Interaction Behind the Scenes: Exploring Knowledge and User Intent in Interactive Decision-Making Processes

Rafael R.M. Brandão[✉], Marcio F. Moreno,
and Renato F.G. Cerqueira

IBM Research, Rio de Janeiro, Brazil
`rmello@br.ibm.com`

Abstract. Logging user interaction data with computational artifacts can be handy in identifying activities and issues associated with interactive decision-making processes. However, while such data commonly results in a temporally linear construction, information involved in such processes is not well structured from a knowledge engineering perspective. Consequently, both its consumption and understanding are not straightforward processes. Considering highly immersive environments with interaction through multiple modalities, the tracking of such knowledge becomes even more complex. Such environments have been increasingly used to support decision-making practices, which may involve cognitive-intense activities and critical thinking. Inferring concepts and knowledge from logging data in such activities is key for improving design of decision support systems, and general systems as well.

Keywords: Interaction log reasoning · Knowledge modeling · User intent · Interpretive trail · Reasoning provenance · Cognitive computing systems

1 Introduction

Interaction logs are valuable resources for exploring knowledge in analytical reasoning processes supported by software artifacts and tools. However, tracking user intent and reasoning from low-level data is a challenging task. Interactive analytical processes involve cognitive-intense activities, where tacit and explicit knowledge are applied to achieve a defined goal. In addition, investigative processes commonly follow an abductive reasoning approach, where analysts or decision-makers test hypotheses using the best available information. That is, these processes commonly comprise uncertainty and incompleteness in data.

An interesting way of seeing decision-making [1] is to consider it as a cognitive process of making choices by setting goals, identifying and gathering information (evidence), reflecting and choosing alternatives to take actions. On the one hand, to decision makers, the process of effectively producing and consuming semantically structured and relevant multi-modal information is crucial. On the other hand, representing knowledge from unstructured data, such as video and audio streams, without a defined semantic model, can be challenging. In the same way, structuring this type of

data comprehends a costly process, since conventionally this is achieved by manual annotation and human interpretation. We identified and classified challenges involving multimedia research in decision-making processes into four categories: challenges related to knowledge extraction in multimedia content, consumption of knowledge through multimedia content, capturing decision makers' intent in multimedia content and modeling decision-making processes with multimedia content [2].

Roberts et al. [3] propose to tackle the problem of knowledge provenance in interactive analytical scenarios at three different conceptual levels: provenance of data, analysis and reasoning. Provenance of data is at the most basic abstraction level. All data and their sources should be registered and associated. They may come from a wide range of sources, e.g. automated capture devices and sensors, or documents written in natural language. Keeping track of data routing is essential to maintain the quality and reliability of the information. Provenance of analysis is related to how the user inter-pretation is carried out, i.e. what actions were performed, what interaction events were triggered. Different techniques may be applied to process and visualize exploration trails. Provenance of reasoning is at the highest level of abstraction, dealing with how decision makers and analysts arrive at their conclusions.

Table 1 shows the conceptual levels discussed by Roberts et al. [3] with a main question related to each provenance level, plus relevant content and useful resources to support activities at each level. The data provenance level can be related to the question "where did this information come from?". Different research fields have varied interests on data provenance. Particularly in the e-Science context, data repository solutions often focus on aspects such as versioning and parameter settings. Generally, data provenance solutions register any changes that may influence on data of interest.

Table 1. Provenance levels, relevant content and useful resources to be captured, based on [3].

Conceptual level	Main question	Relevant content	Useful resources
Reasoning provenance	"How did you arrive at these conclusions?"	Analyst's/decision-maker's interpretation	Audio, video, annotations
Analysis provenance	"How was the analysis performed?"	Actions taken and techniques to process and visualize data	Interaction logs, screen video
Data provenance	"Where did this information come from?"	Data routing	Data models, metadata standards

The analysis provenance level is related to the question "how was the analysis performed?". This level can be supported by instrumenting tools or software artifacts used in the analysis process. By instrumenting these tools, it would be possible to create a history with interaction logs describing user interaction. It is also possible to record this interaction in video, exhibiting the user's screen. One need to balance a trade-off between capturing a massive number of fine-grained actions or registering more coarse-grained, composite actions and associated semantics.

Capturing reasoning provenance can be related to the question "how did you arrive at these conclusions?". In this level, one cannot straightforwardly automatize the

provenance process. It often demands analysts to externalize tacit knowledge and intent, expressing their reasoning through annotation, audio or video. However, some argue that this externalization process by itself can potentially change the reasoning's nature and hinder analysis performance [4].

Investigative and interpretive activities have a strong iterative character. No matter how structured the analysis is, surprises and disappointments will happen. New questions are introduced all the time and these can be promptly investigated or analyzed later. This "revisiting" aspect is a very common trait in qualitative analysis, where analysts need to categorize evidence data fitting it into certain conceptual classifications. The process of capturing and revisiting or accessing data is the focal point of the research topic called Capture & Access (C&A). Per Truong et al. [6], C&A can be defined as the "task of preserving a record of some live experience that is then reviewed at some point in the future. Capture occurs when a tool creates artifacts that document the history of what happened". Where, live experience may comprise any social event or moment whose record can be useful. These artifacts are recorded as streams of information that flow through time and can be accessed later.

The work presented in [7] discusses how a ubiquitous infrastructure for C&A can be used in the context of scientific investigations. The work [7] proposes to structure undertaken investigative procedures into hypermedia documents with analyses and validations, allowing its representation in a theoretical model. This model enables the outlining of the research inquiry, providing semantics to allow relationship between key elements in a qualitative methodology.

On a different perspective, the work discussed by Kodagoda et al. [5] applied machine learning techniques in an attempt to infer and reconstruct interpretive or reasoning trails by statistically classifying activity from log data. Kodagoda et al. used a theory of sensemaking as the basis for inferring reasoning from actions. A training dataset was created through a manual process of coding interactive logs, based on capturing a verbal protocol and interviews with analysts.

In this work, we explore how the approach presented in [8–10] can assist designers and developers when modeling scenarios involving collection and processing of interaction logs carrying unstructured knowledge. Our proposal is to provide a conceptual model geared towards promoting better expressiveness to authors wanting to represent possible relationships among cognitive systems (humans or software), their tools (software tools, devices, physical objects and respective representation), conceptual knowledge and semantics present in perceptual data. Moreover, we argue in favor that the process of capturing and acquisition of data produced in cognitive activities should also be integrated, promoting a better knowledge structuring around the modeled practice.

Our approach is based on the Nested Context Model (NCM) [8] that has been widely applied in the multimedia context. Our recent extensions [8, 9] to this model integrate support for rich knowledge description, along with specification of relationship between knowledge and multimedia data. We named this integration of hypermedia aspects and knowledge engineering as *hyperknowledge*. Through this model, it is possible to specify traditional multimedia features, such as logical structuring and spatiotemporal synchronization among media content, in conjunction with abstract concepts and knowledge structuring, in a single rationale. By integrating such

correlated concerns, we expect to simplify the specification (and eventually developing software to support) of scenarios involving reasoning over data from multiple devices and logging from different interactive software artifacts. Bringing its original features as a hypermedia model, NCM also supports specifying how pertinent data should be presented and navigated according to users' preferences and available resources. That is, it can support the creation of structured narratives expressing implicit and explicit knowledge, as well as material evidence and hypotheses explored in a given analysis. We explore how handling issues related to the three aforementioned provenance levels altogether could support reflecting over users' intent, delineating an interpretive trail from their interaction. In other words, how our approach can support model produced and consumed knowledge by users during their interaction.

2   Background

This section presents the basic concepts of NCM conceptual model, including our recent extensions to enrich knowledge modeling support.

NCM defines an *Entity* class, which has as main attribute its unique identifier. The foundation of NCM is the usual hypermedia concepts of *Nodes* and *Links* [8, 9]. The former, illustrated on Fig. 1, is an *Entity* that represents information fragments, while the latter is an *Entity* that has the purpose of defining relationships among interfaces (*Anchors*, *Ports*, and *Properties*) of *Nodes*. There are two basic classes of *Nodes*: *ContentNode* and *CompositeNode*.

A *ContentNode* represents the usual media objects. *ContentNode* subclasses define the content type (e.g. video, audio, image, text, concepts, etc.). To define its content, a *ContentNode* can use a reference (e.g. URL) to the content or have a byte array of the content (raw data).
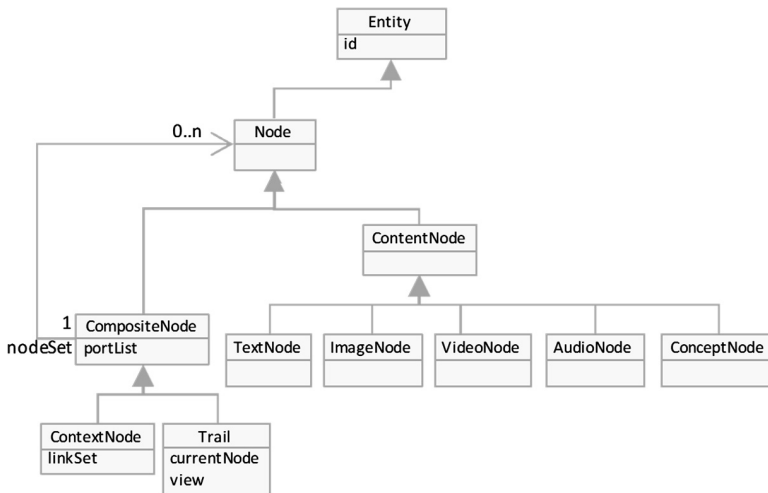

Fig. 1. NCM class hierarchy: the node entity.

A *CompositeNode* is an NCM *Node* whose content is a set of nodes (composite or content nodes). The set of nodes constitutes the composite node information units. In a *CompositeNode*, a *Node* cannot contain itself. *CompositeNode* subclasses define semantics for specific collections of nodes. A *ContextNode* is an NCM *CompositeNode* that also contains a set of links and other attributes [8, 9]. *ContextNodes* are useful, for instance, to define a logical structure for hypermedia and hyperknowledge documents. A *Trail* is an NCM *CompositeNode* that offers content navigation mechanisms. For instance, a *Trail* provides mechanisms to show to a user how the current navigation status (*currentNode* attribute) was achieved showing the navigation history (view attribute). It can be also used to structure the order in which each knowledge was generated, enabling consumers of this knowledge to navigate on a temporal axis (causal and constraint axis can also be considered).

Figure 2 presents the UML diagram of NCM focusing on *Link* and *Connector* entities. A *Link* has two additional attributes: a *Connector* and a set of *Binds*. The *Connector* defines the semantics of a relation through an NCM class named *Glue*, independently of the components that will be included in the relation [8, 9], and a set of access points, called *Roles*. A *Glue* describes how roles must interact and must consider the use of all roles in the connector. The concept of event[1] is the foundation of the *Role* class. Therefore, each role describes an event to be associated to a component of the relation. There are different subclasses of *Role*. Each connector type can use a different set of roles. Back to Fig. 2, in the set of *Binds* of the *Link*, each *Bind* associates each *Link* endpoint (interfaces of *Nodes*) to a *Role* at the referred *Connector*.
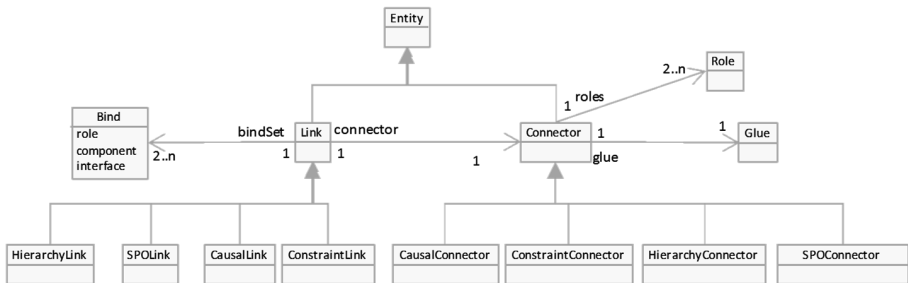

Fig. 2. NCM hierarchy: link and *Connector*.

Theoretically, *Connectors* can represent any type of relation. NCM 3.0 supports the specification of spatio-temporal synchronization relations through causal (*CausalGlue* that can hold *ConditionRoles*, *AssessmentRoles* and *ActionRoles*) and constraint (*ConstraintGlue* holding *AssessmentRoles*) *Connectors*. A condition must be satisfied in a causal relation to execute a group of one or more actions. For instance, a document author can specify a connector that will start (*ActionRole* "start") the presentation of

---

[1] NCM uses the definition of event as stated in the Pérez-Luque and Little work [13]: an event is an occurrence in time that may be instantaneous or may extend over a time interval.

one or more *Nodes* when the presentation of one or more *Nodes* finishes (*Condition Role* "onEnd") or when the *Property* "top" of two or more Nodes receives the same value (*AssessmentRole* evaluating *Property* values). On constraint relations, there is no causality involved. For instance, a *ConstraintConnector* can define that two or more *Nodes* must begin (*AssessmentRole* "begins") their presentation at the same time and must end (*AssessmentRole* "ends") their presentation at the same time.

Besides supporting causal and constraint relations, NCM supports hierarchical descriptions through hierarchy connectors (*HierarchyGlue* that holds *HierarchyRoles*) and SPO (Subject-Predicate-Object [8, 9]) triples descriptions through SPO connectors (*KnowledgeGlue* that can hold *ConditionRoles*, *AssessmentRoles*, *ActionRoles*, *SubjectRoles*, *ObjectRoles*, and *InferenceRoles*) connectors. The *HierarchyRole* defines the participant function ("parent" or "child") in the relation to represent the hierarchy, as in "Ana" is an instance of "Person". The subject and object roles represent, respectively, a Subject and an Object in the traditional SPO relations. For instance, to model the statement that "Ana moved the mouse", the *ConceptNode* "Ana" must be connected to a *SubjectRole*, while the *ConceptNode* "mouse" must be connected to the *ObjectRole* "moved". Note that the names of *ObjectRoles* have semantics, acting as predicates [8, 9]. Finally, the *InferenceRole* indicates which participant in the relation shall be considered to infer (defining the inference direction, "from" or "to") the data according to a knowledge presentation.

3 Knowledge Engineering on Analytical Activities

Both explicit and implicit knowledge can be used to support analysis and decision-making processes. Roberts et al. [3] distinguish such forms of knowledge as hard and soft data. Hard data is typically related to explicit knowledge and quantitative data, with a known source and provenance. In contrast, soft data reflects implicit knowledge such as background information, personal experiences, and tacit knowledge. Roughly speaking, soft data are more related to reasoning provenance, while hard data relates to analysis provenance. To structure and engineer knowledge pertaining to analytical activities, it is key to capture and understand both types of data.

Figure 3 illustrates a typical scenario of cognitive intensive activities supported by our approach. Users interact with a software tool while interaction events along with internal system events are captured in log files. Also, visual interaction may be captured by recording their screens. These data are specifically related to the level of analysis provenance. Simultaneously, users' externalized attitudes and verbal protocol can be recorded by capture devices and various sensors, bringing information that may be related to the level of reasoning provenance. Environmental information may also be considered, adding to the understanding of the context where the analysis was performed. The correlation of these distinct data is performed by a module named Content Understanding.

The Content Understanding module comprises different components. A component for Media Processing is responsible for parsing multimodal content, dealing with data transcoding, fission, and fusion issues. The Knowledge Extraction component addresses issues such as identifying and classifying named entities, in order to extract
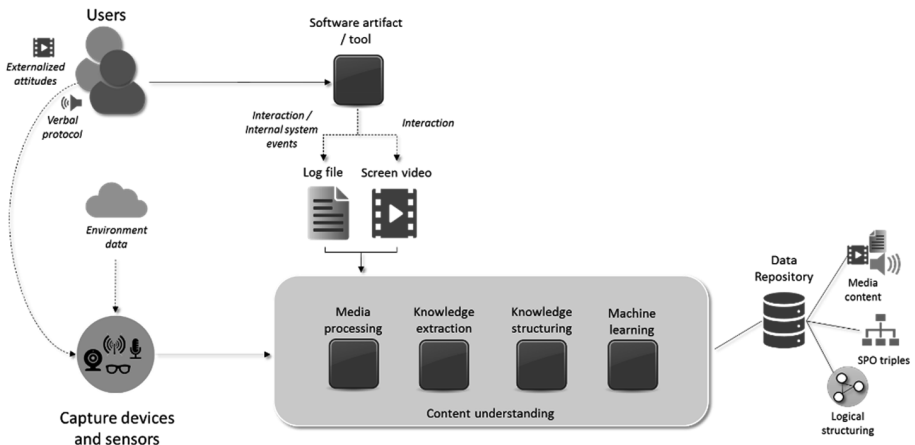
Fig. 3. Proposed knowledge engineering approach to extract and structure knowledge from analytical activities.

and annotate key concepts present in data. It also comprises algorithms for speech processing and for recognizing user sentiment. The Knowledge Structuring component deals with logical organization of knowledge. It is responsible for correlating information from analysis through pre-defined structuring, such as timeline organization or concept similarity. The Machine Learning component abstracts features from algorithms for supervised, unsupervised, and reinforced learning. Considering supervised learning, it is possible for the system to refine knowledge through more accurate classifications through user feedback. If the analysis is carried over unclassified data, clustering techniques may be used to structure data through unsupervised learning. If the analysis involves an iterative process, such as decision-making, reinforcement learning techniques can be applied, given a metric or policy to measure "reward" or "punishment" for the goals of this analysis. A data repository is used to maintain all data (hard and soft), abstract knowledge representation as SPO (Subject-Predicate-Object) triples and logical structuring of this information.

3.1 Structuring and Visualization

In addition to capturing the knowledge generated in a process of analysis or decision, it is necessary to structure the information so that it can be efficiently consumed. Visualizing data through structured narratives or storytelling is an interesting strategy. Narratives make knowledge consumption more natural and compelling. They can explicitly represent addressed hypotheses, and the reasoning applied by presenting different hard and soft data. As highlighted in [3], such narrative structuring can be beneficial not only by facilitating understanding of analysts' reasoning and evidence base, but also by enabling consumers of this "story" to communicate it to others as well. Different narrative styles can be applied for data visualization, including animation/video, slide show, flowcharts among others described in [11].

The most direct approach to present a given reasoning process is to create a temporally linear structuring in the order in which each knowledge was generated, enabling consumers of this knowledge to navigate on a temporal axis. In this manner, a basic support system can present sequences of knowledge, signified in different media content. Annotations and other textual information can be used as background narration or as text notes with arguments throughout the presentation. All in all, a linear structure may not be the best way to present intricate reasoning and multiple analyses, with several ramifications and nesting possibilities. Figure 4 illustrates a hypothetical linear structuring of a narrative.
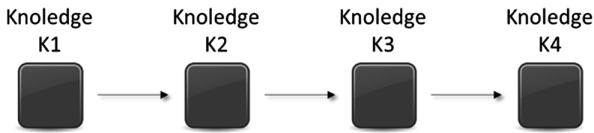

Fig. 4. Narrative with possible linear structuring.

This linear limitation can be easily bypassed through NCM's main abstraction: nested contexts. Contexts allow narratives to be modeled in a range of different ways. They can be nested recursively, creating any desired logical structure to group knowledge and data. For example, grouping knowledge by its similarity. Or, if multiple analysts participated collaboratively in an analysis, knowledge associated with each person could be grouped together allowing navigation in different contexts. Virtually any feature or concept can be used as a parameter for grouping in nested contexts. It is up to the user to model the desired structuring based on his domain knowledge. Figure 5 illustrates a hypothetical grouping through nested contexts considering an arbitrary aspect.
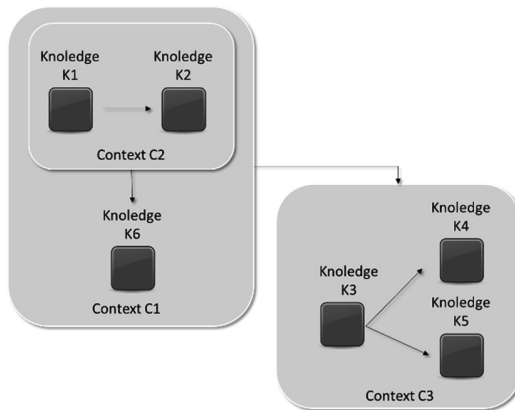

Fig. 5. Nested contexts support for grouping related content and knowledge.

Another potential structuring feature in the model is the *Trail* concept, which is in line with the need of keeping track of the navigation performed by analysts. Trails can represent the path to the current node (content or concept node), or for instance a sequence of knowledge nodes that have been created for a given analysis.

NCM's flexibility supports modeling information relative to the three different levels of abstraction of provenance in a single "notation space". That is, the model provides mechanisms to support description of media content with its sources (data provenance). It supports description of temporal events and relationship between abstract knowledge through its various connectors (analysis provenance). And it supports modeling a capture scenario to record knowledge around analysts' decisions (reasoning provenance).

4   Next Steps

Given the extent and complexity of the challenges involved on structuring interpretive reasoning in decision-making, our research agenda firstly focused on the theoretical and conceptual foundations required for outlining a big picture for the proposed approach. This rationale aims at understanding the basic aspects of a comprehensive domain, which integrates data capture, content understanding and multimedia visualization in a holistic way.

As a next step, we plan to conduct a study to identify common constructs and structures in narrative styles that could be generalized as templates in NCM notation. The idea is to facilitate the creation of new interpretive trail visualization from structured narratives. The narrative styles described by Setel and Heer in [11] is a basis for identifying such structures.

Another future direction is related to using reasoning provenance to enhance analysis during activities. Typically, systems that support reasoning provenance are used at a timeframe subsequent to the analysis activity. That is, they provide mechanisms and constructions to users reflect on the outcomes of an earlier analysis. This is in line with what Schon [12] refers to as reflection-on-action. That is, generating resources to evaluate or audit investigations. However, our proposed approach aims to also support analysts during their analysis, which Schon refers to as reflection-in-action. If decision support systems can infer reasoning provenance on-the-fly, it would be possible to use visual representation of this provenance as an epistemic tool by itself, hopefully generating more knowledge during analysis.

5   Final Remarks

This paper presents our vision and approach around how to model and structure knowledge in cognitive-intense analytic activities. Our proposal aims at integrating concerns that are addressed separately and distinctly into a single rationale. More specifically modeling of data capture, content processing and understanding and data visualization. Through a conceptual model, we explore how aspects of these contexts can be represented and related in a possible knowledge engineering strategy.

The main contribution of our approach is to shed light on this possible holistic view, since what we generally observe in the literature are tools that operate individually and do not support capture and visualization of implicit knowledge. Through this research line, we hope to inspire other researchers and practitioners to reflect and seek to establish a global view on the issue of reasoning provenance.

References

1. Power, D.J., Sharda, R., Burstein, F.: Decision support systems. In: Cooper, C.L. (ed.) Wiley Encyclopedia of Management, pp. 1–4. Wiley, Chichester (2015)
2. Moreno, M.F., Brandão, R., Cerqueira, R.: Challenges on multimedia for decision-making in the era of cognitive computing. In: 2016 IEEE International Symposium on Multimedia (ISM), pp. 673–678. IEEE (2016)
3. Roberts, J.C., Keim, D., Hanratty, T., Rowlingson, R.R., Walker, R., Hall, M., Varga, M.: From Ill-defined problems to informed decisions. In: EuroVis Workshop on Visual Analytics (2014)
4. Hertzum, M., Hansen, K.D., Andersen, H.H.K.: Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload? Behav. Inf. Technol. **28**, 165–181 (2009)
5. Kodagoda, N., Pontis, S., Simmie, D., Attfield, S., Wong, B.L.W., Blandford, A., Hankin, C.: Using machine learning to infer reasoning provenance from user interaction log data: based on the data/frame theory of sensemaking. J. Cogn. Eng. Decis. Mak. **11**, 23–41 (2017)
6. Truong, K.N., Abowd, G.D., Brotherton, J.A.: Who, what, when, where, how: design issues of capture & access applications. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) Ubicomp 2001: Ubiquitous Computing. LNCS, vol. 2201, pp. 209–224. Springer, Berlin (2001). doi:10.1007/3-540-45427-6_17
7. Brandão, R.R.M.: A capture & access technology to support documentation and tracking of qualitative research applied to HCI (Doctoral thesis). PUC-Rio, Rio de Janeiro, Brazil (2015)
8. Moreno, M.F., Brandão, R., Cerqueira, R.: Extending hypermedia conceptual models to support hyperknowledge specifications. In: 2016 IEEE International Symposium on Multimedia (ISM), pp. 133–138. IEEE, December 2016
9. Moreno, M.F., Brandao, R., Cerqueira, R.: NCM 3.1: a conceptual model for hyperknowledge document engineering. In: Proceedings of the 2016 ACM Symposium on Document Engineering, pp. 55–58. ACM, September 2016
10. Moreno, M.F., Brandao, R., Ferreira, J., Fucs, A., Cerqueira, R.: Towards a Conceptual Model for Cognitive-Intensive Practices. IEEE ISM, San Jose (2016)
11. Segel, E., Heer, J.: Narrative visualization: telling stories with data. IEEE Trans. Vis. Comput. Graph. **16**, 1139–1148 (2010)
12. Schon, D.A.: The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York (1984)
13. Pérez-Luque, M.J., Little, T.D.C.: A temporal reference framework for multimedia synchronization. IEEE J. Sel. Areas Commun. **14**(1), 36–51 (1996)