Exploratory Big Data Analytics

Ivo D Dinov

Center for Complexity and Self-management of Chronic Disease (CSCD)
Statistics Online Computational Resource (SOCR)
Michigan Institute for Data Science (MIDAS)
School of Nursing
University of Michigan

www.SOCR.umich.edu
www.MIDAS.umich.edu
University leaders
- Martha Pollack, Provost and EVPA
- Jack Hu, Interim Vice President for Research
- Eric Michielssen, Associate Vice President, ARC
- Brian Athey, Co-Director, MIDAS
- Alfred Hero, Co-Director, MIDAS

Internationally renowned scientists
- Robert Nowak, Engineering, Wisconsin-Madison
- Susan Murphy, Statistics, Michigan
- Kathleen McKeown, Data Sciences and Engineering, Columbia

An introduction to the MIDAS Education and Training Program
- Ivo Dinov, Associate Director, MIDAS
- Erin Shellman, Research Scientist, AWS
- Patrick Harrington, Director of Engineering, WalmartLabs
- Nandit Soparkar, Chief Executive Officer, Ubiquiti

Keynote presentations
- Daniel Goroff, VP, Sloan Foundation
- George Poste, Health Innovation, Arizona State
- Bror Saxberg, Chief Learning Officer, Kaplan
- Kathleen Carley, Computation, Carnegie Mellon
- Jonathan Owen, Operations Research, General Motors
- Ed Seidel, Physics and Astronomy, Illinois
- Kathleen McKeown, Data Sciences and Engineering, Columbia
- Ratna “Babu” Chinnam, Engineering, Wayne State
- Yike Guo, Computing Science, Imperial College London
- Keith Elliston, CEO, tranSMART Foundation
<table>
<thead>
<tr>
<th>Big Data</th>
<th>Information</th>
<th>Knowledge</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Observations</td>
<td>Processed Data</td>
<td>Maps, Models</td>
<td>Actionable Decisions</td>
</tr>
<tr>
<td>Data Aggregation</td>
<td>Data Fusion</td>
<td>Causal Inference</td>
<td>Treatment Regimens</td>
</tr>
<tr>
<td>Data Scrubbing</td>
<td>Summary Stats</td>
<td>Networks, Analytics</td>
<td>Forecasts, Predictions</td>
</tr>
<tr>
<td>Semantic-Mapping</td>
<td>Derived Biomarkers</td>
<td>Linkages, Associations</td>
<td>Healthcare Outcomes</td>
</tr>
</tbody>
</table>

Neuroimaging-Geno-Pheno Development/Aging
Across Time

Across Categories

Multi-Variate

2 Variables

Multi-Variate

Uni-Variate

Many Factors

1 Factor
Demos

• **SOCR Charts** ([http://www.socr.umich.edu/html/cha/](http://www.socr.umich.edu/html/cha/))
  • EDA: Pie Chart (Ring Chart)
  • 3D Chart ([http://socr.umich.edu/html/cha/SOCR_3DCharts.html](http://socr.umich.edu/html/cha/SOCR_3DCharts.html))
  • GIS/Cartography ([http://socr.umich.edu/html/cha/SOCR_Cartography.html](http://socr.umich.edu/html/cha/SOCR_Cartography.html))
  • Motion Charts ([http://socr.umich.edu/HTML5/MotionChart](http://socr.umich.edu/HTML5/MotionChart)), HPI (default) and Ozone data ([http://wiki.socr.umich.edu/index.php/SOCR_Data_121608_OzoneData](http://wiki.socr.umich.edu/index.php/SOCR_Data_121608_OzoneData))

• **Data Dashboard** ([http://socr.umich.edu/HTML5/Dashboard](http://socr.umich.edu/HTML5/Dashboard)), Husain, et al., 2015

• **Brain Viewer** ([http://socr.umich.edu/HTML5/BrainViewer](http://socr.umich.edu/HTML5/BrainViewer))

• **Hierarchical Data Visualization**
  • SOCR Navigators ([http://socr.umich.edu/html/Navigators.html](http://socr.umich.edu/html/Navigators.html))
  • Circos ([http://mkweb.bcgsc.ca/tableviewer](http://mkweb.bcgsc.ca/tableviewer)), AD Data, Moon, Dinov, et al., 2015

• **Cloud Data**
  • Alzheimer’s Disease Neuroimaging Initiative (ADNI), [http://adni.loni.usc.edu](http://adni.loni.usc.edu)
  • Parkinson’s Progression Markers Initiative (PPMI), [http://www.ppmi-info.org/data](http://www.ppmi-info.org/data)

• **Pipeline Workflows** ([http://pipeline.loni.usc.edu](http://pipeline.loni.usc.edu)), PWS, Navigator