

Automatische Analyse Sozialer Medien für das Krisenmanagement im Projekt Slándáil¹

Gerhard Heyer, Sabine Gründer-Fahrer, Antje Schlaf (Universität Leipzig / InfAI)

Alexander Lörch, Sebastian Wustmann (CID GmbH)

Prof. Gerhard Heyer
heyer@informatik.uni-leipzig.de
Universität Leipzig /
Institut für Informatik /
Abteilung Automatische
Sprachverarbeitung
Augustusplatz 10, 04109 Leipzig

Dr. Sabine Gründer-Fahrer
gruender@uni-leipzig.de
und
Antje Schlaf
antje.schlaf@informatik.uni-leipzig.de
Institut für Angewandte Informatik
(InfAI) e. V.
An-Institut an der Universität Leipzig
Harnstraße 11, 04109 Leipzig

Alexander Lörch
a.loerch@cid.de
und
Sebastian Wustmann
s.wustmann@cid.de
CID GmbH
Gewerbepark Birkenhain 1
63579 Freigericht

1. Soziale Medien im Krisenmanagement

Eine Krise hat aus Sicht des Krisenmanagements grundsätzlich drei Dimensionen: 1. das reale Ereignis, 2. das Handeln der betroffenen Organisationen und 3. die öffentliche Wahrnehmung der Krise (vgl. Bundesministerium des Inneren 2008 :10). Kommunikation spielt in jeder dieser Dimensionen eine tragende Rolle, weshalb hohe Anforderungen an eine effiziente Krisenkommunikation gestellt werden. In Bezug auf das reale Ereignis geht es dem Krisenmanager um die schnellstmögliche Verfügbarkeit von zuverlässiger und umfassender Information z.B. zur Ereignis- und Schadenslage. Bei der Organisation und Durchführung von Handlungsabläufen muss die effiziente, unmissverständliche und nachvollziehbare Verständigung zwischen Gesprächspartnern gewährleistet sein, die unter Umständen nicht nur lokal und institutionell, sondern auch sprachlich und kulturell voneinander entfernt sind. Was die öffentliche Wahrnehmung der Krise anbelangt, so eröffnen sich über Vorgänge der Vertrauens- und Meinungsbildung bis hin zum Emotionsmanagement soziale, politische oder psychologische Aspekte von Kommunikation.

Die Digitalisierung der Kommunikation hat eine Steigerung ihrer Qualität in allen drei der oben genannten Krisendimensionen mit sich gebracht, und das neu gewonnene Potential ist bis heute noch längst nicht ausgeschöpft (vgl. Kirchbach Kommission 2013). Eine besondere

¹ Conference Paper zur Internationalen Fachtagung „KommunikationsFluten“ des Europäischen Instituts für Journalismus- und Kommunikationsforschung e.V., 06.11.2105, Leipzig

Rolle innerhalb dieser Entwicklung spielen die sozialen Medien. Im Zuge ihres rasanten Aufstiegs „Emergency management, once the province of official channels, is going where the people are.“ (Rutrell Yasin 2010 :1) In allen drei Dimensionen einer Krise können soziale Medien in vielfacher Weise nutzbringend zum Einsatz kommen. Krisenmanager können Informationen an eine breitere Öffentlichkeit weitergeben. Direkte Interaktion und Kollaboration mit freiwilligen Helfergruppen wird ermöglicht. Durch die Einbeziehung von sozialen Medien als Informationsquelle kann der Krisenmanager die Situation unter Umständen schneller, aktueller, detaillierter und flächendeckender erfassen, als es mit Hilfe herkömmlicher Informationsquellen möglich war. Schließlich erlaubt es die Beobachtung der Kommunikation in den sozialen Medien dem Krisenmanager, sich einen Eindruck davon zu verschaffen, wie das Ereignis oder auch die Handlungen der zuständigen Organisationen in den Communities wahrgenommen und bewertet wird.

Indem sich das Medium des kommunikativen Austauschs verändert, verändert sich auch die Struktur und Kultur der Kommunikation selbst bis hin zu ihren Inhalten. Hieraus ergeben sich für den Krisenmanager Herausforderungen, die über die bloße Anpassung an technische Neuerungen weit hinausgehen. Vielmehr beinhalten sie Änderungen in den Organisationsformen und Umgangsregeln des Austauschs (vgl. Florian Semle 2015) und berühren sogar Fragen der Ethik und des Rechts, vor allem in den Bereichen Datenschutz und Schutz der Privatsphäre. So nutzbringend soziale Medien in allen drei Dimensionen eines Krisenereignisses sein können, so gefährlich kann ihr Einfluss werden, z.B. durch die Verbreitung falscher Informationen oder das Propagieren verzerrter Stimmungsbilder. Die Auswertung der Strukturen und Inhalte der sozialen Medien geschieht also aus Sicht des Krisenmanagements nicht nur mit dem Ziel des Informationsgewinns und -austauschs, sondern auch im Interesse einer verantwortlichen Kontrolle und möglichen Einflussnahme.

Für beide Interessenlagen besteht jedoch eine erste ernste Hürde: Die Digitalisierung der Kommunikation und die einfachen Publikationsmöglichkeiten in Sozialen Medien bringen eine Flut von Informationen hervor, welche in einem einzigen multimedialen, multimodalen, multilingualen und multikulturellen digitalen Raum zusammenfließen. Aufgrund der Menge, Vielfalt und Verschiedenheit der so vorhandenen Daten übersteigt der Versuch deren systematischer Auswertung bereits jetzt jede menschliche Möglichkeit. Will modernes Krisenmanagement die Chancen und Möglichkeiten digitaler und insbesondere sozialer Medien nutzen sowie deren kritisches Potential kontrollieren, so benötigt es für die Analyse dieser riesigen Informationsmengen zuallererst geeignete computerbasierte Werkzeuge.

2. Das Projekt Slándáil

Slándáil (Slandail 2014) ist ein Kooperationsprojekt zwischen 12 europäischen Partnern in Nordirland und Irland, Deutschland und Italien, gefördert im Rahmen des Framework 7 Security Programms der Europäischen Union über einen Zeitraum von 3 Jahren (Projektstart April 2014). Hauptziel des Projektes ist die Entwicklung und Erprobung eines Systems zur Nutzung von Informationen aus sozialen Medien im Katastrophenmanagement, insbesondere im Kontext von Naturkatastrophen. Besondere Merkmale von Slándáil sind die Sammlung und Verarbeitung von Daten in unterschiedlichen Modalitäten (Text und Bild), die Einbeziehung von sprach- und kulturübergreifenden Aspekten von Krisenkommunikation sowie die besondere Beachtung von Fragen der ethischen Korrektheit der Datennutzung. Angefangen von der Entwurfs- über die Entwicklungs- bis hin zur Testphase des Systems werden Katastrophenmanager selbst an wesentlichen Überlegungen und Entscheidungen mitwirken.

Im Projekt sind 4 Organisationen aus dem Bereich des Krisen- und Katastrophenmanagements involviert: *Police Service of Northern Ireland*, *Garda Siochána Southern Ireland*, *Protezione Civile Regione Veneto* und das *Bezirksverbindungskommando Leipzig*. Diese arbeiten zusammen mit kommerziellen und akademischen Experten für Kommunikation und Ethik: *Stillwater Communication Limited*, *University Degli Studi di Padua* und der *Irish School of Ecumenics at Trinity College Dublin*. Gemeinsam wird untersucht, wie Organisationen des Bevölkerungsschutzes und Krisenmanagements untereinander und mit der breiten Öffentlichkeit interagieren und kommunizieren – und wie dieser Austausch verbessert werden kann durch den gezielten Einsatz von sozialen Medien. In einem gesonderten Arbeitspaket werden ethische und rechtliche Richtlinien erarbeitet, welche die Einhaltung der Bestimmungen zur Sicherheit der Daten und Schutz der Privatsphäre garantieren. Die von diesen Arbeitsgruppen formulierten Anforderungen und gewonnenen Einsichten bilden die Grundlage für die Entwicklung eines Softwaresystems, welches die Möglichkeiten zur automatischen Verarbeitung von Text- und Bildmedien in eine Katastrophenmanagementumgebung integriert. Industrielle IT-Partner, deren jeweilige professionelle Systeme in den Bereichen Sprachverarbeitung und Katastrophenmanagement die Grundelemente des Slándáil-Systems bilden, sind die *CID GmbH* in Deutschland und *Data Piano S.r.l.* in Italien. Die detaillierte Weiterentwicklung und Anpassung dieser Grundkomponenten in Hinblick auf die Anforderungen der Benutzergruppe geschieht in

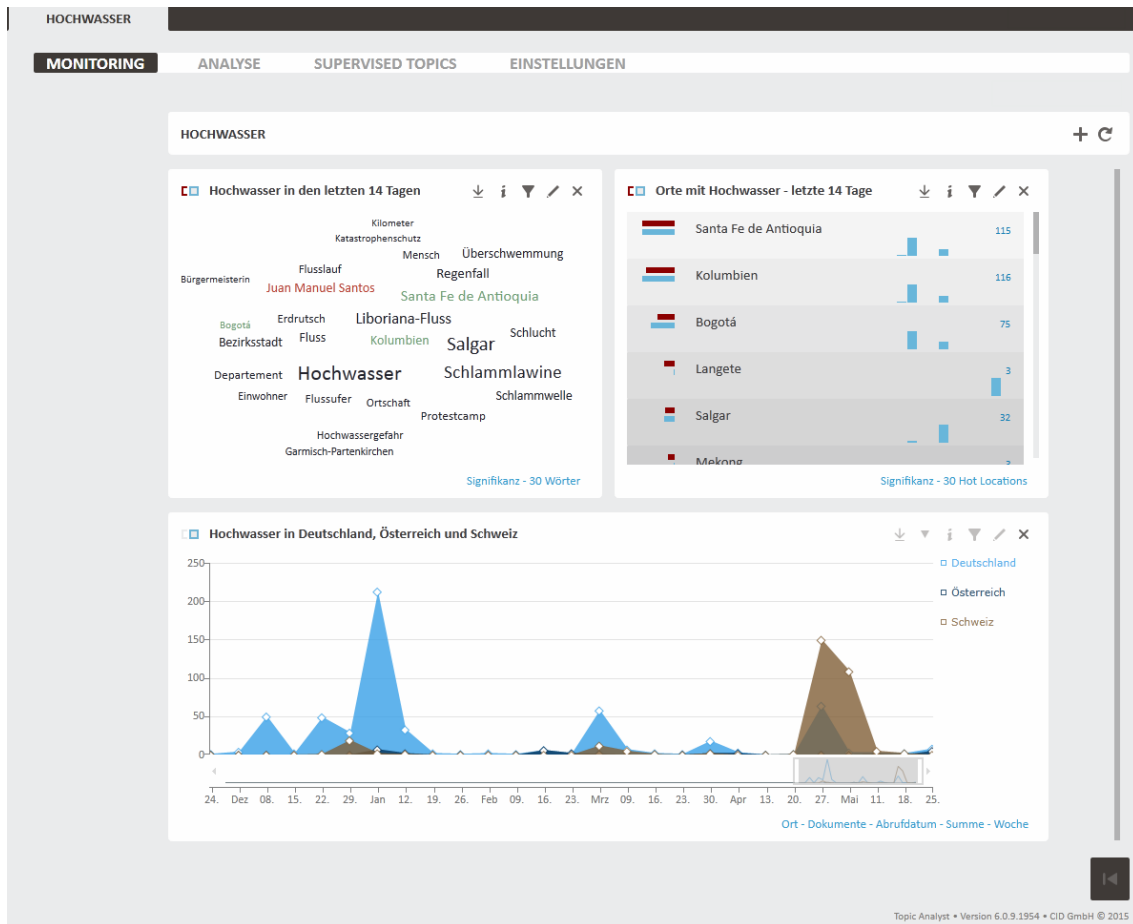
enger Kooperation mit akademischen Partnern aus den Bereichen der Automatischen Sprachverarbeitung für Englisch, Deutsch und Italienisch am *Trinity College Dublin* und der *Universität Leipzig / InfAI* sowie der Bildverarbeitung an der *University of Ulster*. Wie zu erkennen, beinhaltet die Arbeit im Projekt die besonderen Chancen und Herausforderungen einer sowohl fächer- als auch länderübergreifenden Kooperation; bei der Projektorganisation kommt deshalb professionelles Management von *Pintail Limited* in Dublin zum Einsatz.

Im folgenden Abschnitt soll das sprachverarbeitende Modul des Slándáil-Systems (STAM) in seiner ersten und vorläufigen Fassung kurz vorgestellt werden. Die Grundkomponenten und -funktionen der Software werden von der *CID GmbH* eingebracht, aktuelle und noch geplante Erweiterungen erfolgen durch die *Universität Leipzig / InfAI* und das *Trinity College Dublin*. Angesichts der Interessengruppe dieses Tagungspapers sowie aus Platzgründen wird auf detaillierte technische oder theoretische Ausführungen verzichtet und statt dessen eine intuitive Darstellung einiger Hauptfunktionalitäten des Prototypen¹ der Software erfolgen.

3. Automatische Sprachverarbeitung in Slándáil (Prototyp1)

Die Grundlage des sprachverarbeitenden Systems bildet eine große und dynamische Menge von Dokumenten, welche das System als kontinuierlichen Datenstrom in Echtzeit aus verschiedenen externen Quellen wie Websites, Feeds, sozialen Medien oder internen Datenbanken sammelt und speichert. Diese Datenkollektion wird anwendergerecht in verschiedenen Datenpools organisiert, aus denen zu Analyse Zwecken über ein Filterpanel flexible Unterkollektionen erzeugt werden können. Neben Filterkategorien inhaltlicher Art (Person, Ort, Organisation, Keyword, Topics, Tag) existieren Filterkategorien, die auf Metadaten basieren (Veröffentlichungs- bzw. Crawldatum, Geo-Position, Quellen, Quellentypen oder -domänen).

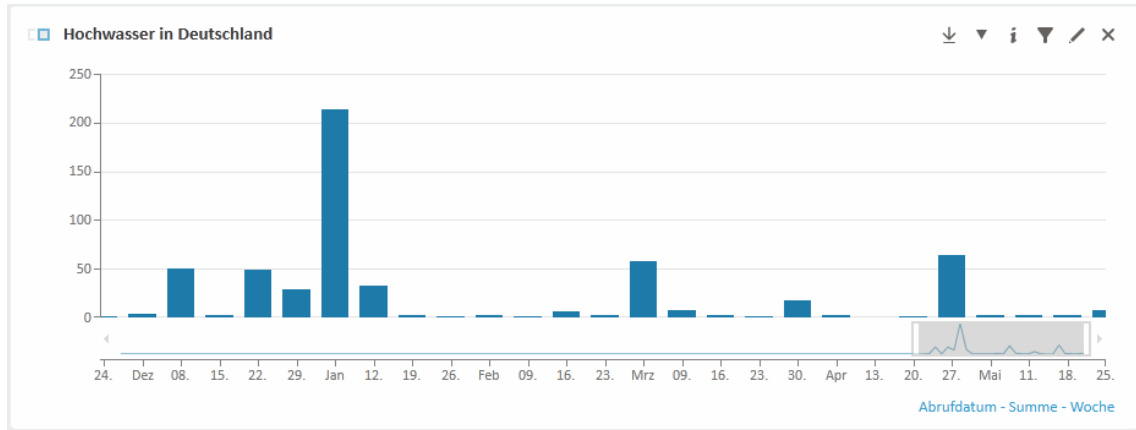
Zur Analyse einer Datenkollektion stehen daraufhin zwei Hauptmodule zur Verfügung: Monitoring (*Dashboard*) und Analyse (*Browser*). Während das Dashboard im Modul Monitoring darauf ausgelegt ist, kontinuierlich Überblick über ein oder mehrere gewählte Themengebiete zu geben, erlaubt das Modul Analyse gezielte ad-hoc Recherchen.



Abbild. 1 Modul Monitoring mit einem Dashboard zum Thema Hochwasser (Ende Mai 2015)

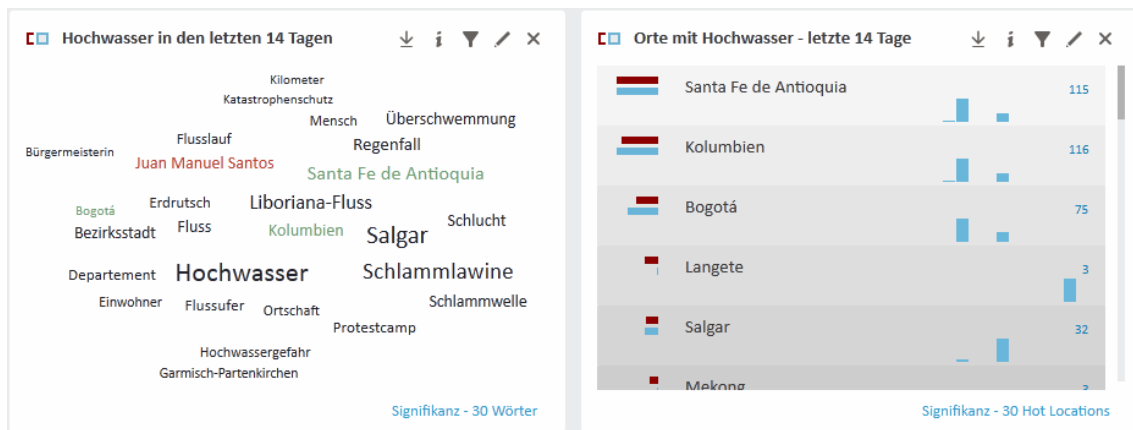
Das Modul Monitoring bietet dem Benutzer die Möglichkeit, ein oder mehrere individuell konfigurierbare Analysefenster (*Dashboards*) zur Beobachtung von Datenkollektionen zu öffnen. Das Layout ist einfach zu konfigurieren, indem per Drag & Drop ein Anzahl von Widgets (*Micro-Analysten*) mit verschiedenen Funktionalitäten hinzugefügt, arrangiert und über das Filterpanel konfiguriert werden können. Die wesentlichen Widgettypen sind:

- *Document Trend*: Anzahl der neu in den Analysekörper aufgenommenen Dokumente über die Zeit.



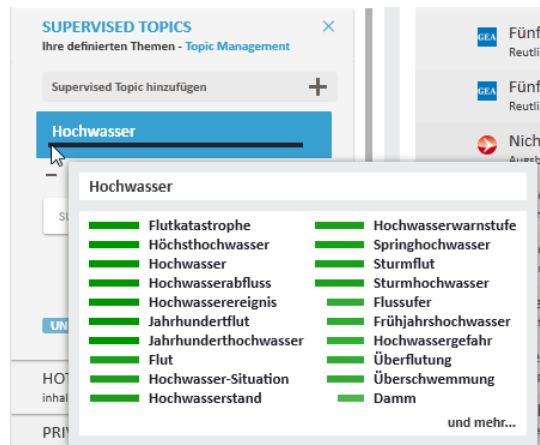
Abbild. 2 Micro-Analyst Document Trend

- *Hot Entities* und *Hot Words*: Entitäten bestimmter Art (z.B. Personen, Orte, Organisationen) oder Schlüsselwörter, welche aufgrund ihrer Frequenz oder Signifikanz besonders repräsentativ für die analysierten Daten sind, werden als Liste oder Word Cloud angezeigt.



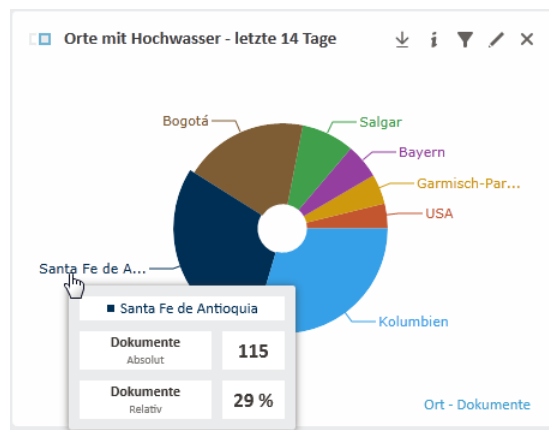
Abbild. 3 Micro-Analysten *Hot Words* und *Hot Entities* für Hochwassermeldungen in der zweiten Maihälfte 2015

- *Hot Topics* und *Supervised Topics*: Gruppierungen von Schlüsselwörtern oder Entitäten, welche Themenkomplexe repräsentieren, die besonders charakteristisch für die betrachtete Datenkollektion sind, werden nach ihrer Signifikanz geordnet angezeigt. Solche Themenkomplexe werden entweder vom System mit Hilfe statistischer Verfahren aus der Datenkollektion abgeleitet oder sie können vom Benutzer selbst definiert werden. Im weiteren Analyseprozess können Topics als flexible Filter zum Einsatz kommen.



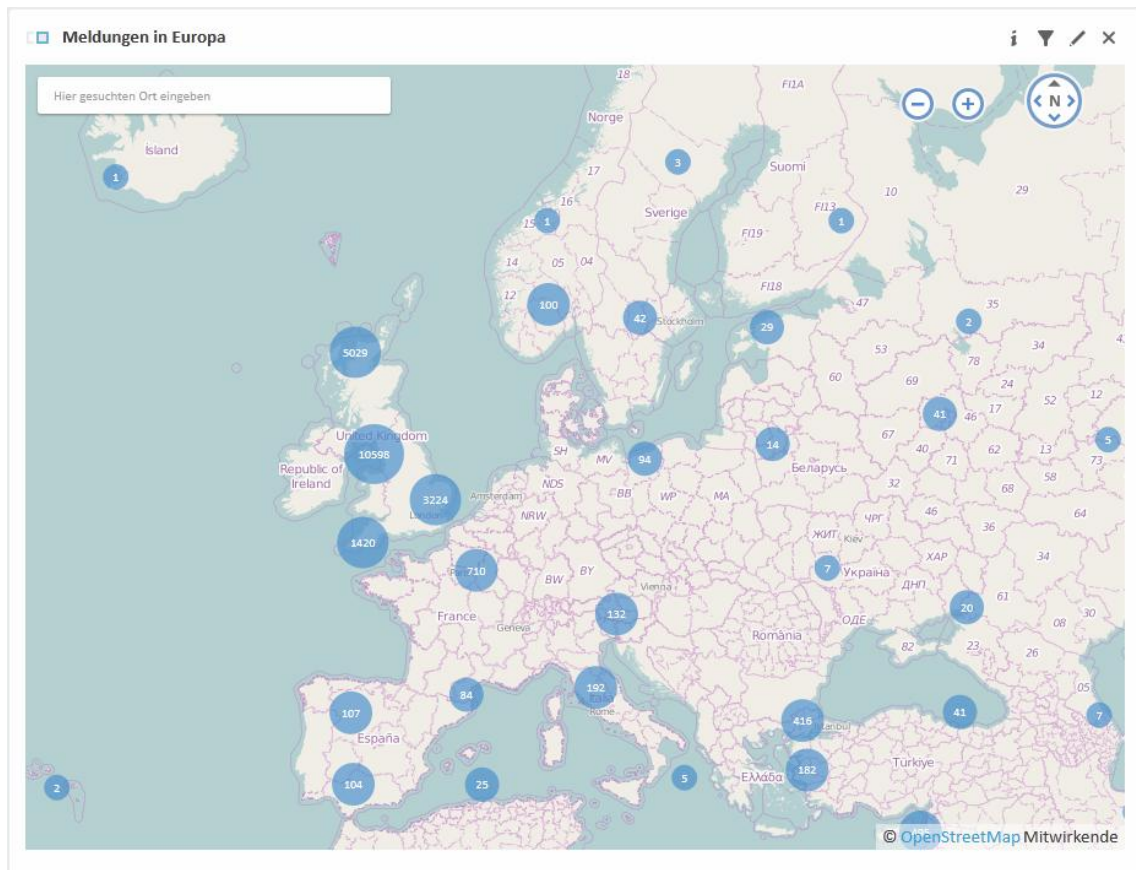
Abbild. 4 Begrifflichkeit für das *Supervised Topic* "Hochwasser"

- *Share Wheel*: Prozentuale Anteile von bestimmten Begriffen eines gewählten Typs (z.B. Schlüsselwörter, Entitäten, Quelltypen oder -domänen) relativ zum Gesamtvorkommen des Begriffstyps werden als Kreisdiagramm dargestellt.



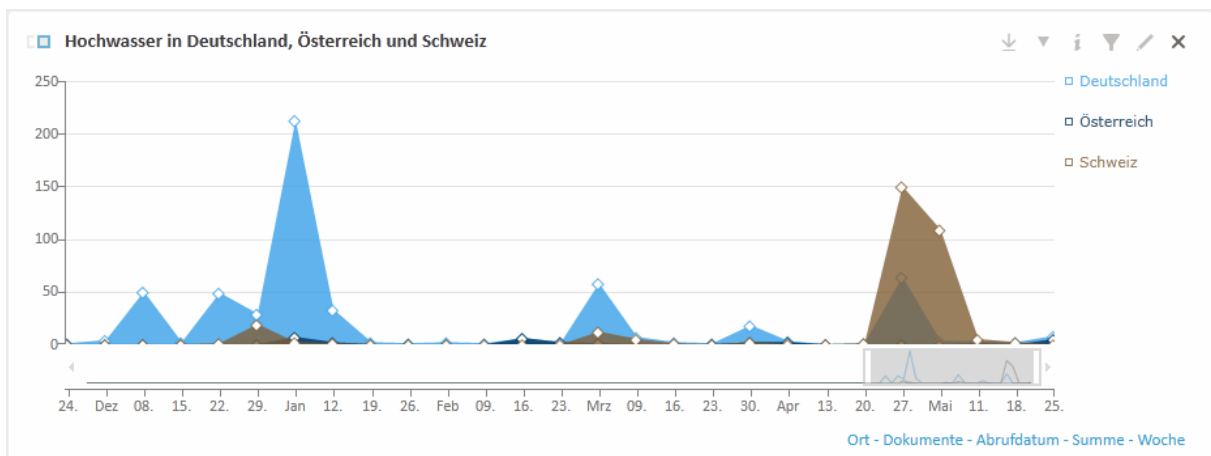
Abbild. 5 Micro-Analyst *Share Wheel* über Orte mit Hochwasser in der zweiten Maihälfte 2015

- *Map*: Die geographische Verteilung geogetaggtter Dokumente wird auf eine Landkarte angezeigt. Die Dokumente werden aufgrund räumlicher Nähe gruppiert und durch Datenpunkte repräsentiert, deren Größe der Anzahl der repräsentierten Dokumente entspricht.



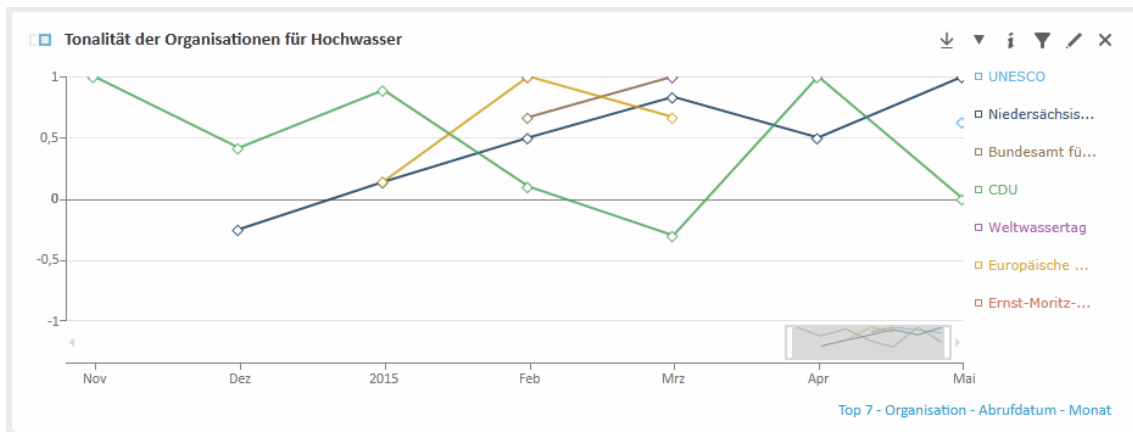
Abbild. 6 Map Micro-Analyst mit georeferenzierten Meldungen aus Europa

- *Value Over Time*: Darstellung des quantitativen Auftretens von Begriffen eines bestimmten Typs (z.B. Schlüsselwörter, Entitäten, Quelltypen oder -domänen) über die Zeit hinweg. Angezeigt werden entweder die 7 wichtigsten Instanzen des Typs oder vom Benutzer gewählte.



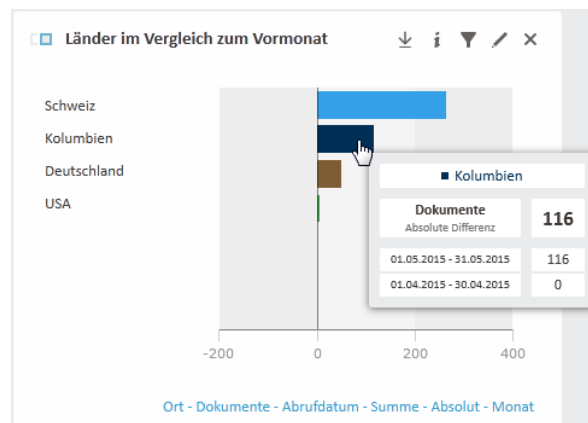
Abbild. 7 Micro-Analyst Value over Time

- *Tonality Over Time*: Anzahl der positiven, neutralen oder negativen Nennungen von Personen, Firmen oder Organisationen über die Zeit.



Abbild. 8 Micro-Analyst *Tonality over Time* für Organisationen zum Thema Hochwasser

- *Growth*: Zahlenmäßiges Wachstum oder Schwund von Begriffsvorkommen in einem gewählten Betrachtungszeitraum relativ zu einem wählbaren Vergleichszeitraum werden als Balkendiagramm dargestellt.



Abbild. 9 *Growth* Micro-Analyst mit Ländervergleich beim Thema Hochwasser im Monat Mai 2015

Jeder *Micro Analyst* wird individuell definiert und jeder *Micro Analyst*-Typ kann beliebig oft in ein und demselben *Dashboard* konfiguriert werden. So wird es beispielsweise möglich, ein bestimmtes Thema parallel in verschiedenen Datenpools zu beobachten.

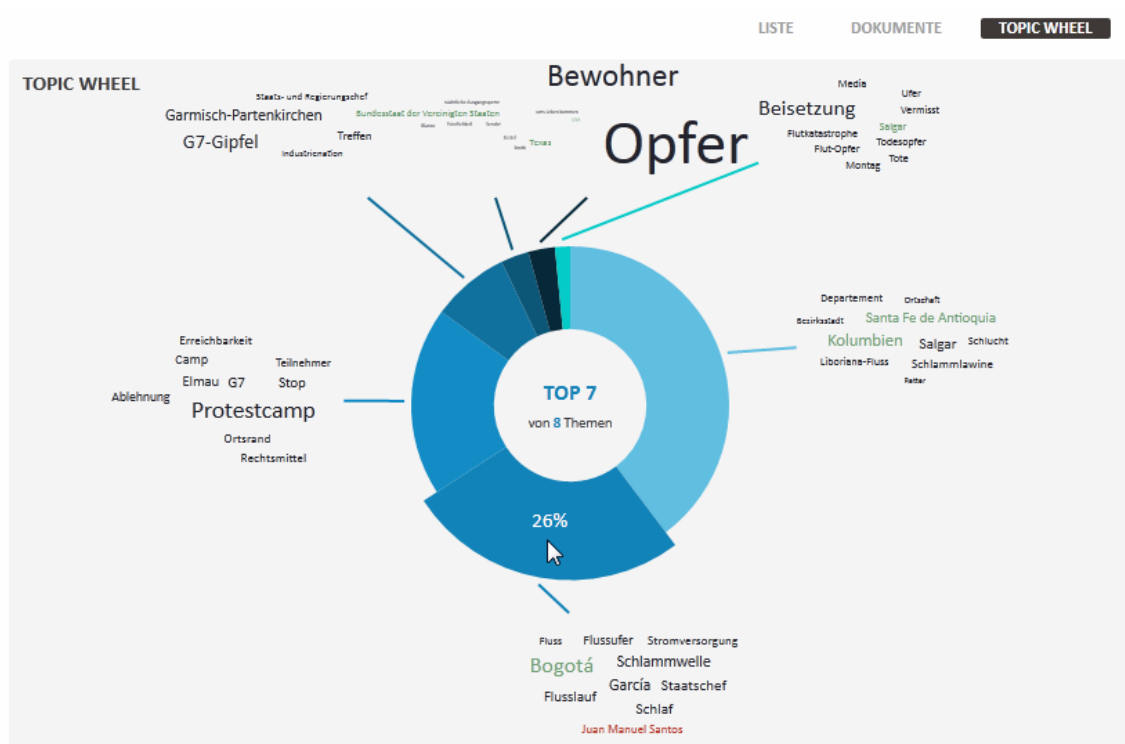
Sobald ein Begriff im *Dashboard*, etwa in einer *Word Cloud* oder auf einer Karte, angeklickt wird, wechselt die Software zum Modul Analyse mit direkter Zugriffsmöglichkeit auf die entsprechenden Dokumente. Dieser Einstieg ist auch im Kontext einer Ad-hoc-Suche möglich. Bei einer großen Anzahl von Treffern kann diese Dokumentenauswahl weiter gefiltert werden; verschiedene Visualisierungsmöglichkeiten erleichtern zusätzlich den Überblick. Das Modul Analyse bietet verschiedene Visualisierungsmöglichkeiten, die über Anklicken von Reitern aktiviert werden:

- *Dokumentliste*: Diese dient zur Voransicht der Dokumente durch Auflistung der Dokumententitel und Kurztexte, sowie Darstellung erkannter Entitäten und weiterer Dokument-Metadaten.

The screenshot shows the 'ANALYSE' module interface. At the top, there are tabs for 'MONITORING', 'ANALYSE', 'SUPERVISED TOPICS', and 'EINSTELLUNGEN'. The main content area is titled '3.084 DOKUMENTE mit Supervised Topics Hochwasser'. Below this, there is a table of documents with columns for 'DOKUMENTE AUSGEWÄHLT', 'SORTIEREN NACH', 'RELEVANZ', 'ÄHNLICHKEIT', 'ABRUFDATUM', 'TITEL', and 'QUELLE'. The table lists several documents, including 'Einschränkung bis Mitte Juni', 'Handy-App „warnt“ vor Katastrophen Alarm aus der Hosentasche', 'Vereinigte Staaten Tote und Verletzte nach Überschwemmungen in Texas', and 'Tote, Vermisste und ein Dammbbruch nahe Dallas'. Below the table, there is a detailed view of the document 'Hochwassergefahr beim Wührplatz' with a date of 27.05.2015 and a word count of 726. The interface also includes a sidebar with 'SUPERVISED TOPICS', 'HOT TOPICS', 'PRIVATE TAGS', 'GETEILTE TAGS', 'ABRUFDATUM', 'KEYWORD', 'PERSON', 'UNBEKANNTE PERSON', 'FIRMA', 'UNBEKANNTE FIRMA', 'ORT', and 'UNBEKANNTER ORT'. At the bottom, there is a footer with 'Topic Analyst • Version 6.0.9.1954 • CID GmbH © 2015'.

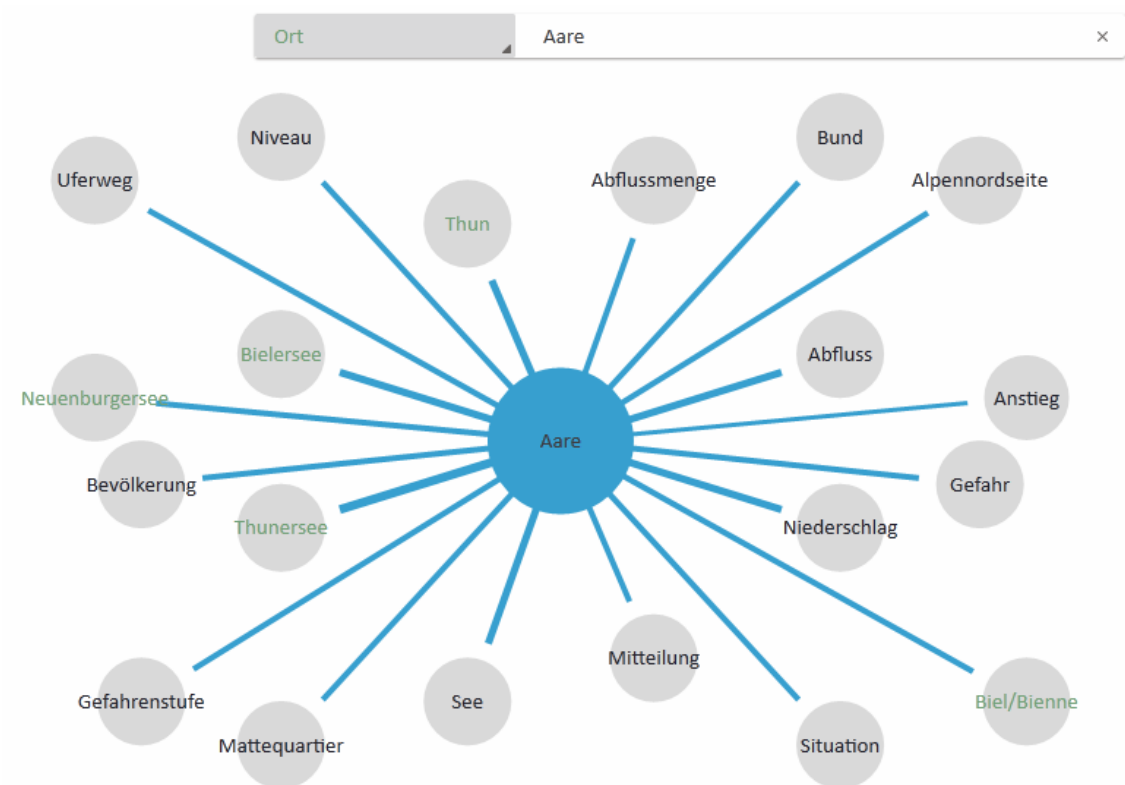
Abbild. 10 Modul Analyse mit der *Dokumentliste* zum Thema Hochwasser

- *Analyseansicht*: Wie im Dashboards stehen nun die verschiedenen *Micro Analysten* zur Verfügung, die im Modul Analyse jedoch nicht fest definiert sind, sondern sich mit jeder neu vorgenommenen Filtereinstellung selbst aktualisieren.
- *Word Cloud*: Wieder erlaubt es die *Word Cloud*-Darstellung der häufigsten oder signifikantesten Begriffe einer gewählten Art, sich einen schnellen Überblick über die wichtigsten Inhalte einer Dokumentenkollektion zu verschaffen.
- *Topic Wheel*: Aus den häufigsten oder signifikantesten Begriffen einer Dokumentenkollektion werden weitergehend Themenkomplexen automatisch generiert. Zu den wichtigsten 7 Themen wird eine *Word Cloud* angezeigt; die prozentualen Anteile der Themen relativ zur Dokumentenmenge werden in Form eines Kreisdiagramms dargestellt.



Abbild. 11 Topic Wheel mit den Themenbereichen zu Hochwasser in der zweiten Maihälfte 2015

- *Korrelationsanalyse/Spektrum:* Untersucht und visualisiert ausgehend von einem gewählten Begriff (z.B. Entität, Schlüsselwort), welche anderen Begriffe häufig gemeinsam in Dokumenten auftreten.



Abbild. 12 Korrelationsanalyse für den Fluss Aare im Monat Mai 2015

- *Schwellenwert-Ansicht*: Wenn eigene Themen (*Supervised Topics*) definiert wurden, kann hier per Regler der Grad der Übereinstimmung von Trefferdokumenten mit dem so festgelegten Filterkriterium angepasst und damit die Größe der weiter zu untersuchenden Dokumentenmenge beeinflusst werden.

Durch diese flexibel einsetzbare Sammlung an Werkzeugen hilft Slándáils Sprachverarbeitung dem Krisenmanager dabei, aus der Menge und Vielfalt digital verfügbarer Informationen konkretes Wissen abzuleiten. Sie erhöht die Qualität und Schnelligkeit von Informationssuche und kann unerwartete oder verborgene Aspekte und Zusammenhänge aufzeigen. Der Krisenmanager wird so bei seiner eigentlichen Hauptaufgabe - den strategischen Entscheidungs- und Organisationsaufgaben zur Bewältigung einer Krisensituation - unterstützt und entlastet.

Zusammenfassung:

Die Digitalisierung der Kommunikation hat zur Verbesserung des kommunikativen Austauschs in verschiedenen relevanten Bereichen des Krisenmanagements geführt. Angesichts der Flut an digital verfügbaren Informationen ist das Krisenmanagement auf geeignete computerbasierte Werkzeuge zum Suchen, Ordnen und Analysieren von Informationen angewiesen. So können Krisenmanager das positive, bislang jedoch unvollständig genutzte Potential der sozialen Medien erschließen wie auch deren mögliche negative Effekte kontrollieren. Im europäischen Kooperationsprojekt Slándáil erstellen akademische und industrielle Partner in enger Zusammenarbeit mit Krisenmanagern Rahmenbedingungen zur faktisch und ethisch korrekten Nutzung sozialer Medien und entwickeln ein Softwaresystem, welches klassisches Katastrophenmanagement mit automatischer Bild- und Sprachverarbeitung kombiniert. Der hier vorgestellte sprachverarbeitende Teil der Software (Prototyp1) eröffnet effiziente Möglichkeiten, um aus großen Mengen an Informationen schnell und zuverlässig Wissen abzuleiten und dem Krisenmanagement wertvolle Entscheidungshilfen an die Hand zu geben.

Literaturverzeichnis

Florian Semle (2015): Das "Social Web" als Herausforderung für die Krisenkommunikation, Krisennavigator (2010). In Krisennavigator (11) 2015, Online verfügbar unter www.krisennavigator.de/Das-Social-Web-als-Herausforderung-fuer-die-Krisenkommunikation.951.0.html, zuletzt geprüft am 25.05.2015.

Kirchbach Kommission (2013): Bericht der Kommission der Sächsischen Staatsregierung zur Untersuchung der Flutkatastrophe 2013 (Kirchbachbericht), Sächsische Staatskanzlei, Dr. Nowak, Erik; Rischke, Lars; Zorn, Hartwig (2013). Online verfügbar unter www.publikationen.sachsen.de/bdb/artikel/20534, zuletzt geprüft am 25.05.2015.

Bundesministerium des Inneren (2008): Krisenkommunikation Leitfaden für Behörden und Unternehmen. Online verfügbar unter www.bmi.bund.de/SharedDocs/Downloads/DE/Broschueren/2008/Krisenkommunikation.pdf?__blob=publicationFile, zuletzt geprüft am 25.05.2015.

Rutrell Yasin (2010): 5 ways social media is changing emergency management. Online verfügbar unter www.gcn.com/articles/2010/09/06/social-media-emergency-management.aspx, zuletzt geprüft am 25.05.2015.

Slandail - Security System for language and image analysis (2014). Online verfügbar unter www.slandail.eu, zuletzt geprüft am 25.05.2015.

The research leading to these results has received funding from the European community's Seventh Framework Programme under grant agreement No. 607691 (SLANDAIL).