# Social Media Information Analysis for Crisis and Disaster Managment

Hermann Huber, Georg Neubauer, Andrea Nowak, Joachim Klerx, Bettina Jager
**Austrian Institute of Technology GmbH**

Klaus Mak, Christian Meurers

**Austrian Armed Forces,**
**National Defence Academy**

Georg Neubauer
AIT Austrian Institute of Technology GmbH
Safety & Security Department
A-2444 Seibersdorf
georg.neubauer@ait.ac.at

TIEMS International Conference on
*Public Alerting and Social Media during*
*Crisis and Disasters*
October 31, 2014

# Table of Contents

- Data – Information - Knowledge

- Challenges

- Applications of Social Media Information Analysis

- Classification of Social Media Plattforms

- Social Media Strategy Lifecycle

- Ubicity – Processing Crowd Sourced Data

# Data – Information - Knowledge (1)



- The net contains huge amounts of unstructured data in different formats

- The amount of existing digital data surpassed 1.8 zettabyte in 2011 (IDC 2011), it is expected that the amount of data doubles every year

# Data – Information – Knowledge (2)

## From Bellinger et al* on Data, Information, Knowledge and Wisdom

- **Data** is raw, simply exists and has no significance beyond it's existence
- **Information** is data that has been giving meaning by way of relational connection, the meaning can be, but must not be useful
- **Knowledge** is the appropriate collection of information being useful
- Understanding is the process of generating knowledge from existing knowledge
- Finally, **wisdom** is based on all previous levels – it gives understanding about which there has previously been no understanding
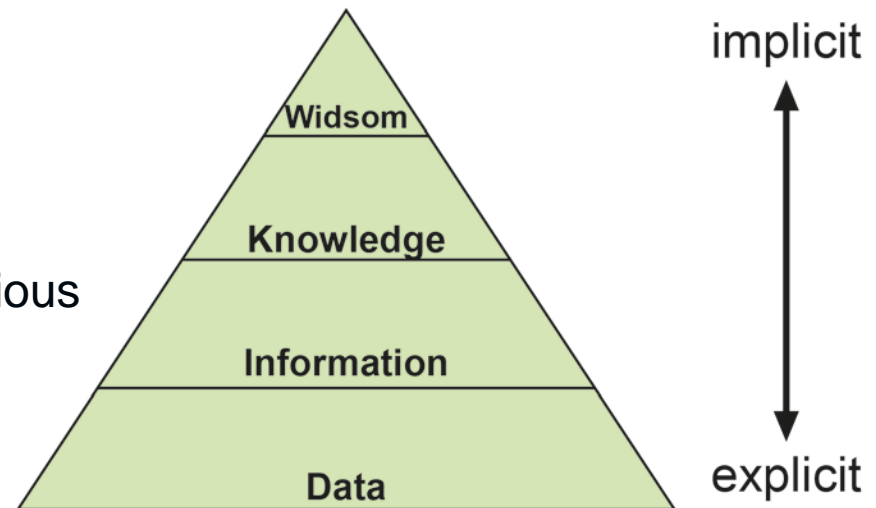
# Non-technical challenges for CDM applications

- Challenges: privacy, legality, reliability, accuracy, etc.

- Approaches:

  - Inform users properly on **data and privacy policy**

  - Leverage the idea of **crowd sourcing** (data acquisition, validation, etc.)

  - Make use of **open source software** (see Ushahidi, Sahana)

  A Social Media strategy should always be accompanied by a comprehensive **community building strategy.**

# Application of Social Media Information Analysis

- **Pre-crisis situation**
  - detect crisis in their early stages

- **In-crisis situation**
  - Analyze the current situation in terms of available resources or humanitarian needs
  - Utilize geo-related data for crisis coordination and communication



- **Post-crisis situation**
  - Analyze crisis activities and derive forecasting models
  - Improve training scenarios and create advanced pre-crisis analysis
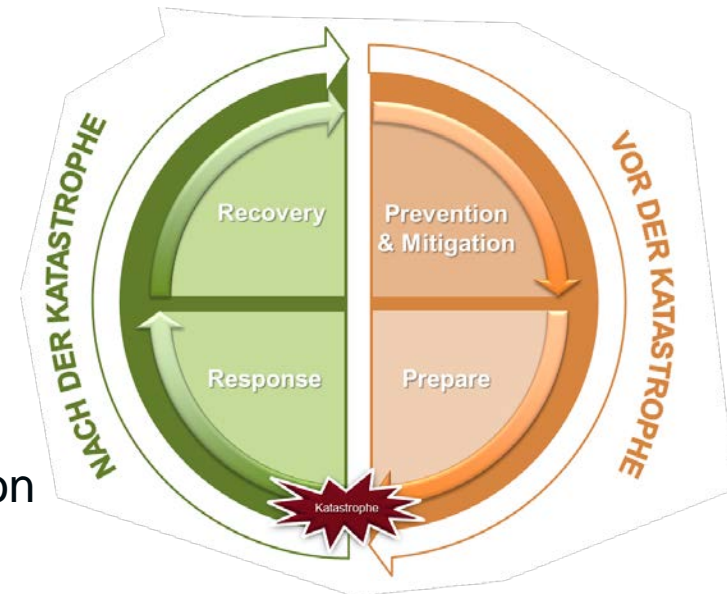
# Social Media Plattforms Classification
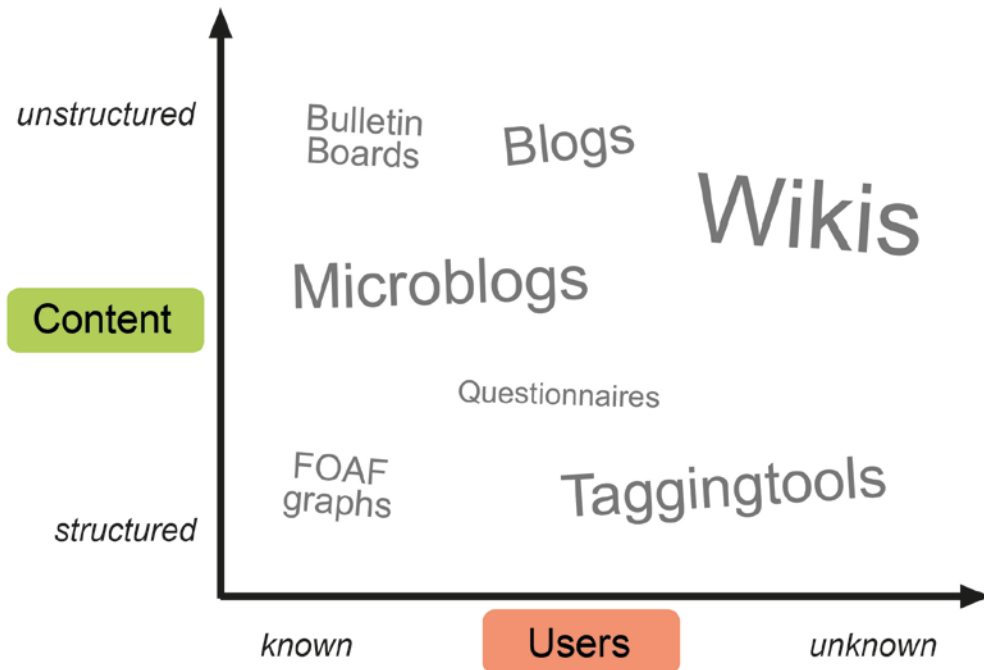


Two basic dimensions

- **Content**: is the content structured or not?

- **Users**: are we able to identify content providers?

- Some use cases may require additional dimensions
  - Communication path (uni-directional, bi-directional, start-like, etc.)
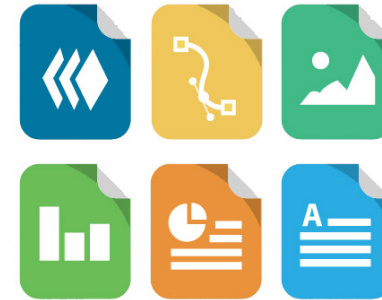  - Geo-spatial resolution (country, city, longitude/latitude, etc.)

# Social Media Content Classification
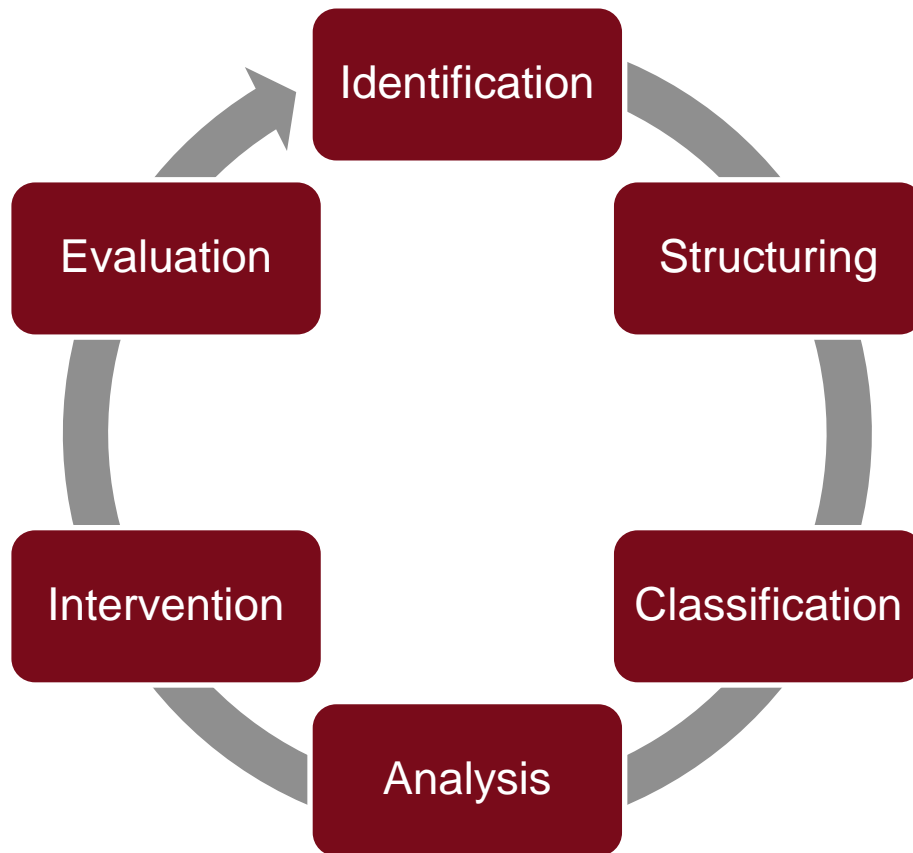
## Network Data
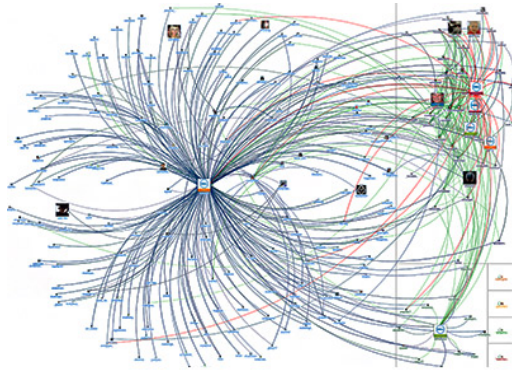


- User profiles can be perceived as a network graph

- Identify authorities and hubs (communication multipliers)

- Use knowledge about network to improve dissemination
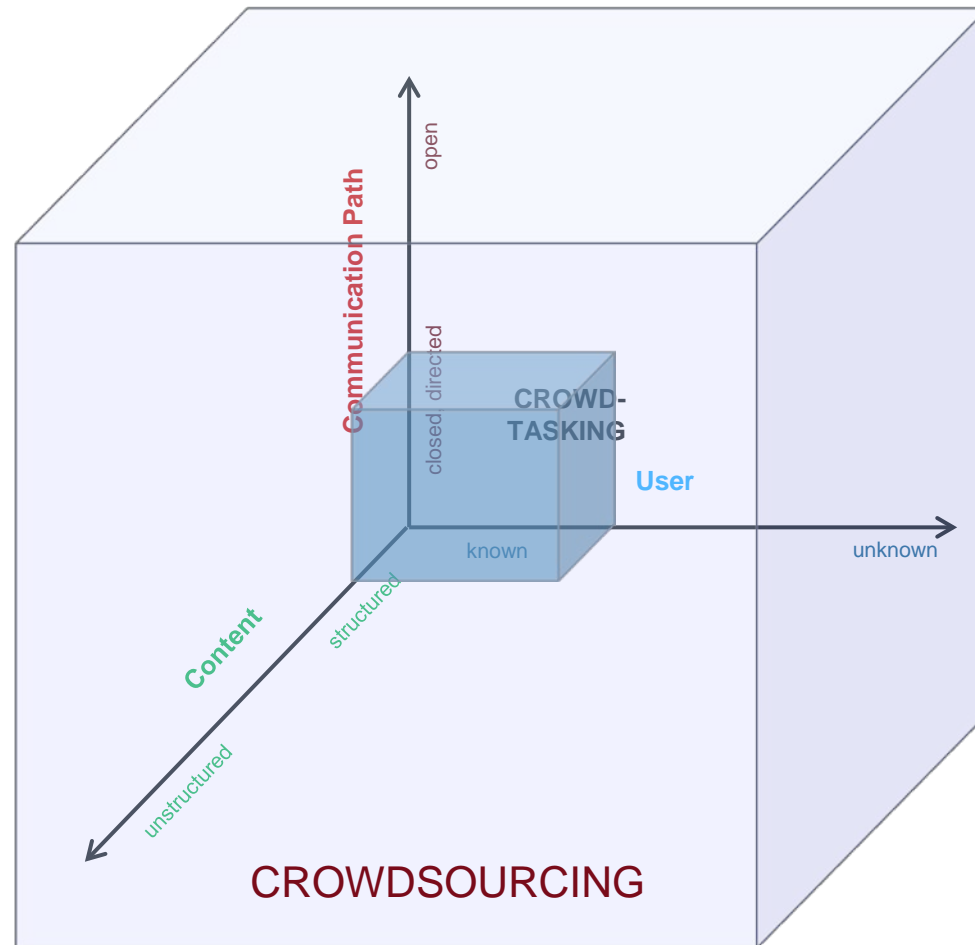
## Content Data



- User-genrated content

- Text, short text, pictures, videos, full HTML, etc.

- Understand thematic structure

- Discover trends and topics etc.

# Network-centric issue management



- Identification of channels

- Structuring of content

- Classification of content

- Analysis of content

    - Emerging topics
    - Influencer groups
    - Opinion leaders

- Intervention e.g. with information campaign, moderation, etc.

- Evaluation of results

# Crowdsourcing – the „Crowdsourcing Space"

# Ubicity – Processing Crowd Sourced Data

- Ubicity will be free and open source software that enables to:

  - collect **large amounts** of geo-enabled data

  - accept **any content provider** (social media, hardware sensors, etc.)

  - create powerful **search indexes**

  - give rise to cross-domain **data interpretations**

  - enable a human **crowd to act**
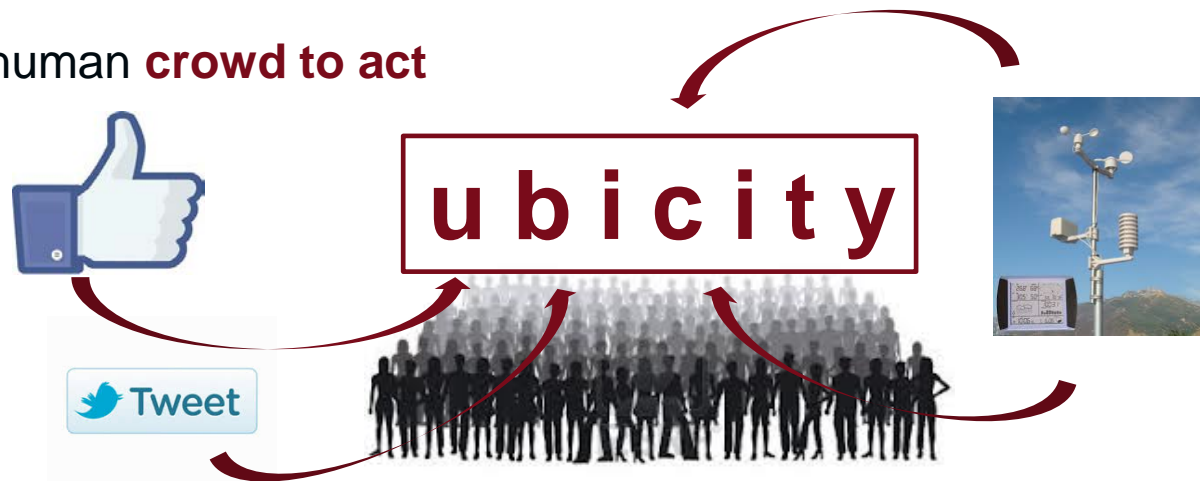
# Ubicity – Processing Crowd Sourced Data

- In the current stadium, we handle
  - Any geo-enabled tweet from Twitter
  - Any meteorological data set
  - Any sensor data set & time series from openUWADAT
    - Traffic, noise, air quality & pollution, etc.

- Currently, we can
  - Store more than 54.000 unstructured documents
  - Index 66.000 unstructured document to leverage search
  - … **per second**

# Conclusion

- Social Media information analysis requires to think about
    - Legal aspects, privacy, copy right
    - Ethical and cultural aspects
    - Reliability, accuracy

- Though, many opportunities for CDM applications
    - sharpen decisions of crisis managers
    - additional approaches to co-ordinate emergency forces
    - psychological assitance of affected people
    - Bi-directional communication path
    - Peer reviewing approaches for crowd sourcing environments

# Thank you for your Attention

Georg Neubauer
AIT Austrian Institute of Technology

2444 Seibersdorf
Austria

georg.neubauer@ait.ac.at