# Groupscope: A Microscope for Large Dynamic Groups Research

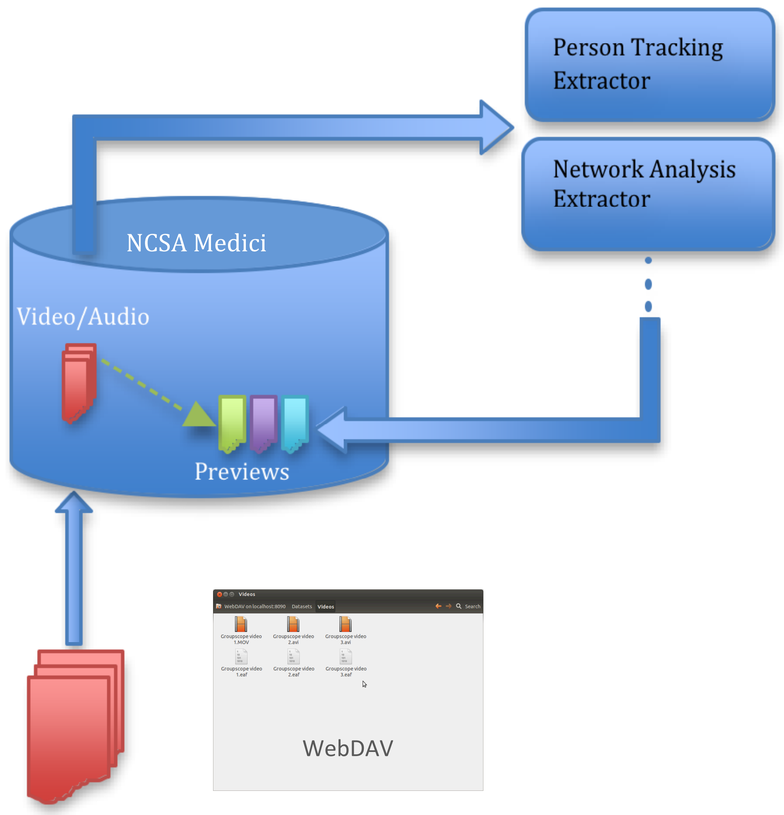
Sandeep Puthanveetil Satheesan, Marshall Scott Poole, Rob Kooper, Kenton McHenry
National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL, USA
{sandeeps, mchenry}@illinois.edu

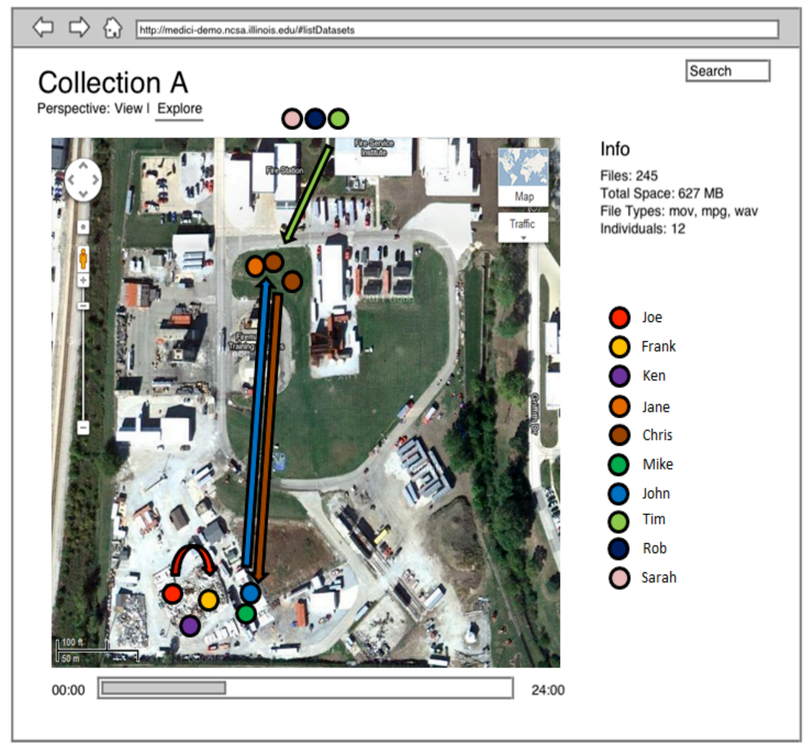
## Abstract

Social Science research on large dynamic groups of people involves many challenges of which a prominent one is working with large amounts of video and audio data. Lack of automation has resulted in a methodological gap on how research is conducted in this area. For example direct observation and annotation becomes a tedious and manual task as the size of the group and the duration of interaction increases. We propose to leverage an extensible and scalable content aware data repository called Medici to incorporate content based analysis tools from computer vision as well as manual annotations tools to provide a more practical solution for storing, managing, processing, and analyzing data for researchers enabling them to make firsthand observations and verification of theories.

## Large Dynamic Groups (LDGs)

- People interacting in dynamic subgroups in indoor or outdoor spaces
- 8 ≤ group size ≤ 200 (approx.)
- Critical to functioning of modern-day society
- Disaster and emergency response
- Medical care
- Product design
- Military operations




## Gaps in LDGs' Research

- Less theory and research on LDGs
- Usually treated as large small groups
- Currently used methodologies:
  - Ethnography
  - Participant observation
  - Analysis of transactional records
  - Survey based network analysis
- Almost NO direct observation
- E.g. disaster simulation by 40 emergency responders for 6 hours (20 cameras and 40 microphones) = 120 hours video + 240 hours audio ≈ 800 GB
- Difficulty in managing, visualizing and analyzing large amounts of data without software tools

## NCSA Medici

- Scalable content aware data repository
- Heterogeneous data
- RESTful services
- Data storage layer abstraction
- Scaling based on study requirements
- Extractors
- Previewers

## Integration with ELAN

- Tool for annotation and transcription
- WebDAV as platform for integration
- Provides a more intuitive interface
- Easier management of data and annotation files


## Person Detection and Tracking

- Identification of communication links
- Using Kinect depth data for indoor scenarios
- Tracking people from a distance
- *"Tracking People by Learning their Appearances", Ramanan et al (2007)*
- *"Detecting interaction links in a collaborating group using manually annotated data", Mathur et al (2012)*

## Use Cases

- Improving efficiency of multi-team interaction
  - Proposing and verifying theories on LDGs
- Studying acts of aggression and bullying in classrooms and playground
  - Additional challenges due to noise and long distance person tracking

## Conclusion

- A scalable content aware data repository based solution for conducting research LDGs and analyzing data
- A microscope to study the world of LDGs
- Could lead to pioneering changes in Social Science research

### Acknowledgement

This work at NCSA/UIUC has been supported by the National Science Foundation under grant BCS-0941268.

For more information, please visit http://isda.ncsa.illinois.edu