Welcome to INFO319:
Research Topics in Big Data

Vimala Nunavath

<Vimala. nunavath@uia.no>

What are we going to do today?

- Some practical information
- Expectations from the course
- About the themes
    - what are we going to focus on?
- About the programming projects
    - how about a coordinated task?
- About the essay
- Introduction to Emergency Management
- Plans for the next session?

Themes:

- Possible themes:

  - Introduction to Emergency management

  - Introduction to Big data

  - Big Data sources for EM

  - Sensors/IOT for EM

  - Social media for EM

  - Machine learning and Natural Language Processing for EM

  - Spatial Intelligence for BDEM (Visualization/Dashboards)

Sessions:

- 8 sessions
    - Thursdays and Fridays from 1015
        - seminar room 548 (SV:S548) in this building
    - the sessions will become highly interactive
        - you present, critique and discuss
            - chapters and papers
            - your essays and projects
    - 80% attendance is required
        - that is: 80% active participation

Prerequisites:

- Need some programming skills.
- Formal requirements:
  - active participation in seminars (80%)
  - graded individual essay (30%)
  - graded group programming project (40%)
- Final exam (30%).

- Grades are given as follows:

    - *Written exam (30%):* An individual

    - Theoretical essay with thoughtful research and discussion of an assigned topic (30%).

    - In addition, there is a practical assignment in groups (40%).

- Both assignments must be done in the teaching semester.

- Dates:

    - Submission of individual essay: 4.12.2019

    - Submission of group project assignment: 13.12.2019

    - Exam date: December 18th 2019.

Curriculum:

- Textbooks:
    - Rob Kitchin (2014). The Data Revolution: Big Data, Open Data, Data Infrastructures & their Consequences. Sage, 208 pages.
    - Carlos Castillo (2016). Big Crisis Data – Social Media in Disasters and Time-Critical Situations. Cambridge, 212 pages.
- Research papers in the wiki
    - will appear in wiki.uib.no/info319 – possibly quite a few!
    - we will make a selection of core paper
- mitt.uib.no   for announcements / discussion
- More information about the course reading list can be found on:
https://wiki.uib.no/info319/index.php/Sessions

The Themes

Practicle Information

Central themes

- Central themes:

  - Introduction to Emergency management

  - Introduction to Big data

  - Big Data sources for EM

  - Sensors/IOT for EM

  - Social media for EM

  - Machine learning and Natural language processing (NLP) for EM

  - Visualization/Dashboards

Introduction to Emergency management

- Lots of emergencies such as natural and technological and man-made emergencies have been happening all around the world, every day and even every single hour.

- What is emergency/disaster.

- What are the phases of EM.

- Who are involved.

- What are the Information needs of stakeholders.

- What are the challenges and obstacles for EM.

- Which new tools and techniques are needed for EM?

Introduction to Big data

- Big (and open) data are becoming enormously important
    - making sense of them is a major challenge for EM.
- What is big data?
- What are existing big data sources?
- What are the dangers of big data?
- Privacy, surveillance, societal impact of big data;
- What are big data challenges?
- Which new tools and techniques are needed?

Emergency Big Data sources for EM

- Big (and open) data are becoming enormously important…

- What are the available EM big data sources?

- How to deal with EM big data sources for making sense of it?

  – semantics can help

- What are the characteristics of big data sources?

- What technical architectures for big data management are available and needed?

Sensors/IOT for EM

- Lots of sensors, actuators and other (smart and dumb) things are on the net.

    - closely related to big data as a source.

    - examples: Radio-frequency identification (RFID), sensor web, wireless sensor network etc..

- What are the characteristics of the IoT?

- Which research efforts target the use of IoT data for EM?

- Which new developments are needed?

Social media for EM

- Social media have played an increasing role in emergencies and disasters. It has become a central source of big data for EM.

- How to collect (process) social media data and present it to needed stakeholder during EM?

- How are social media streams being analyzed?

- How can the analysis results be organized as social network graphs?

- How can the graphs be used, e.g., to detect events and provide information?

Machine learning for EM

- Knowledge building from large data sets
  - available data are used to automatically build predictions, models, explanations(?), theories...
  - machine learning techniques.

- Collect data and test them.

NLP for EM

- Social networks offer a wealth of information for capturing additional information on people's behavior, trends, opinions and emotions during disasters.

- What is NLP? Why sentiment analysis?

- Why it is important?

- Who are beneficiaries?

- How to train a sentiment classifier to categorize messages, which was then used to examine the level of concern and anxiety in the people before and after the disaster?

Visualization/Dashboards

- Spatial Intelligence for BDEM
- How to visualize the processed data on dashboard??

Which themes should we focus on?

- Go to docs.google.com

(https://docs.google.com/forms/d/1UaOM10WwxLv8lCIpJEL65-hhQ7ha2xodEuVZku4ctzg/edit)

- What themes do you like (so far)?

- What type of course do you want?

- And how much do you know about each theme?

The Essays

Individual essays

- The essay shall present and discuss selected theory, technology and tools related to big data technologies and EM, backed by scholarly and other references
    - counts 30% of final grade
    - deadline: December 4rth 1400
    - presentation: December 5th
- Encouraged:
    - more than a paper
    - social media contrib's (Wikipedia, Wikidata...)

The student group Programming Project

Group programming project

- The project shall develop an application that can be used for emergency management. Development and run-time platform is free choice, as is programming language. The project should be carried out in groups of three and not more. Working individually is also allowed.

- Counts 40% of final grade.

- Final presentation: December 6th, 2019.

- Submission deadline: December 13th, 1400