



## SLÁNDÁIL

EIN EUROPÄISCHES SICHERHEITSSYSTEM FÜR DIE SPRACH- UND BILDANALYSE

Universität Leipzig - Lehrstuhl für Automatische Sprachverarbeitung / Institut für Angewandte Informatik (InfAI)

*Prof. Gerhard Heyer*



## DAS SLÁNDÁIL PROJEKT

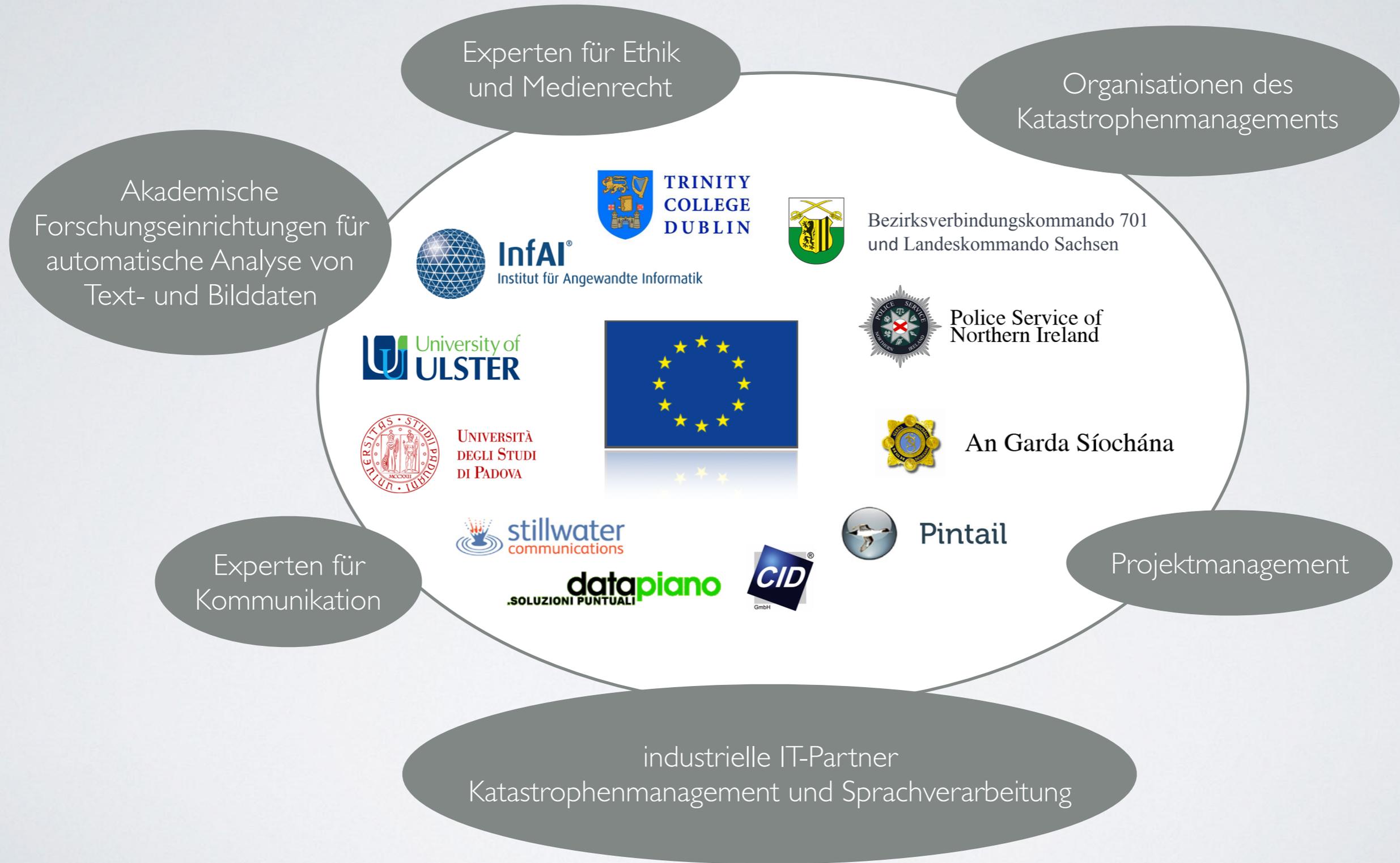
- Kooperationsprojekt zwischen 11 Partnern in Irland, Nordirland, Italien und Deutschland  
Projektzentrale: Dublin (Prof. Khurshid Ahmad)
- gefördert durch EU im Rahmen des FP7 Security Programms  
Projektdauer: 3 Jahre (Beginn April 2014)
- **Hauptziel:** Entwicklung und Erprobung eines Systems (Prototyps) zur Nutzung von Informationen aus sozialen Medien im Katastrophenmanagement (insbesondere Naturkatastrophen)

### besondere Merkmale:

- Verarbeitung von Daten in unterschiedlichen Sprachen (Englisch, Deutsch, Italienisch) und Modalitäten (Text, Bild)
- Einbeziehung von sprach- und kulturübergreifenden Aspekten von Kommunikation
- besondere Beachtung von Fragen der ethischen und rechtlichen Korrektheit der Datennutzung



# DAS SLÁNDÁIL KONSORTIUM



## DAS POTENTIAL SOZIALER MEDIEN FÜR DAS KATASTROPHENMANAGEMENT

- ← schnellere und direktere Weitergaben von Informationen von Behörde an Betroffene
- ← breitere Öffentlichkeit erreichen
- ↻ Interaktionen mit Betroffenen und Kooperation mit Freiwilligen
- Lagebild verbessern mit Hilfe von Informationen von Bevölkerung vor Ort
- Beobachtung von Aktivitäten, Meinungen und Stimmungslagen der Bevölkerung
- ↻ Informationsaustausch und Emotionsmanagement innerhalb der Bevölkerung
  - ! Verbreitung von Falschmeldungen, negativen Stimmungen oder Panik
- ↻ eigenständige Organisation von Spenden und Aktionen Freiwilliger
  - ! Selbstorganisation unangemessener Aktionen Freiwilliger





# DER AKTUELLE GEBRAUCH SOZIALER MEDIEN

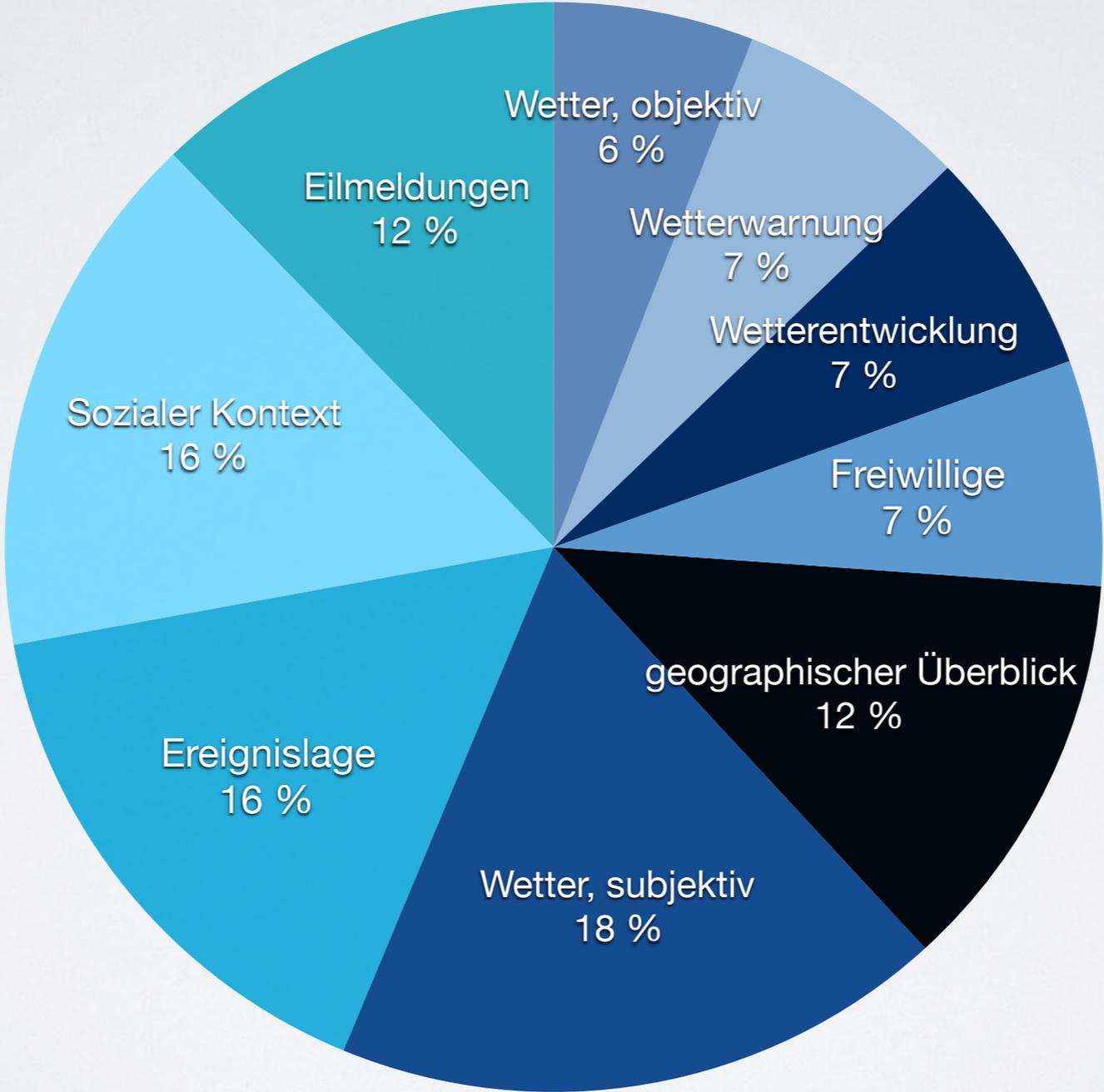
Fallstudie: Flut 2013 Deutschland & Österreich

Topic Model Analyse (HDP CRF)



**Twitter:**

>300.000 Tweets





# DER AKTUELLE GEBRAUCH SOZIALER MEDIEN

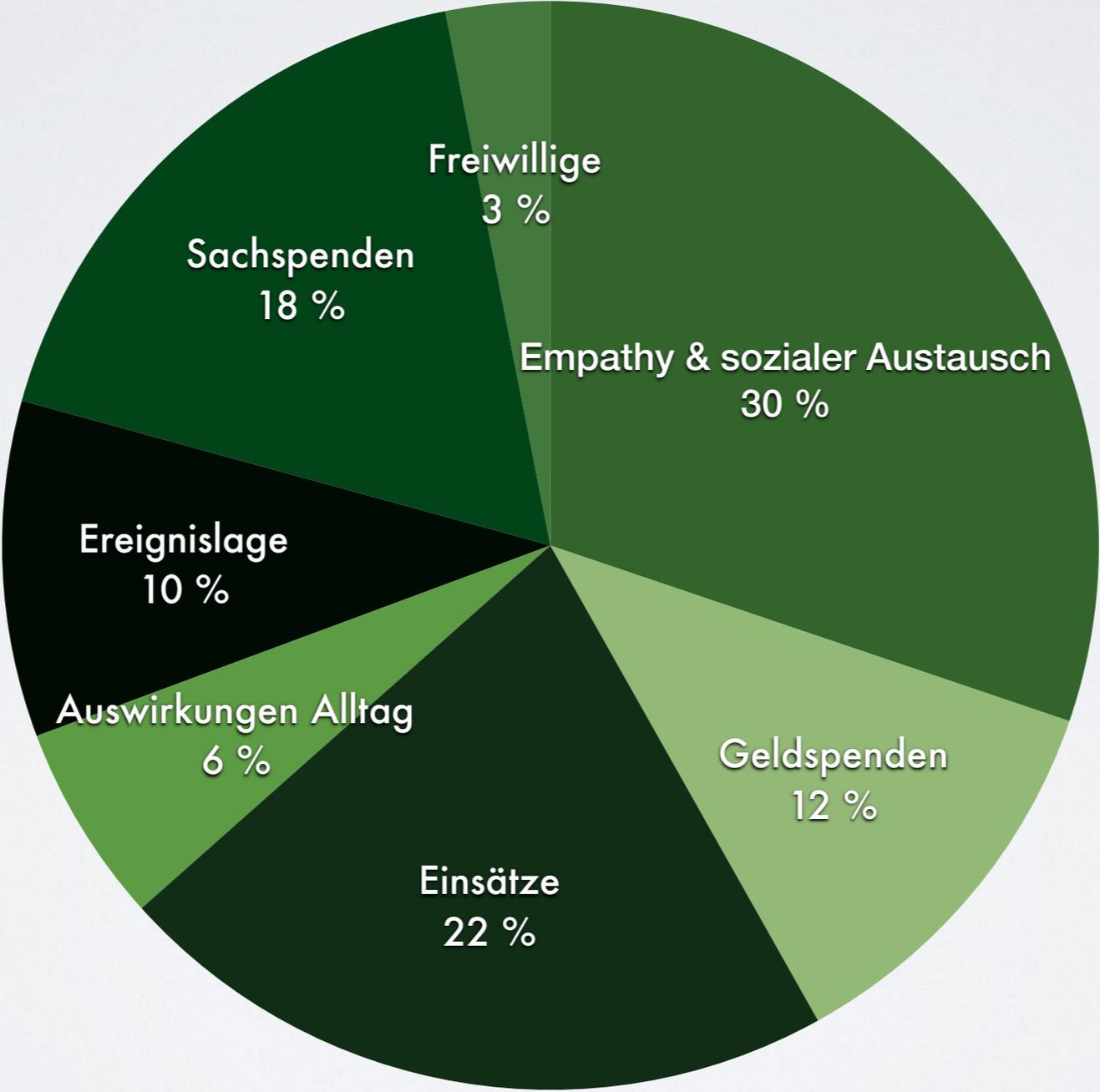
Fallstudie: Flut 2013 Deutschland & Österreich

Topic Model Analyse (HDP CRF)



**Facebook:**

>40.000 Posts



## DIFFERENZIERENDE MERKMALE UNTERSCHIEDLICHER MEDIEN

Fallstudie: Flut 2013 Deutschland & Österreich

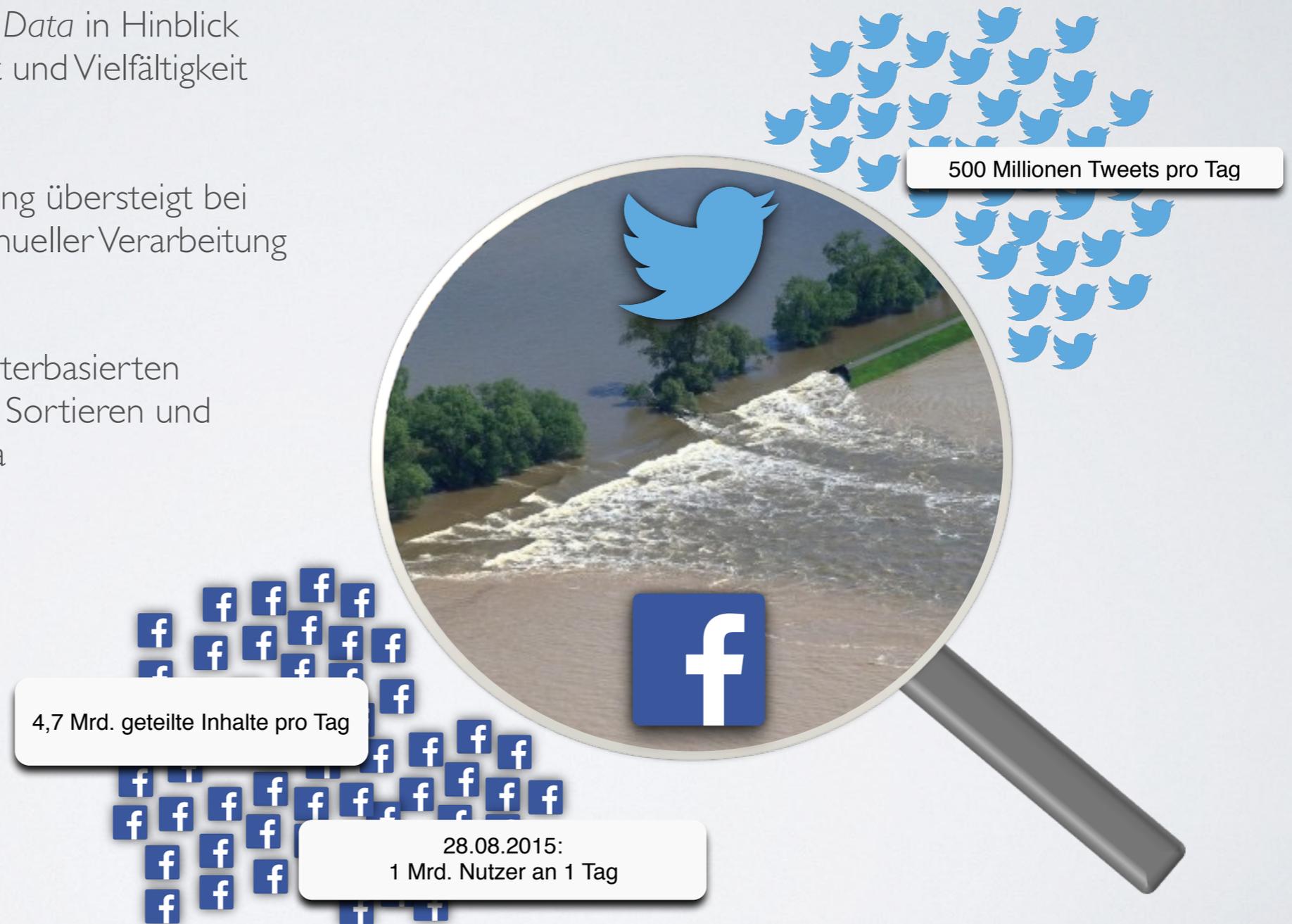
Differenzanalyse (Log-Likelihood)

	 Facebook	 Twitter	 News
Charakteristischer Inhalt	Empathy, sozialer Austausch & Organisation von Hilfe	aktuelle Informationen zur Ereignislage	Ereignis im größeren gesellschaftliche Kontext (z.B. Finanzen, Politik)
vorherrschende Konzeptualisierung	Perspektive der Betroffenen	objektive Perspektive	Perspektive der breiten Öffentlichkeit
dominanter Sprachmodus / Sprachstil	emotional involviert / informell	situativ berichtend / faktisch	informierend, diskutierend / allgemein

# INFORMATIONSÜBERFLUTUNG

Wie kann man relevante Informationen finden?

- Social Media Daten sind *Big Data* in Hinblick auf ihr Volumen, Dynamizität und Vielfältigkeit
  - ihre systematische Auswertung übersteigt bei weitem jede Möglichkeit manueller Verarbeitung
- dringender Bedarf an computerbasierten Werkzeugen für das Suchen, Sortieren und Analysieren von Social Media



## HERAUSFORDERUNGEN

- Standardmethoden und -werkzeuge zur automatischen Sprachverarbeitung müssen an Besonderheiten der Social Media Daten angepasst werden

Kontextinformationen: Metadaten, Bilder, Links

Sprachformat: kurz, fragmentarisch

Sprachform: orthografische Fehler, Abkürzungen, Slang



- Integration eines sprachverarbeitenden Moduls in bestehende Software zum Katastrophenmanagement und Abstimmung der Funktionalität auf spezielle Bedürfnisse der Nutzer



- 24h interaktiver Informationsfluß in one-to-many Setting bringt neue Formen der Kommunikation und neue Inhalte; fordert Anpassung von Organisationsstrukturen, kommunikativen Regeln und Fertigkeiten

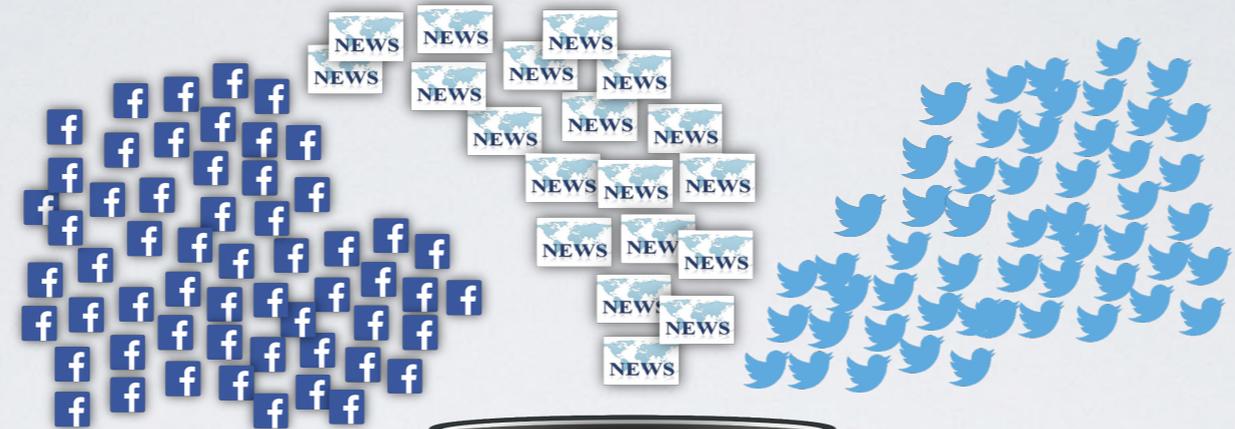


- Fragen der Ethik und des Rechts (Privatsphäre, Urheberrecht, Datenschutz)

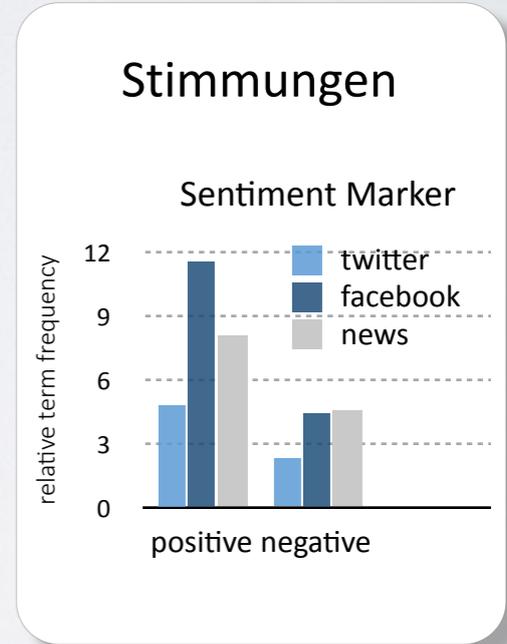
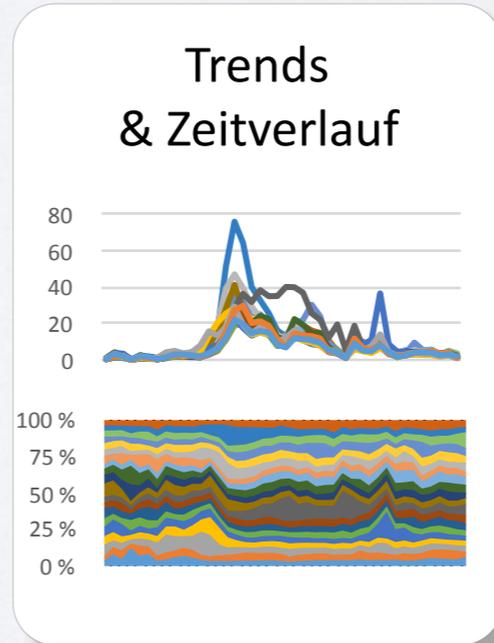
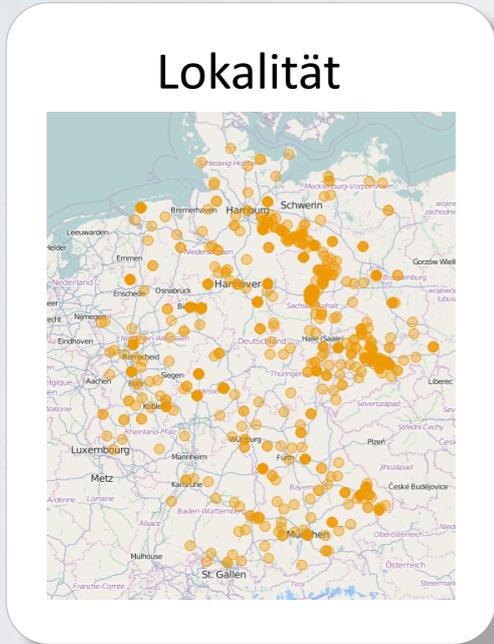




# DAS SLÁNDÁIL SPRACHVERARBEITUNGSSYSTEM



suchen, filtern, sortieren,  
analysieren, visualisieren





***Vielen Dank!***

*Prof. Gerhard Heyer*

[heyer@informatik.uni-leipzig.de](mailto:heyer@informatik.uni-leipzig.de)

*Dr. Sabine Gründer-Fahrer*

[gruender@uni-leipzig.de](mailto:gruender@uni-leipzig.de)

*Antje Schlaf*

[antje.schlaf@informatik.uni-leipzig.de](mailto:antje.schlaf@informatik.uni-leipzig.de)