NewsReader: recording history from daily news streams

Piek Vossen[♣], German Rigau[♢], Luciano Serafini[♠], Pim Stouten[♡], Francis Irving[♣♠], Willem Van Hage[♠♢]

[♣]VU University of Amsterdam, Amsterdam, [♢]Euskal Herriko Unibertsitatea, San Sebastian,
[♠]Fondazione Bruno Kessler, Trento, [♡]LexisNexis BIS, Amsterdam,
[♣♠]ScraperWiki, Liverpool, [♠♢]Synerscope, Eindhoven
[♣]piek.vossen@vu.nl, [♢]g.rigau@ehu.es, [♠]serafini@fbk.eu,
[♡]pim.stouten@lexisnexis.com, [♣♠]francis@scraperwiki.com, [♠♢]willem.van.hage@synerscope.com

Abstract

The European project NewsReader develops technology to process daily news streams in 4 languages, extracting what happened, when, where and who was involved. NewsReader does not just read a single newspaper but massive amounts of news coming from thousands of sources. It compares the results across sources to complement information and determine where they disagree. Furthermore, it merges news of today with previous news, creating a long-term history rather than separate events. The result is stored in a KnowledgeStore, that cumulates information over time, producing an extremely large knowledge graph that is visualized using new techniques to provide more comprehensive access. We present the first version of the system and the results of processing first batches of data.

Keywords: news streams, cross-lingual event extraction, long-term histories

1. Introduction

We believe that we stay informed about the changes in the world by tracking the news and our social networks. However, every working day millions of news articles are produced from thousands of different sources and this number is increasing, as reported by LexisNexis (a large international news brooker). We are thus necessarily extremely selective in the sources we monitor and we simply hope we made the right choices. Besides, we have no idea what all the different sources have to offer. We do not have a good view on what to choose from. Current technology solutions cannot handle these daily streams of news at a very detailed level. They only partially capture (mostly trending) topics in terms of clusters, keywords, named entities and overall opinions but they do no truly represent the changes reported in the news and they cannot compare what is reported across the different sources. Furthermore, these solutions can only measure changes in trendiness of topics, e.g. through timelines or maps showing the rise and fall of a topic or the spread over the world. They do not interpret sequences of specific events as longer term developments or stories as they unfold in time.

To fill this gap, the NewsReader project[1] processes massive amounts of daily news streams in terms of reported changes to reconstruct long-term sequences or stories in time. The visualizations of these storylines are expected to be a more efficient and provide a more natural summarization of the changing world with more explanatory power.

To achieve this, the software automatically reads news in 4 language (English, Spanish, Italian and Dutch), determining *what* happened, *where* and *when*, and *who* was involved. This is done for massive amounts of news articles coming from thousands of different sources, where we compare the news of a single day to find out what they share and where they differ. Furthermore, we merge the news of today with previously stored information, creating a long-term history rather than storing separate events. Through this, we also separate new from old information and speculated information from actual events. The result is stored in a KnowledgeStore that acts as a so-called **history-recorder**, keeping track of the changes in the world as told in the media. The KnowledgeStore represents these changes as triples in RDF and supports reasoning over the data. Since we keep track of the origins of information, the platform also provides valuable insights into who told what story. This will tell us about the different perspectives from which different sources present the news.

The data produced in NewsReader is extremely large and complex: exhibiting the dynamics of news coming in as a stream of continuously updated information with changing perspectives. A dedicated decision support tool suite is developed that can handle the volume and complexity of this data and allows professionals to interact through visual manipulation and feedback and new types of representation. NewsReader will be tested on economic-financial news and on events relevant for political and financial decision-makers.

In this paper, we describe the first implementation of the system and the first results of processing data in the project. In section 2., we provide a more detailed example to explain the problem that we want to solve. In section 3., we give an overall description of the complete system and some of the fundamental design principles. Next, section 4., we describe the use cases defined so far in the project and after that in section 5., we report on the data processed so far. In section 6. we present our first ideas on the visualization and interaction with the data. Finally, we provide some conclusions and future work.

[1]NewsReader is funded by the European Union as project ICT-316404. It is a collaboration of 3 European research groups and 3 companies: LexisNexis, ScraperWiki and Synerscope. The project started on January 2013 and will last 3 years. For more information see: www.newsreader-project.eu/

2. The problem to be solved

Current solutions to monitor news streams tend to use time-lines, maps and clusters to indicate trendiness of topics. This is sometimes combined with sentiment indicators and other meta data, such as the owners of the source, the language of the text. Examples of such systems are the European Media Monitor[2], Yahoo Finance[3], Google News[4], Google Finance[5], Google Trends[6], Reuters[7], Dowjones[8], Factiva[9], LexisNexis[10]. All these solutions try to give users control over large volumes of news from different sources by indicating what is trending, what is the topic in keywords, what (famous) people and organizations are involved and how did the volume develop over time. An example of a timeline display of news around Volkswagen taken from Google-trends can be seen in Figure 1.

In this timeline, we see several peaks between 2004 and 2013 indicating major volumes of news around Volkswagen: something is going on. Whenever Google shows a letter, you can get a snippet indicating the topic of a peak. In Figure 1, we show two topic indicators. The peak G in 2009 shows a snippet from Business Standard saying that Porsche takes over Volkswagen. The peak A in 2012 tells you the exact opposite that Volkswagen takes over Porsche. What truly happened? To find out, you need to start reading the news.

This example illustrates three important aspects of all the systems mentioned above:

1. All systems alert users to trending topics based on the volume of news, where the signal is derived from the cluster as a whole and not from the individual news items.

2. The systems extract additional information from these clusters such as the entities mentioned, topical key words, or in some cases the overall sentiment.

3. Except for the overview results and the extracted data elements, all system provide the results at the document level without a deeper analysis of the event that is reported inside the document and how it relates to what is stated in other documents in the same cluster.

The problem with these approaches is that they do not really define what is reported in the news as an event in the world and do not relate this to what happened in the past. They miss an essential aspect of news, which is their reflection on changes in the world. As a result, none of these solutions can provide a schematic representation of what happened nor can they tell you which news items in a cluster tell the same story or a different story with respect to the changes in the world. They also miss another aspect, which is that news very often refers back to old information to provide a context and speculate about possible consequences in the future. All these aspects are illustrated by the example in Figure 1, in which Google fails to tell us what the status is of each take-over statement.

Figure 2 shows the problem in a more generic way. We shows two time-lines: one for what happened in the assumed world and another timeline for the publication date of the news that reflects on the world at a certain point in time. We want to reconstruct what happened in the world from news that not only reports on recent changes in the world but also connects this to the past and projects it to the future. The continuous stream of news further results in frequent updates with respect to the world, whenever new information comes in or the perspective changes. None of the news items tells the full story and many tell similar stories. Only by combining news over time a more complete story can be built.

The volume of news is therefore far bigger than the volume of changes in the world. LexisNexis estimates that the total volume of English news items on the car industry published since 2003 is about 6 million articles. Currently, we have no idea how many real facts are reported in the news and how much is repeated, refers to the past or is speculated. Our current estimates for entities based on 60K news articles (1% of the total) indicate that there are on average 200K mentions of entities in the news per year and that these mentions involve 10K different entities (on average each entity is mentioned 20 times per year). We have no idea yet what is the case for events. How much did the world change?

In the case of Porsche and Volkswagen, reading the articles reveals that Porsche has been taking stakes in Volkswagen for years, up to a stake of 75% in 2008. Since 2005, the media have been speculating about a take-over of Volkswagen by Porsche, which seemed apparent. Hence news reporting on a take-over by Porsche in 2008 by the Business Standard and many others. The financial crisis had a dramatic impact on this strategy. The year 2009 results in a complete turn around: Volkswagen takes a stake in Porsche of 75% resulting in a real take-over of Porsche by Volkswagen. The CEO of Porsche Wiedeking steps aside and is sued by hedge funds for failing to take-over Volkswagen. This is the storyline of Porsche and Volkswagen that explains the peaks in the trends shown by Google.

In order to be able to derive such storylines, news monitoring applications need to be event-based rather than topic and document-based. This requires the following analysis of text:

1. determine what happened (the low-level event and not just the topic), who was involved, where and when as reported in each news article.

2. determine which news articles report on the same event, how do they complement each other and how do they differ.

3. represent events as instance in the world, abstracting from the mentions in the text and make a distinction between factual events, counter-factual events and

[2] http://emm.newsbrief.eu/overview.html
[3] finance.yahoo.com
[4] https://news.google.nl
[5] https://www.google.com/finance
[6] http://www.google.nl/trends/
[7] www.reuters.com
[8] www.dowjones.com
[9] www.factiva.com
[10] www.lexisnexis.com

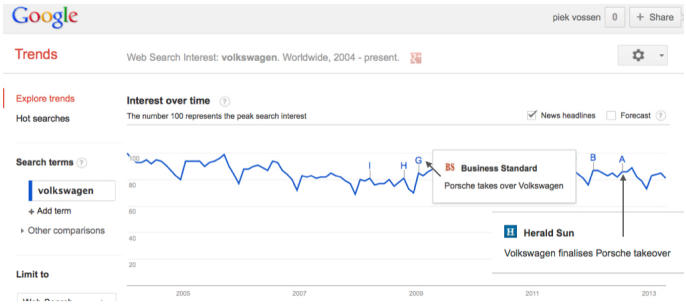

Figure 1: Timeline display of news around Volkswagen from Google-trends

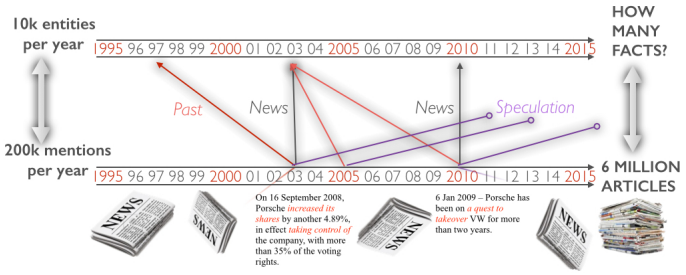

Figure 2: Diverging timelines for news and reality

speculated events (in the future).

4. relate events through time-lines, participants and where possible causal relations unfolding longer term developments.

This will enable us to track events in time that underly the above topics and trends. Such storylines explain what happened and why. They are more intuitive for users to grasp then purely quantitative information. They abstract from many mentions in different sources and can provide a more compact representation than a document based approach. In addition to these requirements, it is very important to keep track of all the sources of information and their opinions on what happened. By capturing this so-called *provenance*, we provide feedback on who told what story, what is the story told by most sources, who differs most from other sources and what are the opinions of the sources reporting in the events.

3. System design

To obtain the above objectives, we first defined the Grounded Annotation Framework (GAF[11], (Fokkens et al., 2014)) that makes a distinction between the mentions of events and participants in text and the instances that they refer to. Whereas, mentions are text-bound representations, instances abstract from these mentions and are identified through unique URIs. For the representation of mentions in text and other output of natural language processing (NLP), we defined the NLP Annotation Format (NAF[12], (Fokkens et al., 2013)). NAF is a sequel of the KYOTO Annotation Framework (KAF, (Bosma et al., 2009)) and is compliant to the Linguistic Annotation Format (LAF, (Ide et al., 2003)). It is a standoff layered representation for the results of a whole series of NLP modules ranging from tokenization, part-of-speech tagging, lemmatization, dependency parsing, named-entity recognition, semantic role labeling, event and entity-coreference to factuality and opinions. NAF is a document based presentation.

From NAF representations of news, we derive instances of events bound in time and place, involving instances of entities. To represent these instances, we use the Simple Event Model (SEM[13], (van Hage et al., 2011)), which is an RDF compliant model for representing events. SEM distinguishes events, actors, places and time, as well as relations between events. GAF adds gaf:denotes and gaf:denotedBy links for pointing to the representations in NAF that provided the evidence for creating the SEM structure. We represent relations between events and participants, place and time as named graphs so that we can express provenance relations between sources and the statements made by these sources. These relations indicate who made what statement but also what the opinion is of the source on the event or towards their participants. The provenance relations are based on the PROV-O model[14].

To obtain the SEM structure from NAF files, we developed an aggregation module that applies cross-document coreference across all mentions in NAF to establish matches across events and entities, as well as places and time. Cross-document coreference results in a single instance representation in SEM for all matched mentions. To establish matches for event mentions, we first cluster news

[11] http://groundedannotationframework.org/
[12] http://wordpress.let.vupr.nl/naf/
[13] http://wordpress.let.vupr.nl/sem/
[14] http://www.w3.org/TR/prov-o/

published on the same day using topic classification and location and time reasoning. Within these clusters, we compare event descriptions. Semantically related events need to be grounded to the same place and time constraints and share a proportion of participants. Two *take-over* mentions at different time points and/or involving different participants cannot refer to the same *take-over* instance. Our measure also allows for loose matches. Events can be described at a high or low abstraction level, places can match across a meronymy axis and time can be very specific or vague. Likewise, participant descriptions can match along hyponymy and meronymy levels, e.g. it can be the Volkswagen Group taking over Porsche divisions or just Volkswagen taking over Porsche. In (Cybulska and Vossen, 2013), we describe the main algorithm for cross-document coreference in more detail.

Whenever we establish a co-reference relation, we aggregate all information on the instance from the different mentions. We maintain the most specific information and complementary information is cumulated. Likewise, different sources can complement each other but also contradict for information that is deemed non-essential for identity. One source may specify specific divisions that are taken-over, another source does not mention the divisions or different sources may disagree on what divisions are taken-over and even the same source may mention different divisions at different points in time. All these mentions still refer to the same global take-over event.

All the sources and results of the processing, represented as NAF and as SEM, are stored in a central KnowledgeStore (Cattoni et al., 2012). The KnowledgeStore has different components for the different type of data. In addition to the original sources, e.g. XML files provided by LexisNexis, pointers to resources, mentions of events and entities are stored in an Hbase and Hadoop platform together with a specification of the context in which entities and events are mentioned in sources. Ultimately, mentions in sources are mapped to relations between instances of events and entities, represented as RDF triples stored in a separate triple store. The systematic separation of event/entity mentions and event/entity instances follows the formal model for semantic interpretation defined in the GAF.

The overall architecture for the NewsReader platform is given in Figure 3. It shows a range of NLP modules deployed around the central KnowledgeStore. The modules use input and output text stream APIs to NAF representations of the text. Whereas most modules for NLP interact with each other or with the Hbase/Hadoop part of the KnowledgeStore, the final modules in the chain need to access the RDF data to compare new event descriptions with past event description, either to establish coreference or to create event sequences.

4. Use cases

Use cases are an integral part of the project; they help us to create circumstances that match a decision maker as closely as possible, and retrieve data sets that have high relevance to the topic they support. Multiple use cases have been discussed and evaluated; four of those have passed our criteria and will be used for evaluations and demonstration. Before working on the actual use cases we started with the definition of a set of criteria, with which all use cases should comply:

1. The key topic(s) is/are in the fields of finance/economics/policy making, thus offering optimal relevance to professional decision making.
2. Availability of large data sets, primarily unstructured, but ideally structured as well, to be used as reference. See the 1st use case below for a structured/unstructured example.
3. Data coverage in at least one of the project languages: English, Italian, Spanish, Dutch.

The above criteria helped us to define the following 4 use cases to process data and carry out user evaluations during the project:

1. TechCrunch/Crunchbase: this consists of two information sources, a well-structured wiki-like database (Crunchbase[15]) and a web site publishing news articles (TechCrunch[16]), both covering an identical topic: information technology. We anticipate that events in the (structured) Crunchbase data will be reflected in the (non-structured) TechCrunch articles.
2. Global Automotive Industries: we created a large (¿6 million documents) and multilingual (English, Italian, Spanish, Dutch) data set that contains news documents on all key players in the automotive industry. This will be used to help News-Reader (re)construct:
   - complex structures, like the ownership structures of automotive conglomerates and
   - complex events, like mergers, acquisitions and corporate restructuring in this industry but also labour conflicts.
3. Business Intelligence: Due Diligence is an umbrella concept, used to describe the gathering of information about companies to evaluate them as potential business partners or customers. This evaluation can be used to ascertain a business ability to repay a loan, to comply with anti-money laundering legislation or other due diligence investigations. We will use information scraped from company web sites by ScraperWiki to test whether NewsReader could be used to support this due diligence process.
4. The Dutch House of Representatives: focusing on the information-intensive Parliamentary Inquiries, we identified several challenges. These are:
   - Coverage: understanding an event, its key actors and entities
   - Mapping the gaps: identifying areas with insufficient information coverage

[15] www.crunchbase.com
[16] techcrunch.com/

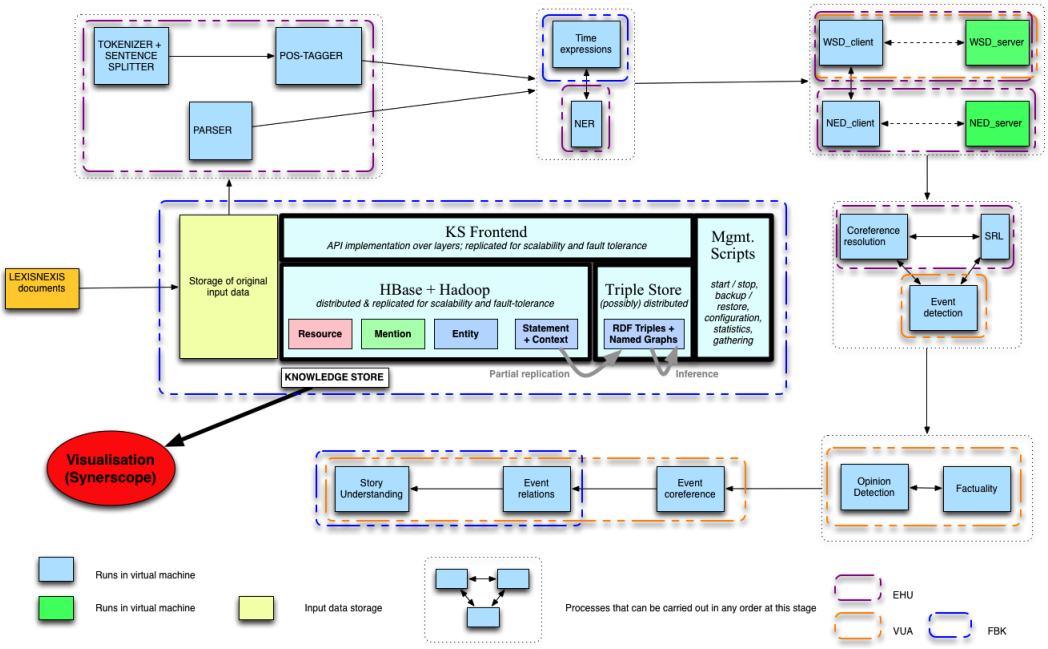

Figure 3: NewsReader system architecture

- Creating networks of events, people and entities (companies, government bodies) and
- Fact checking.

Use cases 1-3 were defined by the various consortium members; the fourth use case was developed in close collaboration with the Dutch Parliaments Information Department. We organized a workshop, interactively going through process mapping with the Parliament's employees to illustrate all information touch points in the process chain of a Parliamentary Inquiry.

5. Processed data

In the first year of the project we processed the following data sets for English news:

- Car Industry news (2003-2013): 63K articles, 1,7M event instances, 445K actors, 62K places, 41K DBpedia entities and 46M triples.
- TechCrunch (2005-2013): 43K articles, 1,6M event instances, 300K actors, 28K DBpedia entities and 24M triples. TechCrunch denoted by links 346K triples, CrunchBase event-relations13M triples.
- DBPedia[17]: 1.8M persons, 467K organisations, 742K places, 152K events, 270M triples.
- WikiNews: 19K English, 8K Italian, 7K Spanish and 1K Dutch. 69 Apple news documents for annotation.
- ECB+: 43 topics and 482 articles from GoogleNews, extended with 502 GoogleNews articles for 43+ topics (similar but different event).

[17] dbpedia.org

The English data sets have been processed using virtual machines (VMs[18]) that pack 13 modules shown in Figure 3. The result is a set of NAF files with various layers of analysis of the text and SEM-TriG files representing instances of events and entities, named-graphs for relations beween events and entities. In addition, we provide provenance and factuality statements on events and the named graphs of their relations. In Figure 4, we show an example of two entities from the entity layer and a predicate from the semantic role layer in NAF. The entities have a type attribute, pointers to the mentions in the text as span elements and a link to DBPedia. The predicate represents an event, typed through PropBank (Bonial et al., 2010), VerbNet (Kipper et al., 2006) and FrameNet (Baker et al., 1998), with a series or roles pointing to expressions in the text through span elements. The roles have PropBank role attributes and VerbNet role references.

In Figure 5, we show a TriG representation for event and entity instances, as well as the event relation, the provenance relation and the factuality. The event instance has a unique identifier, typing coming from the predicate layers, labels used to refer to it and GAF links to the mentions in a NAF file. For the entity, we use a DBPedia URI as the identifier, typing based on the entity category and the role labels from the predicates in which they occur and again GAF links to the mentions. The relation between the entity and the event also got a unique identifier. This identifier is used to express the provenance relation: what source made this statement (Peru_Auto_Report) as well the factuality statement (+CT = certain and factual).

[18] The English VM can be downloaded from the project website. The website also provides online demos to process English text in UTF-8 format through the VM and convert the result to the SEM-TriG format.

Figure 4: Example of the semantic role and entity layers in NAF

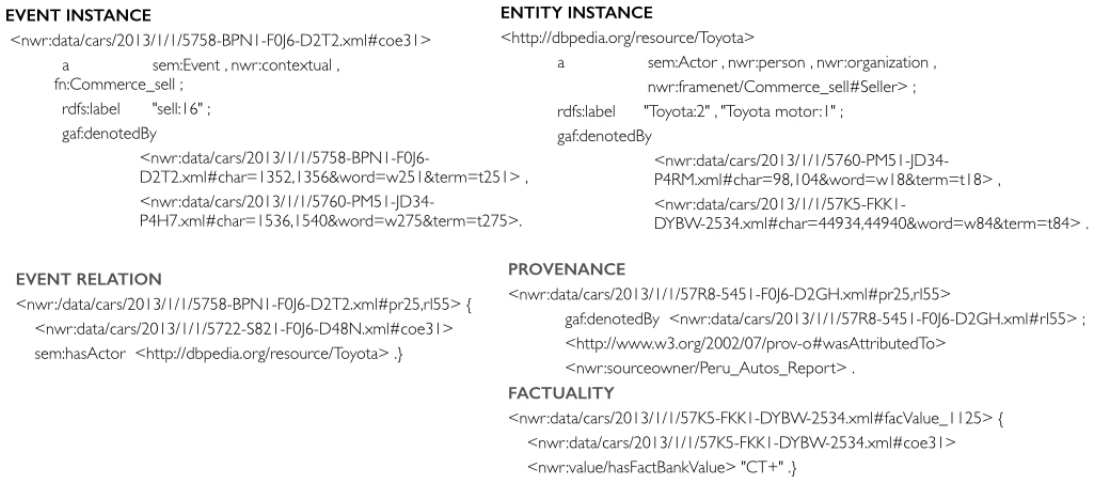

Figure 5: Example of representations of event and entity instances, event relations, provenance and factuality in TriG

Each TriG file represents the aggregation result across a cluster of NAF files, based on the publication date of the news and/or the topic. Aggregation implies that event mentions are coreferential across NAF files. We currently use a baseline system that matches all event descriptions with the same lemma in a cluster. For entities, we use the DBPedia URI to match entities across all NAF files, also extending the clusters. Future versions will use a more elaborate and robust coreference resolution.

The Car Industry data set and the DBPedia background knowledge have been loaded in the KnowledgeStore and can be accessed through the KnowledgeStore website[19]. The other data sets will be loaded soon. The Car Industry and TechCrunch data are used for user evaluations. The original sources can be obtained from LexisNexis and TechCrunch respectively. The processed data are available in SEM-TriG format from the project's website. The WikiNews and ECB+ data sets are manually annotated and will be used to benchmark the technology in the second year of the project. WikiNews is a dump[20] from the public news site from August 2013. From this dump, we took a selection of 69 English articles for annotation according to guidelines based on TimeML (Pustejovsky et al., 2010) and ACE (LDC, 2005). ECB+ is an extension of the Event Coreference Bank (ECB, (Bejan and Harabagiu, 2010). ECB is manually annotated to test cross-document coreference for events and, later also for entities (Lee et al., 2012). Whereas ECB consists of 43 topics and 482 text, the extension ECB+ has 43 additional topics with 502 articles on similar but different events. ECB+ thus has more referential ambiguity for events, see (Cybulska and Vossen, 2014) for more details. Both the WikiNews corpus and the ECB+ corpus are publicly available through the project's website.

We did a statistical analysis of the car data set: 63,810 NAF files for a period of 10 years merged and aggregated per publication date. This resulted in 1.7M event instances with 4.2M mentions in the text. On average events are mentioned 2.94 times in 2.1 sources per day. We also got 445K actor instances with 7.02 mentions in 2.36 sources and 63K place instances with 16.86 mentions in 7.64 sources on average per day. We thus have a stronger reduction from mentions to instances for actors and places than for

[19]https://knowledgestore.fbk.eu
[20]http://dumps.wikimedia.org

events. This is a logical consequence of the fact that actors and places can be matched through URIs in DBPedia and events only through lemma. If we only consider the actors and places that have been found in DBPedia, we get a further reduction to 126K entities (places and persons) with 21.42 mentions, across 8.48 different sources on average. The DBPedia references more or less represent the upper bound of grounding textual mentions to an external resource. The lemma matching within clusters for events represents a lower bound for identifying instances. We expect that future versions will show that the number of event instances will decrease and the average mention ratio will increase, if we improve event-coreference. Event-coreference will move towards entity-coreference although we probably never reach the ratio of the DBPedia mappings.

Actors and places are primarily identified through the semantic role they have in relation to a predicate that denotes an event. Propbank (Bonial et al., 2010) roles such as A0, A1 and A2 are used to define an actor in relation to an event and the role LOC results in a place. This classification is thus independent of the DBPedia URI that may be associated with the same expression. Likewise, we see that countries in the data set mostly have the role of a location but sometimes also have the role of an actor. In some cases this is correct but it can also point to potential errors in the interpretation. We see the same for persons and organizations in DBPedia that are dominantly associated with actor roles but sometimes also as places. Filtering the interpretations of events using background knowledge thus provides a potential powerful method to clean up the data. If we do not filter the participants against a background resource, 48.5% of the participants are persons and 45.3% are organizations according to the semantic roles they take. If we only accepts persons and organizations that are also entries in schema.org most of the participants (89%) are now not known.

6. Visualization and interaction on rich and complex knowledge graphs

The RDF graph that describes the story lines extracted from the news and background knowledge sources might be a simplified, summarized, and unified version of the original texts, but it is still a resource of super human magnitude and complexity. To bridge the gap between the KnowledgeStore holding the RDF storylines and human users, the NewsReader project has a dedicated decision support tool suite (DSTS) illustrated in Figure 6, which is based on SynerScope's interactive visual analytics software. The main goal of the SynerScope tools is to show as much of the data as possible to the user, and to let the user explore and understand the data through fast responsive interaction. This is made possible through visualization techniques that are specifically designed for complex network data (Holten, 2006) with a temporal dimension (van den Elzen et al., 2013), and that run on graphically accelerated systems. The DSTS consists of a large number of coordinated views that provide different perspectives on the same data, such as networks, sequences, scatter plots, bar charts, tables, search, Web browsing, etc. For the purpose of research on storyline visualization and rapid prototyping of additional views the DSTS includes a Web-based plugin system. The current SynerScope based DSTS allows interactive exploration of changing networks consisting of around 500,000 edges of a handful of different types and nodes that can be hierarchically classified based on any attribute. Although the scalability of the DSTS is about three orders of magnitudes greater than that of alternative visualization tools there still remains a large gap between the size of the RDF graph in the KnowledgeStore and the size of the graph that can be dealt with by the DSTS. This gap is closed by means of a graphically assisted data importer that allows the user to interactively limit the perspective and scope of the data to limit the data that is shown.

7. Conclusions and future work

In this paper, we outlined the main objective of the NewsReader project, the design and methodology adopted and the results of processed massive amounts of news in the first year of the project. The project aims at modeling dynamic news streams that report on the changes in the world in complex ways. We seek to add functionality to current solutions that do not address the dynamic sequencing of events and the way news reports on these changes. We designed and implemented a complex platform for processing large volumes of news in different languages and storing the result in a KnowledgeStore that supports the dynamic growth and reasoning over the data. We also discussed the use cases that we have defined for evaluation and the visualization and interaction of users on the large data sets that produced. The project shows an effective marriage between NLP and SemanticWeb, enabling us to develop reasoning technologies on top of the data that is generated from raw text. We started the analysis of the data but proper benchmarking and evaluations are scheduled for the next year in the project. In the future, we will improve the current modules through benchmarking and process more data. We will also implement cross-lingual event extraction representing the results in a language-neutral format. Current processing is fully generic. In the next year, we will incorporate background knowledge and modeling in the processing and use the reasoning capacity of the KnowledgeStore.

8. Acknowledgements

This work has been supported by the EC within the 7th framework programme under grant agreement nr. FP7-IST-316040.

9. References

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The berkeley framenet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 1*, ACL '98, pages 86–90, Stroudsburg, PA, USA. Association for Computational Linguistics.

Bejan, C. and Harabagiu, S. (2010). Unsupervised event coreference resolution with rich linguistic features. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1412–1422,

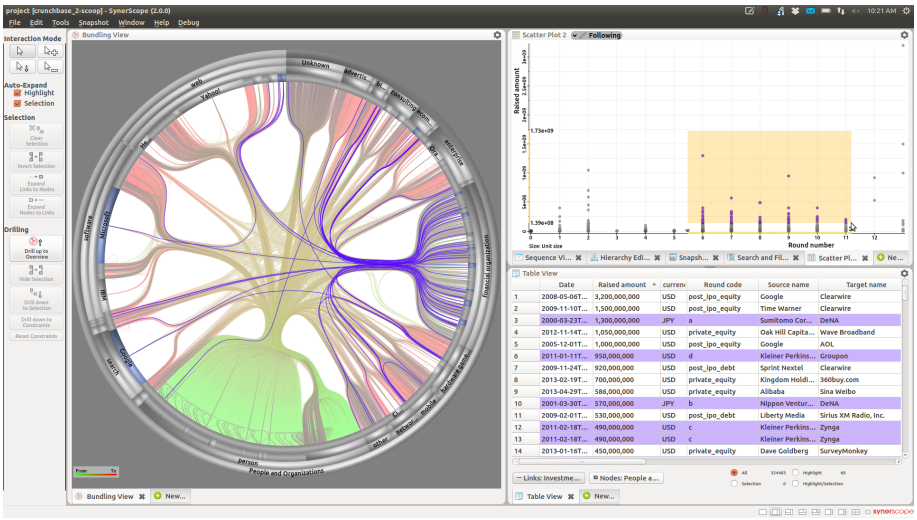

Figure 6: Multiple and coordinated views showing a hierarchical edge bundling view of a dynamic network of investments, a scatter plot showing magnitude of the investment per round, and a table view with additional attributes.

Uppsala, Sweden, July. Association for Computational Linguistics.

Bonial, C., Babko-Malaya, O., Choi, J. D., Hwang, J., and Palmer, M. (2010). Propbank annotation guidelines, version 3.0. Technical report, Center for Computational Language and Education Research, Institute of Cognitive Science, University of Colorado at Boulder, Pisa, Italy. `http://clear.colorado.edu/compsem/documents/propbank_guidelines.pdf`.

Bosma, W., Vossen, P., Soroa, A., Rigau, G., Tesconi, M., Marchetti, A., Monachini, M., and Aliprandi, C. (2009). KAF: a generic semantic annotation format. In *Proceedings of the 5th International Conference on Generative Approaches to the Lexicon GL 2009*, Pisa, Italy.

Cattoni, R., Corcoglioniti, F., Girardi, C., Magnini, B., Serafini, L., and Zanoli, R. (2012). The knowledgestore: an entity-based storage system. In *Proc. of 8th Int. Conf. on Language Resources and Evaluation (LREC'12)*, LREC '12.

Cybulska, A. and Vossen, P. (2013). Semantic relations between events and their time, locations and participants for event coreference resolution. In *Proceedings of Recent Advances in Natural Language Processing (RANLP-2013)*, pages 156–163.

Cybulska, A. and Vossen, P. (2014). Using a sledgehammer to crack a nut? lexical diversity and event coreference resolution. In *Proceedings of the 9th Language Resources and Evaluation Conference (LREC2014)*, Reykjavik, Iceland, May 26-31.

Fokkens, A., van Erp, M., Vossen, P., Tonelli, S., van Hage, W. R., Serafini, L., Sprugnoli, R., and Hoeksema, J. (2013). GAF: A grounded annotation framework for events. In *Proceedings of the first Workshop on Events: Definition, Dectection, Coreference and Representation*, Atlanta, USA.

Fokkens, A., van Erp, M., Vossen, P., Tonelli, S., van Hage, W. R., Serafini, L., Sprugnoli, R., and Hoeksema, J. (2014). NAF: A grounded annotation framework for events. In *Proceedings of the LREC2014 Workshop ISO10*, Reykavik, Iceland.

Holten, D. (2006). Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):741–748.

Ide, N., Romary, L., and de La Clergerie, É. V. (2003). International standard for a linguistic annotation framework. In *Proceedings of the HLT-NAACL 2003 Workshop on Software Engineering and Architecture of Language Technology Systems (SEALTS)*. Association for Computational Linguistics.

Kipper, K., Korhonen, A., Ryant, N., and Palmer, M. (2006). Extending verbnet with novel verb classes. In *Fifth International Conference on Language Resources and Evaluation*.

LDC. (2005). Ace (automatic content extraction) english annotation guidelines for events ver. 5.4.3 2005.07.01. Technical report, Linguistic Data Consortium.

Lee, H., Recasens, M., Chang, A., Surdeanu, M., and Jurafsky, D. (2012). Joint entity and event coreference resolution across documents. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 489–500, Jeju Island, Korea, July. Association for Computational Linguistics.

Pustejovsky, J., Lee, K., Bunt, H., and Romary, L. (2010). Iso-timeml: An international standard for semantic annotation. In *LREC*.

van den Elzen, S., Holten, D., Blaas, J., and van Wijk, J. J. (2013). Reordering massive sequence views: Enabling temporal and structural analysis of dynamic networks. In *Visualization Symposium (PacificVis), 2013 IEEE Pacific*, pages 33–40. IEEE.

van Hage, W. R., Malaisé, V., Segers, R., Hollink, L., and Schreiber, G. (2011). Design and use of the Simple Event Model (SEM). *J. Web Sem.*, 9(2):128–136.