledge derived from the empirical validation of "Big Data" processing. Hence they improve both on tradi-

tional symbolic approaches, that lack data validation, and on connectionist approaches, that lack capability to represent knowledge [6].

Topic 247			Topic 5			Topic 43			Topic 56	
word	prob.		word	prob.		word	prob.		word	prob.
DRUGS	.069		RED	.202		MIND	.081		DOCTOR	.074
DRUG	.060		BLUE	.099		THOUGHT	.066		DR.	.063
MEDICINE	.027		GREEN	.096		REMEMBER	.064		PATIENT	.061
EFFECTS	.026		YELLOW	.073		MEMORY	.037		HOSPITAL	.049
BODY	.023		WHITE	.048		THINKING	.030		CARE	.046
MEDICINES	.019		COLOR	.048		PROFESSOR	.028		MEDICAL	.042
PAIN	.016		BRIGHT	.030		FELT	.025		NURSE	.031
PERSON	.016		COLORS	.029		REMEMBERED	.022		PATIENTS	.029
MARIJUANA	.014		ORANGE	.027		THOUGHTS	.020		DOCTORS	.028
LABEL	.012		BROWN	.027		FORGOTTEN	.020		HEALTH	.025
ALCOHOL	.012		PINK	.017		MOMENT	.020		MEDICINE	.017
DANGEROUS	.011		LOOK	.017		THINK	.019		NURSING	.017
ABUSE	.009		BLACK	.016		THING	.016		DENTAL	.015
EFFECT	.009		PURPLE	.015		WONDER	.014		NURSES	.013
KNOWN	.008		CROSS	.011		FORGET	.012		PHYSICIAN	.012
PILLS	.008		COLORED	.009		RECALL	.012		HOSPITALS	.011

Fig. 1. Example of topics extracted from the TASA corpus [5].

More technically, PTM, exploits LDA (Latent Dirichlet Allocation) [7] to automatically extract topics (concepts) from document corpora. Each topic is associated with a list of words and each document is associated with a mixture of topics. The process of topic extraction returns the probability distribution ϕ^j of the words of the document corpora for each topic j and the probability distribution θ^i of the topics for each document i. By exploiting techniques of statistical inference and sampling, these probability distributions are inferred by observing the frequency of words within documents. Figure 2 illustrates the probabilistic generative process and the statistical inference process for topic extraction.

The idea behind the use of PTM for monitoring open-world processes is to exploit the flow of information that accompanies the different steps of the process to identify situations where topics that are most closely associated with contents exchanged during process execution become densely populated. We take this as a signal that the process is going through a critical phase and that the information that accumulates around it during that period of time can contain key indicators of possible changes and re-adaptations.

To give an example on which we will return with further details later on, an intercompany process of the highest relevance is the one that governs the transfer of a business from one ownership to another. Clearly, this is an open-world business process in the sense meant in Section 1, in fact:

1. It is a set of linked activities that take an input and transform it to create an output. In this case the input is the existing ownership of a business or of a business unit, and the output is a new ownership, with a downstream recipient corresponding to the new owner, to whom the business is transferred, and an upstream recipient corresponding to the former owner, who gets the proceeds of the transaction;

4

2. It is open-world, since its control does not reside within the boundary of a single organization and its execution is contextually influenced by a number of factors and stakeholders which may vary and change over time.
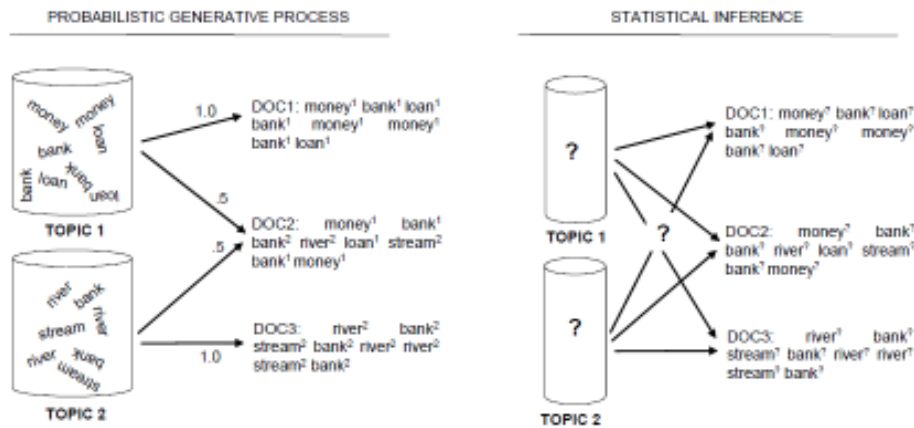

Fig. 2. From the generative process to the statistical inference [5].

For companies that operate in countries that adhere to the European Union such contextual factors and stakeholders include, respectively, the "Transfers of Undertakings Directive" of the European Union, that protects the contracts of employment of people working in transferred businesses, and the decision-makers with powers on the interpretation and application of this law. Each member state is responsible for the implementation of the Directive with respect to transfer operations of its pertinence and the rules of implementation have been generally clarified through the intervention of decision-makers of last resort, who have arbitrated conflicting interpretations during the run-in period and thus have defined the standards of implementation from then on. Considering the time generally required for a European law to stabilize within the legislative framework of the member states, and the fact that this legislation replaces since the beginning of the 2000s a previous European law in force during the last 25 years of the previous century, we can say that the Directive is reaching just now the final stage of its running-in. As we shall see, this situation is clearly reflected by the high heat of the topic most directly associated with the Directive in relatively recent times, when the Italian Court of Cassation, which is the decision-maker of last resort on this subject in Italy, made a number of judgments that interpret and define the criteria for its implementation. These documents are clustered within the topic during that interval of time, causing its heating up. After that the process of business transfer in Italy appears, as regards the Directive of Undertakings, to have reached a stable state, and this is reflected in the corresponding cooling down of the topic, as shown by the fact that in the successive time intervals the amount of content clustered within it drops very sharply.

Thus, spotting when a topic becomes "hot" and, conversely, when it "cools down", requires plotting the content carried by the topic through time. This will be illustrated in the next section.

3   Methodology

In order to substantiate the relationship between open-world business processes and hot topics we have employed LDA to classify 20.600 rulings issued by the Italian Court of Cassation in matters pertaining labor law between 2009 and 2014. This period saw several innovations of the Italian labor law, some of which attributable to the implementation of European directives in the field. In the Italian civil law system the Court of Cassation is called upon to address and validate the work of the lower courts as well as to fix the interpretation of the legislation. We can therefore expect that processes, which typically involve businesses, trade unions and workers as stakeholders, made possible by these innovations have gone through a period of adjustment solved through the deliberating activity of the Court of Cassation; such activity can in turn be reconstructed by tracking hot topics within the corpus. To achieve this, we need to plot the evolution of the measured probability of each topic against time. Let us define t as the time frame considered, $p(z|j)$ as the probability of topic z given the judgment j, $p(j|t)$ as the probability of judgment j given the time frame t and T_j as a function that associates the judgment j with the corresponding time frame. The empirical probability that an arbitrary judgment j issued at time period t was about topic z is indicated with $p(z|t)$ and it is defined in Equation 1:

$$p(z|t) = \sum_{j:T_j=t} p(z|j)p(j|t) = \frac{1}{C} \sum_{j:T_j=t} p(z|j) \qquad (1)$$

Since $p(j|t)$ is the probability that the judgment is assigned to the time frame t, that term can be substituted with $\frac{1}{C}$, where C is the number of judgments in the time frame t. The function T can be parameterized to yield time intervals corresponding to one month, two months, four months, six months and one year periods.

We ran LDA setting the number of topics to extract equal to 10, 20, 50, 100 and 200 in order to find the best granularity of topics. The standard LDA model assumes that the topic structure is flat, and tries to assign a unique theme to each topic. However, if the number of topics is too big themes are split across many topics, while if the number is too little more themes can be aggregated in a single topic. While selecting the correct number of topics is still an open issue [8], an empirical analysis can be conducted to assess the topic quality. In our case domain experts, namely labor lawyers, chose the 50 topics experiment as the best candidate and specifically reviewed and graded the 50 topics set. In all 18 topics turned out to be good performers, 14 were considered noise with remaining ones being somewhat uncertain. Of the 18 good performers a further selection can be made by taking out 4 topics that are so close to other ones to correspond substantially to clones. Best performing topics have been manually labelled on the basis of the domain specific sense of topic words (Table 1).

The characteristics of a good performer, in the eyes of domain experts, can be summarily characterized in the ability to identify concepts specifically attributable to a particular legislative and / or decision-making context, e.g. "collective dismissal /

union agreement / selection criteria" or "rotation / redundancy funds / Fiat agreement".

Table 1. Best performing topics from the 50 topics extracted.

Best performing topics	Relevant words
Transfer of business	judge (giudice), convinction (convincimento), evidence (prova), irregularity (irregolarità), company (azienda)
Collective contract	collective (collettivo), contract (contratto), agreement (accordo)
Collective dismissal	employees (dipendenti), union (sindacali), criteria (criteri), mobility (mobilità), collective (collettivo)
Work injury	injury (infortunio), liability (responsabilità), damage (danno), insurance (assicurazione)
Dismissal for just cause	contestation (contestazione), sanction (sanzione), just_cause (giusta causa), conduct (condotta), justified (giustificato)
Overtime work	compensatory_rest (riposo compensativo), damage (danno), availability (reperibilità)
Journalistic job provision	provision (prestazione), activities (attività), journalists (giornalisti), nature (natura), guarantee (garanzia)
Nature of the enterprise	cooperative (cooperativa), family (familiare), tax (tributario), administration (gestione), protection (tutela), shareholder (socio)
Fixed-term employment contracts at the Italian Post	contract (contratto), fixed-term (termine), Italian Post (Poste Italiane), damage (danno)
Impact on severance indemnities of overtime work	overtime_work (lavoro_straordinario), indemnities (trattamento), compensation (compenso), national_collective_labor_contract (CCNL)
Notice and indemnity in the agency contracts	contract (contratto), agent (agente), indemnity (indennità), notice (preavviso)
Criteria of rotation in the extraordinary wages guarantee fund	extraodinary_wages_guarantee_fund (CIGS), criteria (criteri), rotation (rotazione), agreement (accordo), Fiat
European directive on transfer of undertakings	transferee (cessionario), European (europea), court_of_justice (corte di giustizia), directive (direttiva), transfer (trasferimento), seniority (anzianità)
Duties and qualifications of company directors	national_collective_labor_contract (CCNL), qualifications (mansioni), category (categoria), speriore (higher), director (dirigente)

Conversely, the characteristics of a bad performers may be multiple, some of which amendable through an improved morpho-syntactic analysis of the text, but the most frequent and endemic one resides in a composition of the topic in terms of elements too general to lead to significant identification, such as "law", "burden", "responsibil-

ity" and so on. We have then applied Formula [1] to monitor the trends in the topics. Topic Dismissal for just cause (the fifth from the top of Table 1) makes for an interesting and a relevant case. The theme of the topic has been in fact substantially revised by the most recent labor reform in Italy, which entered in force in June 2012, among other things by introducing relevant modifications in the open-world process related to the retaining of workers by businesses, in particular regarding so called small and medium enterprises (SMEs). We can therefore expect that immediately after that the topic would heat up. This could not be related much to the possibility to open and finalize new legal procedures on the basis of the recent legislation, which would not be possible in such a short period of time, but rather on the ability to make decisions on extant procedures by also taking into account the new norms. So it turns out to be the case: the graphic in Figure 3 shows a peak in the topic trend (probability evolution) during the second half of 2012, that on a bimonthly split can be exactly located in September 2012. After this peaking the topic progressively cools down, an indicator that the corresponding process has for the time being readjusted and stabilized.
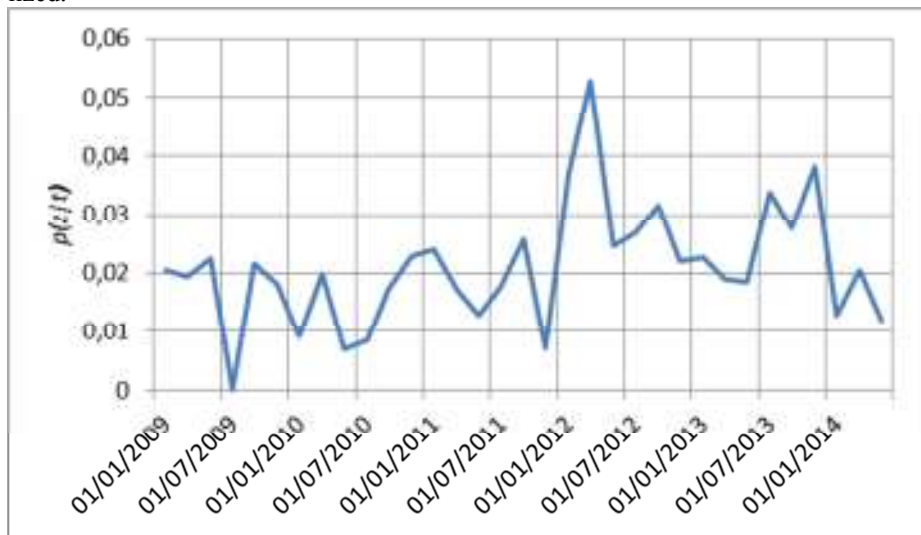

Fig. 3. "Dismissal for just cause" topic evolution.

As representative of the typical development of a hot topic, and an associated open world process, as Dismissal for just cause is, it is far from being the hottest topic among those that we have identified. In fact at the top of the "hit parade" of hot topics we find the second last item from Table 1, namely European directive on undertakings. We can compare how far hotter it is with respect to Dismissal for just cause by plotting the trends of the two topics one against the other as in the graphic in Figure 4, where we can also notice that the latter topic peaks up at its highest probability value during the second half of 2011 and then resurges sharply yet again, even if without reaching the previous heights, for a longer period encompassing most of the second half of 2012 and of the first half of 2013. Indeed, it is in this period that the Italian

Court of Cassation issues a number of rulings that have become fundamental benchmarks for the implementation in the Italian context of the European directive on transfer of a business (or of a business unit). We can also ask ourselves why European directive on undertakings is so much hotter as a topic than Dismissal for just cause. This may find a plausible answer in the fact that the scope of European directive on undertakings, that concerns companies of all sizes, and touches an issue of foremost importance (sometimes decisive for the fate of thousands of workers), is so much wider than the changes affecting the scope of Dismissal for just cause, where the actors mostly concerned are SMEs and the dealt cases are about individual workers. As an example of a mid-flyer we can find topic Overtime work, dealing with theme of compensation for overtime work, a subjects that is well-known and established, but given its numerous interpretations and social relevance, is bound to heat up from time to time, with the Court of Cassation acting an actor of arbitration and regulation for the diverse options open in the execution of the processes.

Finally, we can see in Figure 4 these three topics against two topics that have been deemed as non-performing by the domain experts. As can be expected, the topics develop in a very flat way, by spreading evenly among all processes, and thus lack the capability of highlighting major turning points in their definition and implementation.

As a note on related work, a somewhat similar equation has been applied in [9] with a different purpose, namely the statistical and quantitative reconstruction of the history of ideas in a variety of scientific areas. Indeed, the focus of [9] is a case study on the evolution of research directions in computational linguistics through the topic-based analysis of 12,500 articles published in major international conferences in the field between 1980 and 2005. It is interesting to note that the trends reported in this work show characteristics significantly different from those reported here. In fact, the graphics contain mostly gentle slopes as opposed to the abrupt peaks with steep ascents and descents characterizing, albeit with varying levels of abruptness, our graphics. This difference has a variety of reasons, the most obvious and paramount of which is that the analyzed documents are, in this case, typically relevant for the advancement of a scientific discipline, a phenomenon that is indeed distributed over time, but with none of the characteristics of processes aimed at attaining specific goals, with which the documents analyzed in our case are associated. Thus, the gradual emerging and waning of ideas, caught by such softer uphills and downhills, is precisely what we expect, in contrast to the sharp phases of adjustment, as required by the practical needs of the moment, that characterize our topics. [10] do address the notion of hot topic in a vein very similar to ours, but their formal and computational treatment falls completely outside PTM and LDA, and in fact is term-oriented rather than topic-oriented. Thus, it does not appear suitable for the spotting of hot topics from large content corpora which is our purpose here, while it may be particularly suitable for their identification in the context of short texts, such as the twits or the threads of social networks.
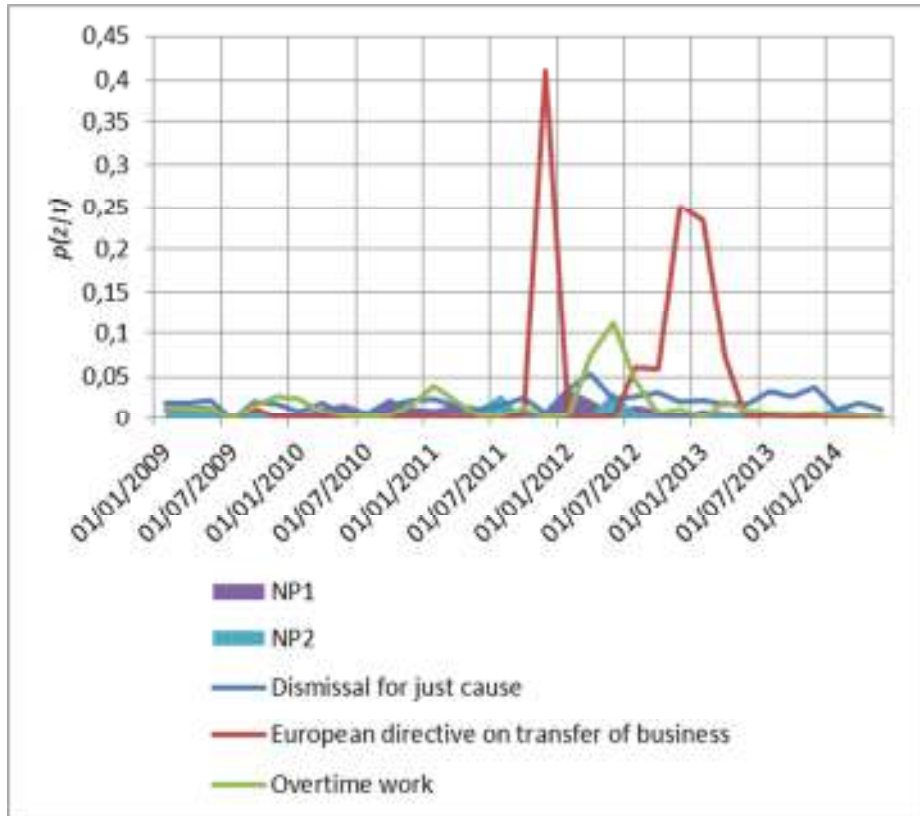
Fig. 4. Topic evolution of significant topics compared to non-performing topics (NP).

4   Conclusions

We have introduced the notion of open-world process in order to capture large pro-
cesses that involve multiple organizations, and are influenced by a number of contex-
tual factors and stakeholders. Furthermore, we have introduced the notion of hot topic
so as to provide a computationally effective way to track context evolution around
open-world processes in terms of the information that flows into such processes at
various degrees of density over time. Hot topics are themselves an application of sta-
tistical inference in the form of PTM, and thus are firmly rooted into empirical evi-
dence, without sacrificing high level representations of the inferred meanings. Hence
they show promise to be effectively combined with existing high-level representations
of business processes. There are a number of research directions that can and, in order
to obtain useful and concrete results, must be pursued to evolve this initial contribu-
tion. A most immediate one is to carve out a framework for Open-world Business
Process Management within the wider field of Business Process Management [11],
the discipline that encompasses the established computational framework for the

10

management of closed-world processes, namely Workflow Management. One clear direction to achieve this is to combine our statistical approach to the monitoring of information flows in open-world processes with a general framework for the monitoring of events over time in global systems, like, for instance, the reactive version of the Event Calculus proposed in [12]. Once verified the heating up of a topic over a defined interval of time such a calculus may then trigger the inspection of the pertinent contents, and the consequent retuning of the closed-world processes in the participating organizations that are synchronized around the relevant open-world process, through adaptive technologies like adaptive workflow management [13], [14] and case handling management [15].

Another crucial step is to survey and consequently map the existing open-world processes from the informal sources through which they are currently documented into a more rigorous notation. Given the inherent definitional fluidity of open-world processes, in place of the formal notations commonly used for closed-world processes, like Petri Nets and Workflow Nets, it may be preferable to use one aimed at handling incomplete information and soft constraints, such as the Generalized Process Structure Grammars (GPSGs) presented in [16]. GPSGs that encode open-world processes could in fact include rules containing symbols as names for topics generated through LDA/PTM, thus acting as interfaces between process definitions and contextual information flows. Another approach that could be similarly adopted, being itself based on a declarative formalism for the definition of flexible processes, is the one presented in [2].

As far as the topics are concerned, there are further possible constructions that can help us to gain insight about the processes they are associated with. In particular, one possibility we plan to explore is the generation of links among semantically related content objects clustered within the topics, following the methodology presented in [17]. In the context of the specific case study presented here, based on the implementation of labor law, this would allow us to rate the relevance of judgments on the basis of how many other judgments make reference to it, this being a clear case of semantic proximity. However, given that semantic proximity between content objects is itself probabilistically computed, other less obvious relationships would emerge. In this way, we could gain access not just to hot topics, but also to hot objects.

5   Bibliography

1. L. Baresi, E. Di Nitto e C. Ghezzi, «Toward Open-World Software: Issues and Challenges,» IEEE Computer 39(10), pp. 36-43, 2006.
2. M. Pesic e W. M. van der Aalst, «A Declarative Approach for Flexible Business Processes Management,» in Business Process Management Workshops, 2006.
3. H. J. Johansson, P. McHugh, J. Pendlebury e W. A. Wheeler, Business Process Reengineering: Breakpoint Strategies for Market Dominance, John Wiley & Sons, 1993.
4. D. M. Blei, «Probabilistic Topic Models,» Commun. ACM 55(4), pp. 77-84, 2012.
5. M. Steyvers e T. Griffiths, «Probabilistic Topic Models,» in Handbook of Latent Semantic Analysis, 2007, pp. 424--440.

6. J. B. Tenenbaum, C. Kemp, T. L. Griffiths e N. D. Goodman, «How to grow a mind: statistics, structure, and abstraction,» Science 331 (6022), pp. 1279-1285, 2011.

7. D. M. Blei, A. Y. Ng e M. Jordan, «Latent Dirichlet Allocation,» The Journal of Machine Learning Research, pp. 993--1022, 2003.

8. E. H. Ramirez, R. Brena, D. Magatti e F. Stella, «Topic Model Validation,» Neurocomputing, pp. 125-133, 2012.

9. D. Hall, D. Jurafsky e C. D. Manning, «Studying the History of Ideas Using Topic Models,» in EMNLP, 2008.

10. K.-Y. Chen, L. Luesukprasert e S.-c. T. Chou, «Hot Topic Extraction Based on Timeline Analysis and Multidimensional Sentence Modeling,» IEEE Trans. Knowl. Data Eng. 19(8), pp. 1016-1025, 2007.

11. W. M. P. van der Aalst, A. H. M. ter Hofstede e M. Weske, «Business Process Management: A Survey.,» in Business Process Management, 2003.

12. S. Bragaglia, F. Chesani, P. Mello, M. Montali e P. Torroni, «Reactive Event Calculus for Monitoring Global Computing Applications,» in Logic Programs, Norms and Action, 2012.

13. U. Borghoff, P. Bottoni, P. Mussio e R. Pareschi, «Reflective Agents for Adaptive Workflows,» in Second International Conference on the Practical Application of Intelligent Agents and Multiagent Technology, London, 1997.

14. M. Leitner, S. Rinderle-Ma e J. Mangler, «AW-RBAC: Access Control in Adaptive Workflow Systems,» in ARES, 2011.

15. W. M. P. van der Aalst, M. Weske e D. Grünbauer, «Case handling: a new paradigm for business process support,» Data & Knowledge Engineering 53 , pp. 129-162, 2005.

16. N. S. Glance, D. Pagani e R. Pareschi, «Generalized Process Structure Grammars for Flexible Representations of Work,» in CSCW, 1996.

17. M. Rossetti, R. Pareschi, F. Stella e F. Arcelli , «Integrating Concepts and Knowledge in Large Content Networks,» New Generation Comput. 32(3-4), pp. 309-330, 2014.

18. T. L. Griffiths e M. Steyvers, «Finding scientific topics,» Proceedings of the National academy of Sciences of the United States of America, pp. 5228-5235, 2004.