Towards a Big Data Platform for News Angles*

Marc Gallofré Ocaña, Lars Nyre, Andreas L. Opdahl,
Bjørnar Tessem, Christoph Trattner, and Csaba Veres

Dept. of Information Science and Media Studies, University of Bergen, Norway
{Marc.Gallofre,Lars.Nyre,Andreas.Opdahl,Bjornar.Tessem,
Christoph.Trattner,Csaba.Veres}@uib.no
http://www.uib.no/en/rg/ssis

Abstract Finding good *angles* on news events is a central journalistic and editorial skill. As news work becomes increasingly computer-assisted and big-data based, journalistic tools therefore need to become better able to support news angles too. This paper outlines a big-data platform that is able to suggest appropriate angles on news events to journalists. We first clarify and discuss the central characteristics of news angles. We then proceed to outline a big-data architecture that can propose news angles. Important areas for further work include: representing news angles formally; identifying interesting and unexpected angles on unfolding events; and designing a big-data architecture that works on a global scale.

Keywords: Big data · Journalistic tools · News · Semantic technologies.

1 Introduction

Journalistic work is becoming increasingly reliant on computers and the internet [20]. Miroshnichenko [25] argues strongly for *artificial intelligence (AI)* in journalism and points to four areas of impact: data mining, topic selection, commentary moderation, and news writing. Journalistic robots developed by commercial companies such as Narrative Science and Automated Insights can already generate news stories in areas like finance and sports automatically [18]. According to [25], Automated Insight's Wordsmith tool wrote and published 1.5 billion news stories in 2016 alone, possibly more than all the human journalists in the world combined.

These developments in AI are driven in part by the availability of *big and open data sources* that are relevant for journalism. For example, researchers have investigated how news events can be extracted from big-data sources such as Tweets [15] and other texts [13]. Maiden et al. [21] propose the INJECT tool to

* Supported by the Norwegian Research Council IKTPLUSS project 275872 *News Angler*, which is a collaboration with Wolftech AB, Bergen, Norway.

support journalistic creativity during the early phases of news work. Their tool suggests relevant news stories to trigger new ideas for story angles more quickly and efficiently, and it has been tested in Norwegian and German newspapers.

Researchers have also used *semantic technologies*, such as RDF and OWL [1], to make big and open data sources more readily available for journalistic purposes [32] — and for journalistic AI tools [9]. Fernandez et al. [6] propose an ontology for streamlining news production and distribution processes. Heravi et al. [12] advocate *social semantic journalism*, which uses natural-language processing (NLP) and semantic metadata together to detect news events from socially-generated big data, verify information and its sources, identify eyewitnesses, and contextualise news events and their coverage. Leban et al. [19] present a platform that collects news messages and lifts them into a semantic *knowledge graph* (in RDF) in order to detect and describe news events in real time. In collaboration with *Wolftech AB*, a software company that delivers newsroom systems to the international market, our research group has developed *News Hunter*, an architecture and proof-of-concept prototype that supports journalists and other news professionals by building and mining semantic knowledge graphs that represent news-related information [27,3] (see Section 3).

All newsworthy events have remarkable qualities whether or not journalists are aware of them. Certain events are so remarkable that no effort is needed to find the best angle, whereas other events need to be probed, explored, and criticised to identify an angle that will interest the readers (or listeners, viewers). Finding good angles on news events thus resembles topic selection [25], but also is technique for presenting news stories in interesting ways. The task has traditionally been the responsibility of professional journalists [30, p. 115] and is considered a journalistic "trick-of-the-trade". It is covered in most introductory textbooks, but appears to have been little theorised in the research literature.

As news work becomes increasingly computer-assisted and big-data based, journalistic tools must become better able to identify and propose suitable angles too. This paper therefore investigates whether and how our News Hunter architecture and tool can be evolved to handle *big data* and extended to provide support for *news angles*. We ask: *Which characteristics of news angles need to be captured and represented for them to be supported by journalistic tools?* and *What are the open research issues related to creating a big-data platform that supports news angles?* Our aim is not to automate, but to support: we want to aid journalists by detecting new events and by suggesting newsworthy angles on them, along with relevant background information.

To investigate these questions, the rest of the paper is organised as follows: Section 2 first clarifies and discusses the central characteristics of news angles and related terms. Section 3 proceeds to describe News Hunter, our evolving big-data architecture and tool for journalistic work. Finally, Section 4 concludes the paper by reviewing open research issues.

Table 1. Alternative angles on the same football event.

Event:	***Football team A beats team B 2–0 in city C on date D.***
Impact:	"Historically important team B is now relegated and on the brink of bankruptcy."
Influence:	"The results of team A correlate with civil unrest and domestic violence in their home town."
Conflict:	"Coach A publicly insults rival coach B!"
Conflict:	"Supporters of these two teams have been fighting in the past."
Recency:	"Join our feed for live results."
Actionability:	"Join our newspaper's campaign to get rid of coach B!"
Proximity:	"Goalkeeper B grew up down the street from our editorial office."
Milestone:	"38 minutes into this match, team B will be the first team ever in the series to play a 1000 minutes with no penalty against them."
Human interest:	"Left midfielder B plays in honour of his terminally ill sibling."

2 What is a News Angle?

Certain events are so remarkable that they are newsworthy in themselves. Other events need to be presented in a certain way to become interesting for its readers (listeners, viewers). Several decades ago already, Altheide [2] observed that reporters rely on "'angles,' or story lines, which give the specific events new meaning", to which Shoemaker and Reese [30, p. 115] add that "[a] predefined story 'angle,' [. . .] provides reporters a theme around which to build a story". They also mention "news values [that] distil what people find interesting and important to know about" [30, p. 106].

2.1 Definition

We define a *news angle* tentatively as *how a journalist or other news professional makes an event interesting for an audience.* As an example, Table 1 lists alternative angles on the same event: a football game (we will go on to analyse the *impact* angle in more detail below). In addition to gaining the audience's attention, a news angle such as these serves several additional purposes:

- it provides a *criterion* for selecting events that are worth reporting;
- it points towards *additional facts* to report;
- it suggests which *information sources* to use; and
- it can serve as a *template* for how to present the event.

We focus more on the first three than on the fourth. Using basic concepts from literary theory [10], we focus more on *what is told* (the *fabula*) than *how it is told* (the *discourse*), which together form a *narrative.* Hence, finding an angle on an event is a creative but fact-based task. It takes as input a limited factual description of an event and produces as output a richer description that contains additional facts that are related to and augment the event and that connects the core facts to the interests of the audience.

Of course, there is no such thing as neutral factual content. Journalists and editors continuously choose which events to report, how visible to make them, who to interview, which other data sources to use, and how to word the final story — a phenomenon often referred to as *news framing* or *slanting*. Even seemingly objective big data collected by surveillance cameras or other sensors are, in the end, products of human choices of whether and where to place the cameras and sensors and of how to analyse and disseminate the captured data [16]. Yet computer-assisted journalism may in the future serve to limit — or offer alternatives too — human framing and slanting of the news.

2.2 Example: The impact angle

Several researchers, such as [30], have listed common news angles used by journalists. Additional lists have been provided by practitioners [33,29]. In future work, we want to synthesise these and other reviews into a taxonomy of news angles. As an example, this paper will discuss one of them, *impact*, in a little more detail according to: how it is described in the literature, its most common subtypes, its indicators, the data sources available to assess the indicators, and whether and how the angle amplifies and/or is amplified by other angles. The purpose is to better understand the requirements for a big-data platform that can support this and other angles.

Description The literature describes the impact angle in various ways (of which some are perhaps angle subtypes):

- Prominence: "The importance of a story is measured in its impact: how many lives it affects. Fatalities are more important than property damage." [30]
- Disaster: "Describes the impact of negative situations (and usually either what brought them about, how it's affecting the new subject, or what's being done about it)." [33]
- An incident: "Anything that goes wrong has the potential to become newsworthy, such as an industrial explosion, car crash or school shooting." [29]

Types of impact Events can be impactful in several ways, including: loss of life, physical injury, mental distress, damage to the environment, loss of property, and damage to property, including public infrastructure. Impact can thus be subdivided accordingly into: human impact, environmental impact, damage to property, etc. We envisage a big-data architecture where specialised agents for each subtype (or subsubtype etc.) continuously crawl a knowledge (RDF) graph in search of impactful events that can trigger a variant of the impact angle. Other agents can search for indicators of other angles, such as groups of reports that describe the same event with very different sentiments (potentially a subtype of *conflict*) or events that are related to an influential person.

Indicators and data sources Indicators of *human impact* are: loss of life, physical injuries, and mental distress, which can be gleaned from analysis of small as well as big data sets. Loss of life and physical injuries can be lifted in real-time from the official social-media feeds and online hospital logs if they are available. Otherwise, they must be synthesised from other news reports or, using triangulation, from less trusted social-media sources. Mental distress in an area can also be identified through large-scale sentiment analysis of social-media messages.

Damaged infrastructure can be indicated by and triangulated from a range of sources, such as surveillance cameras and other sensors, citizen reports on social media, messages from public authorities, deviating arrival times of and timetable changes for public transport. *Environmental impact* and *damage to property* can be derived from many of the same sources.

Estimating past impacts from archival materials can be much easier, as authorities and open data sources maintain statistics, for example, of accidents and disasters by type and various measures of impact.

For impact types such as these to be identifiable by the agents that operate on the knowledge graph, the represented events must be continuously *enriched* with additional types of information both from small-data sources like public authorities, trusted news sources, and official social media accounts and from big-data sources like social media and the Internet of Things (IoT).

Interactions High-impact events are newsworthy in themselves, and the core facts established by the agents can be presented to the audience more or less as is. Lower-impact events can also turn out to be interesting: either because there are (potential) *secondary consequences*, such as a limited avalanche blocking a train line during the holiday season, or because they are *amplified* through interaction with other angles, such as proximity or influential people: a minor flood in a residential area can become global news if it fills the basement of a celebrity's home.

2.3 Audience and genre expectations

A news angle is (almost always) relative to an audience: in case of the influence angle, different audiences may have widely different views of which people are famous and, to a lesser extent, powerful. News angles rely on the type of events that interest the intended audience [23]. For example, the angles and topics that interest people who read a local newspaper context are quite different from those of the international news section of BBC World. Indeed, analyses of media users in order to better understand their preferences and habits is itself a big-data analysis problem.

News angles are also influenced by the general characteristics of the news market. Traditional journalism is undergoing an economic crisis due to online news competition, and many newsrooms have had to trim their staff while producing more news than ever [31]. This leads to variations of copy-paste journalism and click baits. Adjustments have also been made to adapt to the online news market [5,11]. Higher-level journalistic tools that support news angles is a promising way of improving both quality and productiveness in a time of crisis and hard competition.

For each angle type and indicator, newsworthiness criteria can be established, taking into account: the market addressed by the newspaper, the characteristics of the audience, and the genre of the given news story. Optimising newsworthiness criteria for different media forms and genres is an empirical problem that can potentially be answered with big-data analytics, comparing factual descriptions of past events with news criteria most prevalent in the audience of the corresponding news reports. Multiple angles on the same event can be possible. Sometimes, only the best one should be chosen; other times, two or more of them could be combined to suggest a better story or to reach different niches of readers (listeners, viewers).

3 Towards a Big-Data Architecture for News Angles

News Hunter is an evolving architecture and proof-of-concept prototype for supporting journalistic work, which has been developed by our research group in collaboration with Wolftech AB, a supplier of newsroom software systems for the international market [27,3]. It has been designed to continually harvest news items and social media messages from the web; analyse and represent them semantically in a knowledge graph; classify, cluster, and label them; enrich them with additional information from encyclopedic and other reference sources; and present them in real time to journalists — either as tips about new events or as background material for stories they are already working on.

3.1 Current News Hunter architecture

The current version of News Hunter comprises the following components:

- *Harvesters* continuously download news texts and other relevant data items, such as social-media messages, from the web.
- *Uploaders* load harvested data into the appropriate database.
- The *TextDB* stores textual data items such as news stories and social media messages in raw form.
- A (currently online) *Translator* translates other-language texts into the canonical language, which is currently English.
- The *GraphDB* represent harvested data items semantically as knowledge graphs in RDF format. (In the GraphDB, each data item is also known as an event, although they may not be important enough to be called news events.)
- The *Lifter* represents each text in the TextDB as a knowledge graph in the GraphDB with the coordinated aid of several more specific analysers: *concept extractors* identify the central keywords in the text and disambiguate their meaning. *Topic analysers* identify the central topics the text is about,

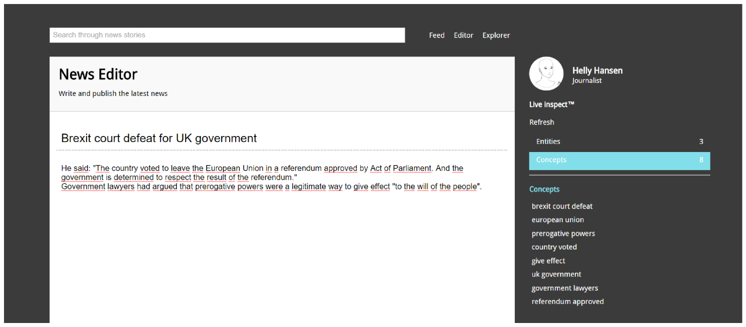

Figure 1. The front-end of the current News Hunter prototype. Relevant (named) entities and concepts that are extracted from the editor panel on the left-side are listed in the right-side panel, in which clicking an entity or concept returns a list of stories and other texts that are related to the one being typed into the editor.

independent of the keywords that are used. *Named-entity analysers* identify the own names of individuals, such as the people, organisations, and places that are mentioned. *Sentiment analysers* identify the positive and negative emotions in the text and its various phrases. *Categorisers* (or *Labellers*) assess how well the text fits predefined taxonomies, such as the IPTC News Codes [14].

- An *Event detector* identifies bursts and other changes in the occurrence frequencies of concepts, topics, named entities, and sentiments in a geographical or social region.
- A (currently limited) *Enricher* extends the core knowledge graphs produced by the *Lifter* with additional semantic reference data retrieved from the LOD cloud and from proprietary sources.
- A *Social networker* performs basic social-network analyses on the graph (currently limited to focussing on affinities).
- The *Editor* lets journalists write up new stories, which the Lifter continuously analyses semantically.
- A (currently limited) *Retriever* uses the semantic analyses of the new story to identify relevant background information in the GraphDB and retrieve related stories and other texts from the TextDB.
- The *Front end* (Figure 1) contains the editor and presents relevant background information and related stories and other texts to journalists and other news professionals.

3.2 Current News Hunter technologies

The current prototype [26,4] is mainly written in *Python* and *C#* as an *ASP.NET* application. Its components are interconnected through REST APIs [7] in a Flask-based micro-service architecture.

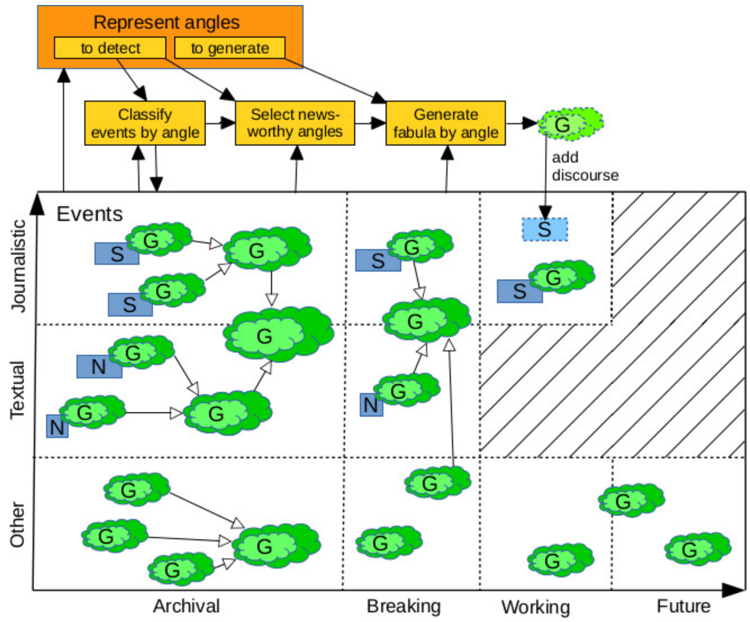

Figure 2. Extending News Hunter to support news angles.

The *Harvester* component uses a *Python* script to collect news-related texts (also called data items) from a variety of sources, such as Facebook, RSS, and online newspapers. Current focus is on downloading and parsing *RSS* feeds into JSON files using *Feedparser*. The JSON files are then stored in raw form in an *Elasticsearch TextDB* and, if necessary, in English using *Microsoft's Translate API*. The (English-language) JSON files are sent to a *C# .NET* pipeline that analyses the texts semantically in terms of: concepts, topics, named entities, sentiments, and categories. The lifted data is then stored in a *BrightstarDB GraphDB* (or *triple store*), which can be queried using SPARQL though a *Microsoft LINQ .NET* component. The news-related texts (data items) represented in the knowledge graph are also clustered in order to detect new events.

The different analysers use a variety of tools and techniques such as *TextRank*, *TF-IDF*, *SVM* (support vector machines), *MLP* (multi-layer perceptron), *RAKE* (Rapid Automatic Keyword Extraction), and *DBSCAN* clustering, among others. Both *SVM* and *MLP* were implemented using the *Keras* Python library. *DBSCAN* clustering was implemented using the *Scikit-learn* Python library. *TF-IDFs* were calculated with the *Textacy* and *Spacy* Python libraries. Sentiment analysis was done using the *AFINN* Python library.

The front-end application was written in HTML and CSS combined with *AngularJS*, and it was prototyped using *Sketch* and *Marvel*. *Froala* was used as text editor. An overview of News Hunter is presented in [3]. For further details on the tools, techniques, configurations, and evaluations we have used, see [26,4].

3.3 Leveraging big data for news angles

The current prototype does not yet scale to big data and does not support news angles. We are therefore evolving News Hunter into a big-data architecture that can be extended with components that support news angles.

The new architecture must be *open-ended* in two ways: It must allow user organisations to interface with their existing back-end tools, such as existing semantic and other analysis services and existing text and graph databases. It must also allow them to use open and proprietary data sources in combination. For open data, News Hunter may provide storage and analysis as a service. But news organisations may also want to use News Hunter to store and analyse proprietary data, either self-produced or licensed. The architecture should therefore be able to combine cloud and local data storage and analysis as seamlessly as possible.

The (orange/yellow) graph on top of Figure 2 illustrates some of the high-level reasoning steps needed to support news angles, whereas the (blue/green) grid at the bottom shows the different types of information that must be dealt with. Along the vertical axis in Figure 2, information can be either: *Journalistic*, meaning that it is text written by professional journalists. *Textual*, meaning that it is textual, but not written by professional journalists. *Other*, meaning that is is non-textual, which currently means that it is represented as knowledge graphs represented in RDF. Of course, future versions of our architecture may also cover other information types than texts and knowledge graphs, such as other types of structured data, along with images, audio, and video, introducing additional rows in Figure 2 and requiring additional analysis and storage techniques such as speech-to-text conversion, image/video analysis, and other types of databases.

Along the horizontal time axis, the information can be either: *Archival*, representing past events. *Breaking*, representing currently unfolding events. *Working*, representing not-yet-reported events. *Future*, representing anticipated, predicted, scheduled, or recurring events.

Each event is represented as a small *event graph*, represented as RDF, with both a *core* that describes the event directly and an *extension* that provides context. The central resource (or node) in the core graph represents the event itself, with related resources that result from *lifting* event data to semantic form. For example, journalistic stories and other texts can be lifted using techniques such as concept extraction, topic identification, named-entity recognition (of people, places, organisations, works, etc.), and sentiment analysis. The extension graph results from *enriching* the core graph with additional information in RDF format, for example from open semantic reference data sets in the Linked Open Data (LOD) cloud — such as DBpedia, Wikidata, GeoNames, and LinkedGeoData — or from proprietary data sources that have been lifted to semantic format.

Event graphs will usually overlap to a large extent because they include the same RDF resources for: people, places, and organisations; concepts, topics, and categories; RDF types; etc. Figure 2 therefore indicates that multiple overlapping (or otherwise similar) smaller graphs will be clustered and merged to form larger,

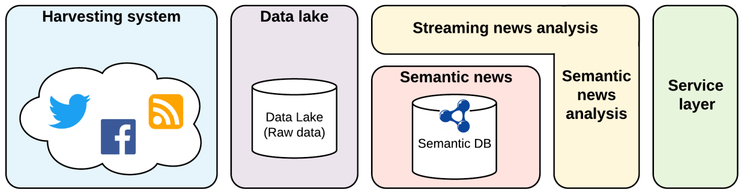

Figure 3. Overview of the new architecture for News Hunter.

more detailed, and reliable graphs. Exploiting overlapping and similar event graphs in this way is essential both for generating richer (more complete and detailed) event descriptions and to corroborate them. In particular for social media messages, unless the originator is known and trusted, triangulation of information from several independent sources is essential to ensure that only reliable event data are reported as news.

3.4 A lambda architecture that supports news angles

We are currently evolving News Hunter into a big-data architecture that can be extended with components that support news angles along the lines shown in Figure 3, on top of Apache's Kafka platform[1].

The new architecture is based on the *Lambda architecture* pattern [22], which is designed for service-oriented big-data processing and is able to analyse big data from social media sources with satisfactory performance [28]. An advantage of the Lambda architecture — as opposed to the alternative *Kappa architecture* [17] — is that it supports both real-time streaming analysis of all incoming data items and batch-oriented deeper analyses (and re-analyses) of selected data items that later turn out to be particularly interesting.

Harvesting system The new architecture is designed for continuously gathering news-related information from a variety of sources through the harvesting system, which is conceived as a message publishing and subscribing system. This pub-sub system lets News Hunter connect to a wide variety of external data sources, from social networks via commercial news services to the Internet of Things (IoT). It will filter and prioritise the incoming data streams and store the raw data in a data lake. Built on top of Apache Kafka[1], it provides a scalable and parallel messaging mechanism. The harvesting system will comprise the *Harvester* components in the current architecture and add a new *Filter* and/or *Prioritiser* component.

Data lake The data lake stores incoming data items in their raw form as Kafka *topics*, along with their English translations. It will comprise the *TextDB* and *Translator* components in the current architecture, and we will consider adding more powerful big-data storage technologies as the tool evolves.

[1] `https://kafka.apache.org/`

Knowledge graph (Semantic news) The knowledge graph contains semantic triples that are *lifted* from the data lake in real time as new data items arrive. Data lifting consists of concept extraction, topic identification, named-entity recognition, sentiment analysis, categorisation/labelling, and relation extraction. The graph will comprise the *GraphDB* and *Lifter* components in the current architecture, and we plan introducing a *Relation extractor* component.

News analysis The analysis layer analyses the lifted data items further, in real time and possibly as batch. *Streaming news analysis* in real time takes semantically-lifted data straight from lifting and is intended to provide journalists with both (1) real-time updates of the stories they are working on and (2) potentially newsworthy new events. *Semantic news analysis* in batch takes semantically-lifted data stored in the knowledge graph and is intended to (3) provide journalists with background information related to the stories they are working on and (4) organising and enriching the knowledge graph with data from other sources; performing social network analysis on the graph to identify super-nodes, sub-networks, affinities; and detecting clusters of overlapping or similar events.

The streaming news analysis will comprise the *Event detector*, *Enricher*, and *Social networker* components in the current architecture. In order to support news angles and other higher-level services to journalists, the analysis layer will also explore new components such as *Organisers* that continuously assess and improve the structure to the knowledge graph. *Analogy reasoners* that aim to identify other less obvious but semantically deeper connections between past and present events and stories and related background information. *Anglers* leverage the semantic analyses and background information to identify, generate, and rank candidate angles on potential news events that a journalist is already working on or that have been detected in the knowledge graph.

Service layer The service layer is in charge of making all the knowledge and analysis results available to newsrooms through a GUI and a REST API. It is this layer that makes the architecture and tool available to journalists and other news professionals. It will offer dashboards and user interfaces that present potentially relevant angles, stories, and other background information to journalists based on their current activities and preferences. The service will comprise the *Editor* and *Front end* components, and it will extend the *Retriever*.

4 Conclusion

The paper has discussed how journalistic tools can be improved by combining open and big data sources with the concept of *news angles*. This is an important research problem, because news work is becoming increasingly computer-assisted and big-data based, creating opportunities for a new generation of tools that provide even higher-level support for journalistic work. We have clarified and discussed the central characteristics of news angles and related terms and outlined how our architecture and tool for journalistic work, *News Hunter*, can be evolved and extended into a big-data platform that supports news angles.

Our work is part of a research project, *News Angler*, that is carried out in collaboration with Wolftech AB, a supplier of newsroom software systems for the international market. The News Angler project has two primary goals: (1) to improve and evolve the News Hunter architecture towards a big-data architecture that scales to the needs of international news organisations and (2) to extend News Hunter to support news angles. The project thus combines a moderate-ambition research and development goal (1) with a higher-ambition basic research goal (2). While eyeing the second, longer-term goal, the present paper also lays out concrete steps towards the first.

Our work on news angles is only beginning, and a long line of research and development issues remain. On the *architecture level* (1), are porting the News Hunter prototype to Linux on top of Apache's big-data stack, leveraging tools such as Kafka, Cassandra, and Spark. Most of the current components will have to be reengineering or reimplemented as part of this effort. Many of them should also be extended and improved in the process (in particular the *Enricher*, *Social networker*, and *Retriever*), and some new components introduced (for example the *Filter*, *Relation extractor*, *Organiser*, *Analogy reasoner*, and a locally-running *Translator*). In order to make our platform open-ended, we need to define clear-cut APIs between our components, (a) to make it easier for user organisations to interface with their existing back-end tools, (b) to make it easier for user organisations to use open and proprietary data sources in combination, and (c) to make it easier for ourselves to combine and compare alternative component implementations, such as multiple named-entity recognisers.

On the *news-angle level* (2), we want to develop new components that continuously analyse the input stream and knowledge graph for new events and newsworthy angles. One component type will manage prerequisites for angles, for example agents that identify approaching anniversaries; conflicting descriptions of an event; natural disasters and their impact; swift changes in popularity or political power; etc. Another component type will match angles to (breaking or historical) events, possibly combining multiple angles on the same event.

We think our work on the News Hunter prototype and its News Angler components can contribute to the wider research literature in several ways. *Relation extraction* is an research area that is central for our project. Currently, our event graphs tend to have a star-like structure, with a central RDF node representing the event itself and with lots of related annotation nodes that are connected to it, but less often to one another. Yet it is the relations between concepts, topics, named entities, and sentiments that describe an event most precisely. Gangemi et al. [8] have recently proposed FRED, a library and online tool for relation extraction that may be useful for our purposes. The emerging generation of neural-network base compositional vector representations of word meanings [24] may also offer new ways to extract relations from text. *Relating events by analogy* is another interesting research task, along with, e.g.: collecting and creating taxonomies of news angles; developing a user-friendly way of reading and writing news angles; identifying interesting and unexpected news angles; and fully exploiting open data, in particular linked open data.

Acknowledgement

Early development of News Hunter was supported by NCE (Norwegian Centre of Expertise) Media. News Angler is funded by the Norwegian Research Council's IKTPLUSS programme as project 275872. The authors are indebted to Arne Berven and Bjarte Djuvik Næss at Wolftech AB for fruitful discussions and to Kamal Alipour, Ole Andreas Christensen, Kjetil Jacobsen Villanger, and Sindre Moldeklev who made central contributions to the earlier versions of News Hunter.

References

1. Allemang, D., Hendler, J.: Semantic web for the working ontologist: Effective modeling in RDFS and OWL. Elsevier (2011)
2. Altheide, D.L., Rasmussen, P.K.: Becoming news: A study of two newsrooms. Sociology of Work and Occupations **3**(2), 223–246 (May 1976). https://doi.org/10.1177/073088847600300206, `http://journals.sagepub.com/doi/10.1177/073088847600300206`
3. Berven, A., Christensen, O.A., Moldeklev, S., Opdahl, A.L., Villanger, K.J.: News Hunter: Building and mining knowledge graphs for newsroom systems. NOKOBIT — Norsk konferanse for organisasjoners bruk av informasjonsteknologi **26** (2018)
4. Christensen, Ole Andreas, Villanger, Kjetil Jacobsen: News Hunter: A semantic news aggregator. Master's thesis, Univ. of Bergen (2017), `http://hdl.handle.net/1956/16192`
5. Ekdale, B., Singer, J.B., Tully, M., Harmsen, S.: Making change: Diffusion of technological, relational, and cultural innovation in the newsroom. Journalism & Mass Communication Quarterly **92**(4), 938–958 (2015)
6. Fernández, N., Fuentes, D., Sánchez, L., Fisteus, J.A.: The NEWS ontology: Design and applications. Expert Systems with Applications **37**(12), 8694–8704 (2010)
7. Fielding, R.T.: Architectural styles and the design of network-based software architectures. Ph.D. thesis, University of Californa, Irvine (2000), `https://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf`
8. Gangemi, A., Presutti, V., Reforgiato Recupero, D., Nuzzolese, A.G., Draicchio, F., Mongiovì, M.: Semantic web machine reading with FRED. Semantic Web **8**(6), 873–893 (2017)
9. García, R., Perdrix, F., Gil, R., Oliva, M.: The semantic web as a newspaper media convergence facilitator. Web Semantics: Science, Services and Agents on the World Wide Web **6**(2), 151–161 (2008)
10. Gervas, P.: Computational approaches to storytelling and creativity. AI Magazine **30**(3), 49–62 (Jul 2009). https://doi.org/10.1609/aimag.v30i3.2250, `https://aaai.org/ojs/index.php/aimagazine/article/view/2250`
11. Gynnild, A.: Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets. Journalism **15**(6), 713–730 (2014)
12. Heravi, B.R., McGinnis, J.: Introducing social semantic journalism. The Journal of Media Innovations **2**(1), 131–140 (2015)
13. Hogenboom, F., Frasincar, F., Kaymak, U., De Jong, F.: An overview of event extraction from text. In: Workshop on Detection, Representation, and Exploitation of Events in the Semantic Web (DeRiVE 2011) at Tenth International Semantic Web Conference (ISWC 2011). vol. 779, pp. 48–57. Citeseer (2011)

14. International Press Telecommunications Council: Newscodes — IPTC, https://iptc.org/standards/newscodes/
15. Jackoway, A., Samet, H., Sankaranarayanan, J.: Identification of live news events using Twitter. In: Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks. pp. 25–32. ACM (2011)
16. Kitchin, R.: The data revolution: Big data, open data, data infrastructures and their consequences. Sage (2014)
17. Kreps, J.: Questioning the lambda architecture, https://www.oreilly.com/ideas/questioning-the-lambda-architecture
18. Latar, N.L.: The robot journalist in the age of social physics: The end of human journalism? In: The new world of transitioned media, pp. 65–80. Springer (2015)
19. Leban, G., Fortuna, B., Brank, J., Grobelnik, M.: Event Registry: Learning about world events from news. In: Proceedings of the 23rd International Conference on World Wide Web. pp. 107–110. ACM (2014)
20. Machill, M., Beiler, M.: The importance of the internet for journalistic research: A multi-method study of the research performed by journalists working for daily newspapers, radio, television and online. Journalism Studies **10**(2), 178–203 (2009)
21. Maiden, N., Zachos, K., Brown, A., Brock, G., Nyre, L., Nygård Tonheim, A., Apsotolou, D., Evans, J.: Making the news: Digital creativity support for journalists. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. p. 475. ACM (2018)
22. Marz, N.: How to beat the CAP theorem, http://nathanmarz.com/blog/how-to-beat-the-cap-theorem.html
23. McNair, B.: The sociology of journalism. Arnold, London (1998)
24. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: Advances in neural information processing systems. pp. 3111–3119 (2013)
25. Miroshnichenko, A.: AI to bypass creativity. Will robots replace journalists? (The answer is "yes"). Information **9**(7) (Jul 2018). https://doi.org/10.3390/info9070183, http://www.mdpi.com/2078-2489/9/7/183
26. Moldeklev, S.: Improving usefulness and ease of use for a prototype tool for journalists. Master's thesis, University of Bergen (2018)
27. Opdahl, A.L., Berven, A., Alipour, K., Christensen, O.A., Villanger, K.J.: Knowledge graphs for newsroom systems. NOKOBIT—Norsk konferanse for organisasjoners bruk av informasjonsteknologi **24** (2016)
28. Persico, V., Pescapé, A., Picariello, A., Sperlí, G.: Benchmarking big data architectures for social networks data processing using public cloud platforms. Future Generation Computer Systems **89**, 98–109 (2018). https://doi.org/10.1016/j.future.2018.05.068, http://www.sciencedirect.com/science/article/pii/S0167739X17328303
29. Phillips, B.: 16 story angles that reporters relish. https://www.prdaily.com/Main/Articles/16_story_angles_that_reporters_relish_17748.aspx (2014)
30. Shoemaker, P.J., Reese, S.D.: Mediating the message: Theories of influences on mass media content (1995)
31. Sjøvaag, H.: Homogenisation or differentiation? The effects of consolidation in the regional newspaper market. Journalism Studies **15**(5), 511–521 (2014)
32. Troncy, R.: Bringing the IPTC news architecture into the semantic web. In: International Semantic Web Conference. pp. 483–498. Springer (2008)
33. Upchurch, W.: Ten common news angles for media releases. http://www.streetdirectory.com/etoday/ten-common-news-angles-for-media-releases-uuofou.html (2018)