Acknowledgements

• Mentors (MIT: Licklider, Kessler; Cornell: Salton)

• Virginia Tech, CS (ECE), Digital Library Research Laboratory (DLRL)

• Department Head 2008-2015, Barbara Ryder, VP for ACM, many awards

• NSF and many other sponsors

• Students, colleagues, co-investigators (selected):

• DL Curriculum: Sanghee Oh, Jeffrey Pomerantz, Barbara Wildemuth, Seungwon Yang
Working in CS Areas with Less Gender Imbalance

1. Information retrieval / digital libraries (closely related to library and information science, move toward iSchools)

2. Human-computer interaction (including social networks and informatics)

3. Collaboration with health, medical, biological and bio-informatics areas (+Reiki)

4. Collaboration with humanities and social/behavioral sciences (e.g., NLP / computational linguistics)
Current Funded Grants

1. IMLS LG: Opening Books and the National Corpus of Graduate Research, Bill Ingram, Jian Wu (ODU), E. A. Fox
2. IMLS RE: Continuing Education to Advance Web Archiving, Zhiwu Xie, E. A. Fox; GWU, LANL, IA, ODU, U. Waterloo
5. NSF I-Corps: Automated Summarization Technology, E. A. Fox, Naren Ramakrishnan
6. NSF IIS: Global Event and Trend Archive Research (GETAR), VT: E. A. Fox, Chandan Reddy, Andrea L. Kavanaugh, Donald J. Shoemaker; Internet Archive: Jefferson Bailey
INFORMATION

Design of

Access Extraction Representation Retrieval Systems Technology Theory Viz

Multimedia

Images

Videos

Database

 Tables

Multimedia

Search Engine

Crawling

Analytics

NLP

Relational

Mining

Deep Learning

AI

Statistics

Library

Archives

Hypermedia

Hypertext

Search Engine

Webpage

Machine Learning

Deep Learning

Data

Knowledge
Information Life Cycle

- Authoring
- Modifying
- Organizing
- Indexing
- Storing
- Retrieving
- Distributing
- Networking
- Using
- Creating
- Accessing
- Filtering
- Retention / Mining
- Accessing
- Filtering
- Utilization
- Inactive
- Active
- Semi-Active
- Searching
- Creation
- Social Context
## Infrastructure Services

### Repository-Building

<table>
<thead>
<tr>
<th>Creational</th>
<th>Preservational</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquiring</td>
<td>Conserving</td>
</tr>
<tr>
<td>Cataloging</td>
<td>Converting</td>
</tr>
<tr>
<td>Crawling (focused)</td>
<td>Copying/Replicating</td>
</tr>
<tr>
<td>Describing</td>
<td>Emulating</td>
</tr>
<tr>
<td>Digitizing</td>
<td>Renewing</td>
</tr>
<tr>
<td>Federating</td>
<td>Translating (format)</td>
</tr>
<tr>
<td>Harvesting</td>
<td></td>
</tr>
<tr>
<td>Purchasing</td>
<td></td>
</tr>
<tr>
<td>Submitting</td>
<td></td>
</tr>
</tbody>
</table>

### Add Value

- Annotating
- Classifying
- Clustering
- Evaluating
- Extracting
- Indexing
- Measuring
- Publicizing
- Rating
- Reviewing (peer)
- Surveying
- Translating (language)

### Information Satisfaction Services

- Browsing
- Collaborating
- Customizing
- Filtering
- Providing access
- Recommending
- Requesting
- Searching
- Visualizing
Informal 5S & DL Definitions

DLs are complex systems that
• help satisfy info needs of users (societies)
• provide info services (scenarios)
• organize info in usable ways (structures)
• present info in usable ways (spaces)
• communicate info with users (streams)
Communication Analysis in the Social Interactome

Text Classification

- Multinomial, naïve-Bayes classification considers the count for each feature name in making classifications
- Training the classifier: built a corpus of 150 documents—75 of which were sentences that were clearly indicative of belonging to a success story and 75 of which were sentences that were not indicative of a success story
- Acknowledgements to Victoria Worrall for her efforts on this classifier last semester

Samples of Story Classification

"Since being in recovery I have not been around any drugs or alcohol but if I had to, such as a wedding or something I wouldn’t have a problem saying that I don't drink or I’m in recovery." => success

'Drove very drunk.' => not_success

Network Structures

- Queried the Friendica database to see who the participants wrote text to and who the participants received text from
- Generated graph of the private messaging communication in the lattice social network

<table>
<thead>
<tr>
<th>Lattice Network</th>
<th>Small-world Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>128 participants</td>
<td>128 participants</td>
</tr>
<tr>
<td>22 users in the most connected component</td>
<td>4 users in the most connected component</td>
</tr>
</tbody>
</table>

Abigail Bartolome, Advised by Dr. Edward A Fox
NIH Grant: 1R01DA039456-01
The Social Interactome of Recovery: Social Media as Therapy Development
Acknowledgements to Dr. Chris Franck, Prashant Chandrasekar, Lexie Mellis
Virginia Tech CS 4994, April 2016
ETANA-DL Architecture
DigBase and DigKit

ETANA-DL UNION CATALOG

- Search
- Browse
- Recommend
- Note
- Personalize
- Review
- Visualizations
- Archaeology Specific

DATABASE WRAPPERS

Lahav
Nimrin
Umayri
Hisban
Megiddo
Jalul
...
Data Mapping Framework in a Digital Library with Computational Epidemiology Datasets
S.M. Shamimul Hasan, Sandeep Gupta, Edward A. Fox, Keith Bisset, Madhav Marathe -- Virginia Tech (CS, BI)
IMLS: Developing Library Cyberinfrastructure Strategy for Big Data Sharing and Reuse

3 patterns for Library Big Data Services
Integrated Digital Event Archive and Library (IDEAL) stakeholders

- Affected
  - Victims
  - Victims Family and Friends
  - General Public

- Researcher
  - Analyzed Data Reports
  - Tragedy Updates
  - Personal Event Accounts

- Policy Makers
  - Preventative Laws
  - Coping Resources

- Community Service Workers
  - Tragedy Updates
  - Analyzed Data Reports

- Supporting

CTR Toolkit

Content

React
Support
Reactions
Personal Event Accounts
Preventative Laws
Victim Information
Personal Event Accounts
Coping Resources
Coping Resources
Tragedy Updates
Tragedy Updates
Tragedy Updates
Tragedy Updates
Tragedy Updates
Tragedy Updates
Who is involved in a WMB?

- Fix water pipe
  - Water utility
  - City/town utility
- Traffic
  - Police
- Affected
  - Citizen
- Others …

Lakewood, NJ, June. 2014
West Philadelphia, PA, June. 2015
What Causes Water Main Breaks? Earthquakes (USGS)

Mar. 1 – Apr. 5, 2012
West Village water main break causing traffic problems: The main broke near Greenwich and 14th St. Use numbered lines if possible. #MTA

Rep from Comptroller @scottmstringer's office at scene of water main break on 13th St and 5th Ave. He says doesn't know much besides it's a 'big one.'

Watch out world. I've got a water main break, and the pix stitch app. #oea #oea

Looks like a water main break at 14th and 6th, lots of MTA emergency presence at W 4th and 14th St. Something big. New York! New York!

@watermainbreak (@ Water Main Break) http://bit.ly/OT0eRAFYC

Between my apartment, the NYU bookstore pipe burst, and this Union Square water main break. #NYCity #NewYork #NYU
GETAR Architecture
Global Event and Trend Archive Research

Data/Info/Knowledge

Events

Sources
Searching
Visualizing
Recommending
Selecting
Browsing

Utilization
NLP

Trends

Collections
Model Building

Correcting/Revising
Organizing

Users

Linked Data

Archives

Phases

Source Identification
Analysis

Info Extraction

Geolocation
Classifying

WWW
Twitter
<table>
<thead>
<tr>
<th>Area</th>
<th>Description</th>
<th>Investigators</th>
<th>Courses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Core Research</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis, access</td>
<td>Databases, HCl, information visualization, machine learning, ontologies, statistics</td>
<td>Fox, Franck, Huang, North, Sheetz</td>
<td>BIT 4524, 4544, 4614; CMDA/CS/ STAT3654; CS5764</td>
</tr>
<tr>
<td>Library, information, data</td>
<td>Archives, big data, curation, data management, decision support, exploring, knowledge engineering, searching</td>
<td>Fox, French, Nicholls, Speer, Thomas, Zobel</td>
<td>CS4624, 5604, 6604; FOR3604; GRAD5134</td>
</tr>
<tr>
<td>NLP</td>
<td>Arabic, document analysis, errors, information extraction, summarization, topic identification</td>
<td>Eubank, Fox, Rozovskaya</td>
<td>CS4624, 4984, 5984, 6804</td>
</tr>
<tr>
<td><strong>Applied Research Across Disciplines</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geospatial</td>
<td>Car crashes, data structures, GIS, maps, queries, traffic, tweets, weather and crops</td>
<td>Baird, Lu, Sforza</td>
<td>GEOG1115, 1116</td>
</tr>
<tr>
<td>Simulation</td>
<td>NDSSL: epidemiology, diffusion in networks, planning response</td>
<td>Eubank, Lewis, Swarup</td>
<td>GBCB5874, 7994</td>
</tr>
<tr>
<td>Climate change</td>
<td>Adaptation, biodiversity, conservation, ecology, ecosystems, effects on plants &amp; animals, environment, sea-level rise</td>
<td>Bukvic, Jelesko, Kalkstein, Quinn</td>
<td>GEOG2994, 4974, 4994; PPWS4994</td>
</tr>
<tr>
<td>Economics</td>
<td>Development, families, game theory, Middle East, smart cities, social networks</td>
<td>Ball, Korkmaz, Salehi-Isfahani</td>
<td>ECON3004</td>
</tr>
<tr>
<td>Energy</td>
<td>Green engineering, nuclear policies</td>
<td>Avey, McGinnis</td>
<td>ENGR3124</td>
</tr>
<tr>
<td>History</td>
<td>Globalization, Soviet history</td>
<td>Ewing</td>
<td>HIST2124, 1215, 2124, 3394, 3554</td>
</tr>
<tr>
<td>Innovation</td>
<td>CIE: entrepreneurship, impact of resources, industry collaboration, social and technology based ventures</td>
<td>Junkunc</td>
<td>MGT3064, 3074, 4064, 4094</td>
</tr>
<tr>
<td>Resilience</td>
<td>Concentrations, dependency, disasters, evacuations, planning, policy, relocation, supply chains, urban and regional, vulnerability (incl. in Maasai society)</td>
<td>Bohland, Bukvic, Lawrence, Murray-Tuite, Zobel</td>
<td>CEE5620, 5660; GRAD5134</td>
</tr>
<tr>
<td>Sociology</td>
<td>Crises, global issues, social inequality, social movements, social participation, violence, social networks, communication behavior and effects</td>
<td>Baird, Kavanaugh, Shoemaker, WImberley</td>
<td>SOC2034, 2044, 3304, 3504, 3854, 4354, 4424, 4444, 4764, 5424</td>
</tr>
<tr>
<td>Political Science</td>
<td>National security, world politics, nuclear policies</td>
<td>Avey, Nicholls</td>
<td>PSC11004, 1024, 1034, 2034, 2054, 2064, 3114, 3514-6, 3524, 3544, 3564, 3515-6, 3624, 3634, 3684-5, 3794, 4734, 5254, 5264, 5284, 5384, 5424, 5444, 5464, 5474, 5514, 5524, 5584, 5624, 5634, 5644</td>
</tr>
</tbody>
</table>
Where Can You Fit in CS?

CS Looking Outward:
- Interaction: Games, Graphics, HCI, VR/AR
- Programming: Algorithms, Languages, Problem Solving, Workflows
- Simulation: Agents, Modeling: Epidemiology
- KID: Knowledge, Information, Data: AI, Machine Learning

CS – Looking Inside:
- HPC <-> PC <-> GPU
- Networking
- Programming
  - Algorithms,
  - Languages,
  - Problem Solving
  - Workflows
- Systems
- Theory
Women in CS Faculty

1. Jin-Hee Cho  
2. Hoda Eldardiry  
3. Margaret Ellis  
4. Andrea Kavanaugh  
5. Aisling Kelliher  
6. Wenjing Lou  
7. Na Meng  
8. Tanushree Mitra  
9. Deborah Tatar  
10. Danfeng (Daphne) Yao  
11. Liqing Zhang
Questions?

At least learn: Computational Thinking

Volunteers and students from high school to Ph.D.

http://fox.cs.vt.edu

Feel free to contact: fox@vt.edu