

Chapter 132

Representation Languages for Unstructured 'Narrative' Documents

Gian Piero Zarri
University Paris-Est, France

Category: Technologies for Knowledge Management

INTRODUCTION

A big amount of important, 'economically relevant' information, is buried into unstructured, multimedia '*narrative*' resources. This is true, e.g., for most of the corporate knowledge documents (memos, policy statements, reports, minutes etc.), for the news stories, the normative and legal texts, the medical records, many intelligence messages, the 'storyboards/historians' describing sequences of events in industrial plants, the surveillance videos, the actuality photos for newspapers and

magazines, lot of material (text, image, video, sound...) for eLearning etc., as well as, in general, for a huge fraction of the information stored on the Web. In these 'narrative documents', or 'narratives', the main part of the information content consists in the description of 'events' that relate the real or intended behavior of some 'actors' (characters, personages, etc.) – the term 'event' is taken here in its more general meaning, covering also strictly related notions like fact, action, state, situation etc. These actors try to attain a specific result, experience particular situations, manipulate some (concrete or abstract) materials, send or receive messages, buy, sell, deliver etc. Note that, in these narratives, the actors or personages are not necessarily human beings; we can have narrative documents concerning, e.g., the vicissitudes in the journey of a nuclear submarine

DOI: 10.4018/978-1-59904-931-1.ch134

(the ‘actor’, ‘subject’ or ‘personage’) or the various avatars in the life of a commercial product. Note also that, even if a large amount of narrative documents concerns natural language (NL) texts, this is not necessarily true, and ‘narratives’ are really ‘multimedia’. A photo representing a situation that, verbalized, could be expressed as “The US President is addressing the Congress” is not of course an NL text, yet it is still a narrative document.

Because of the ubiquity of these ‘narrative’ resources, being able to represent in a general, accurate, and effective way their semantic content – i.e., their key ‘meaning’ – is then both *conceptually relevant and economically important*: narratives form, in fact, a huge underutilized component of organizational knowledge, and people could be willing to pay for a system able to process in an ‘intelligent’ way this information and/or for the results of the processing. This type of explicit yet unstructured knowledge can be, of course, indexed and searched in a variety of ways, but it requires, however, *an approach for formal analysis and effective utilization that is neatly different from the ‘traditional’ ones*.

BACKGROUND

Usual ontologies – both in their ‘traditional’ and ‘semantic web’ versions, see the “Knowledge Representation for Knowledge Management” and “RDF and OWL for Knowledge Management” Chapters in the Encyclopedia – are not very suitable for dealing with narratives. Basically, ontologies organize the ‘concepts’ – that we can identify here with the important notions to be represented in a given application domain – into a hierarchical structure, able to supply an elementary form of definition of these concepts through their mutual generic/specific relationships (‘IsA’ links). A more detailed definition of the concepts is obtained by associated with them a set of binary relationships of the ‘property/

value’ type (e.g., a ‘frame’). The combination of these two representational principles is largely sufficient to provide a *static, a priori* definition of the concepts and of their properties.

Unfortunately, this is no more true when we consider the *dynamic behavior* of the concepts, i.e., we want to describe their mutual relationships when they take part in some concrete action, situation etc. (‘events’). First of all, representing an event implies that the notion of ‘role’ must be added to the traditional generic/specific and property/value representational principles. If we want to represent adequately a narrative fragment like “An important media company ... will develop a new model of Internet cellular phone...”, besides asserting that MEDIA_COMPANY_1 is an instance of the concept media_company and that we must also introduce a specific instance of a concept like internet_cellular_phone, we have to create a sort of ‘threefold’ relationship including a ‘predicate’ (like DEVELOP or PRODUCE), the two instances, and a third fundamental component, the ‘roles’ (like SUBJECT or AGENT for MEDIA_COMPANY_1 and OBJECT or PATIENT for the new cellular phone), used to specify the exact function of these two instances within the formal description of the event. Moreover, in an event context, we must also deal with those ‘*connectivity phenomena*’ like causality, goal, indirect speech, co-ordination and subordination etc., that link together the basic ‘elementary events’. It is very likely, in fact, that, dealing with the sale of a subsidiary, the global information to represent could be something like: “Company X has sold its subsidiary Y to Z *because* the profits of Y have fallen dangerously these last years *due to* a lack of investments” or, returning to our previous example, that “MEDIA_COMPANY_1 will develop a new model of Internet cellular phone *to reply* to a direct competitor move” or that, dealing with the relationships between companies in the biotechnology domain, “X made a milestone payment to Y *because* they decided to pursue an *in vivo* evaluation of the candidate compound

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/representation-languages-unstructured-narrative-documents/49083

Related Content

Knowledge Power

(2014). *Harnessing Dynamic Knowledge Principles in the Technology-Driven World* (pp. 1-13).

www.irma-international.org/chapter/knowledge-power/83667

Managing Knowledge in Digital and Creative Design-Decision-Making for Heritage Innovation

Hui Jiang (2026). *International Journal of Knowledge Management* (pp. 1-19).

www.irma-international.org/article/managing-knowledge-in-digital-and-creative-design-decision-making-for-heritage-innovation/400251

Social Media Knowledge

(2014). *Harnessing Dynamic Knowledge Principles in the Technology-Driven World* (pp. 219-227).

www.irma-international.org/chapter/social-media-knowledge/83682

Who are Effective Time Managers? Bivariate Correlation Analysis and Hypotheses Testing

Dezhi Wu (2010). *Temporal Structures in Individual Time Management: Practices to Enhance Calendar Tool Design* (pp. 116-138).

www.irma-international.org/chapter/effective-time-managers-bivariate-correlation/36651

Open Innovation Maturity

Hagen Habicht, Kathrin M. Möslein and Ralf Reichwald (2012). *International Journal of Knowledge-Based Organizations* (pp. 92-111).

www.irma-international.org/article/open-innovation-maturity/61430