Proposed Data Science and Engineering PhD Program

Lee Riedinger
Director, Bredesen Center
Professor of Physics

Russell Zaretzki
Associate Professor
Haslam College of Business

November 21, 2016
PhD program started in 2011 - Energy Science and Engineering - UTK and ORNL

• Interdisciplinary PhD in *Energy Science and Engineering* (ESE)
  – Specialty areas: nuclear energy, bioenergy, energy conservation and storage, renewable energy, distributed energy and grid management, environmental and climate sciences related to energy, transportation science
  – Entrepreneurial and policy components

• PhD program is administered in the UT/ORNL *Bredesen Center for Interdisciplinary Research and Graduate Education*

• Task force formed in August 2015 by Chancellor Cheek and ORNL Director Mason to study need for an interdisciplinary PhD in *Data Science and Engineering* (DSE) - chaired by Lee Riedinger (UTK) and Jeff Nichols (ORNL)
## Task force members - data science and engineering

<table>
<thead>
<tr>
<th>Member</th>
<th>Institution</th>
<th>Position</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeff Nichols</td>
<td>ORNL</td>
<td>Associate Laboratory Director</td>
<td>Computing and Computational Sciences</td>
</tr>
<tr>
<td>Shaun Gleason</td>
<td>ORNL</td>
<td>Division Director</td>
<td>Computational Sciences and Engineering</td>
</tr>
<tr>
<td>Arjun Shankar</td>
<td>ORNL</td>
<td>Team Leader</td>
<td>Computational Sciences and Engineering</td>
</tr>
<tr>
<td>Budhu Bhaduri</td>
<td>ORNL</td>
<td>Group Leader</td>
<td>Computational Sciences and Engineering</td>
</tr>
<tr>
<td>Gina Tourassi</td>
<td>ORNL</td>
<td>Director</td>
<td>Biomedical Sciences and Engineering Center</td>
</tr>
<tr>
<td>Thomas Proffen</td>
<td>ORNL</td>
<td>Division Director</td>
<td>Neutron Data Analysis and Visualization</td>
</tr>
<tr>
<td>Jack Fellows</td>
<td>ORNL</td>
<td>Director</td>
<td>Climate Change Science Institute</td>
</tr>
<tr>
<td>Lonnie Love</td>
<td>ORNL</td>
<td>Group Leader</td>
<td>Energy and Transportation Science</td>
</tr>
<tr>
<td>Mike Leuze</td>
<td>ORNL</td>
<td>R&amp;D Staff Member</td>
<td>Computer Science and Mathematics</td>
</tr>
<tr>
<td>Peter Tortorelli</td>
<td>ORNL</td>
<td>Deputy Director</td>
<td>Materials Science and Technology</td>
</tr>
<tr>
<td>Stacy Prowell</td>
<td>ORNL</td>
<td>Team Leader</td>
<td>Computational Sciences and Engineering</td>
</tr>
<tr>
<td>Lee Riedinger</td>
<td>UTK</td>
<td>Director</td>
<td>Bredesen Center</td>
</tr>
<tr>
<td>Russ Zaretzki</td>
<td>UTK</td>
<td>Professor</td>
<td>Business Analytics and Statistics</td>
</tr>
<tr>
<td>Mark Dean</td>
<td>UTK</td>
<td>Professor</td>
<td>Electrical Engineering and Computer Science</td>
</tr>
<tr>
<td>Chris Cox</td>
<td>UTK</td>
<td>Head</td>
<td>Environmental and Civil Engineering</td>
</tr>
<tr>
<td>John Kobza</td>
<td>UTK</td>
<td>Head</td>
<td>Industrial and Information Engineering</td>
</tr>
<tr>
<td>David Keffer</td>
<td>UTK</td>
<td>Professor</td>
<td>Materials Science and Engineering</td>
</tr>
<tr>
<td>Colleen Jonsson</td>
<td>UTK</td>
<td>Director</td>
<td>NIMBIOS</td>
</tr>
<tr>
<td>Ralph Lydic</td>
<td>UTK</td>
<td>Professor</td>
<td>Psychology</td>
</tr>
<tr>
<td>Vasileios Maroulas</td>
<td>UTK</td>
<td>Professor</td>
<td>Mathematics</td>
</tr>
<tr>
<td>Brent Mallinckrodt</td>
<td>UTK</td>
<td>Associate Dean</td>
<td>College of Arts and Sciences</td>
</tr>
<tr>
<td>Suzie Allard</td>
<td>UTK</td>
<td>Associate Dean</td>
<td>College of Information and Communications</td>
</tr>
<tr>
<td>Carolyn Hodges</td>
<td>UTK</td>
<td>Dean</td>
<td>Graduate School</td>
</tr>
<tr>
<td>Bob Davis</td>
<td>UTHSC</td>
<td>Professor</td>
<td>Governor’s Chair with ORNL</td>
</tr>
<tr>
<td>Joseph Kizza</td>
<td>UTC</td>
<td>Head</td>
<td>Computer Science and Engineering</td>
</tr>
</tbody>
</table>

Wide enthusiasm for an interdisciplinary PhD program in *data science and engineering*, building on model of the existing *energy science and engineering* PhD

---

**Bredesen Center**

for Interdisciplinary Research and Graduate Education
Proposed PhD in Data Science and Engineering

- National focus and need
- State need
- Expanding area of academic research
- Essential for improved discovery and productivity in many sectors of industry productivity, laboratory R&D, and academic research programs

- Proposed areas of research:
  - Applied computing
  - Data analytics
  - Statistics and machine learning
  - Health & Biological Science
  - Environmental/Climate Science
  - Materials Science
  - Urban Systems Science
  - Transportation Science
  - National Security
  - Advanced Manufacturing
Data Science and Engineering curriculum overview

• Core Curriculum - 21 credit hours:
  – Understand and apply modern statistics and data analysis
  – Understand computing and coding skills in serial and parallel modes
  – Understand information storage and retrieval
  – Develop comprehensive foundation in statistics, data mining, machine learning along with the ability to read current literature and implement software tools
  – Have an understanding of other modeling and simulation approaches from applied mathematics

• Knowledge Breadth Curriculum - six credit hours relating to policy and/or entrepreneurship

• Knowledge Specialization Curriculum - deep dive into area of specialization in a domain area

• DSE 599 Seminar: topical seminars in the focus areas of the DSE PhD

72 hours of graduate credit required for PhD, including at least 36 hours of course work for a student with a BS degree, 24 with a Master’s
Bredesen Center current and past PhD students in ESE doctoral program

- Graduate students - total count is now 137, including 25 new ones in August
  - First class started in August of 2011
  - Also 5 part-time grad students from ORNL
  - Students who have left the program - 19

- PhD completed - 11 in ESE, 4 in departments with ESE concentrations

- National fellowship awards - 11:
  - NDSEG Fellowship - Ryan Ginder (2013), Patrick Shower (2014)
  - Hydro Foundation Fellowship - Mark Christian (2013)
  - Fulbright Fellowship - Kassie Ernst (2016)
  - Nuclear Energy University Program Fellowship from DOE - Brianne Heisinger (2016)
Matching the best students with the best faculty

126 Energy Science and Engineering faculty - all are based in departments or divisions - we have no faculty lines of our own:

- 76 based at ORNL
- 47 based at UTK, including 8 Governor’s Chairs
- 3 based at UTIA

ORNL-based faculty become joint faculty with our program.
Expectations of faculty are to engage with the center, sometimes work on a committee, be available for a seminar or general student interaction, and mentor and support a grad student if desired by both parties.

Areas of research of our ESE graduate students:

- Transportation Sciences - 5
- Nuclear Energy - 29
- Energy Conversion and Storage - 14
- Environmental and Climate Sciences - 14
- Bioenergy and Biofuels - 26
- Cross-Cutting Energy Sciences - 7
- Energy Materials - 17
- Renewable Energy - 9
- Distributed Energy and Grid Management - 6
- Computational energy systems - 10
- Energy Geography - 2
‘Diplomacy Lab’ course on the nuclear fuel cycle - policy aspect of ESE doctoral program

- Course this semester in conjunction with U.S. State Department and Baker Center
- State Department chose our course about the nuclear fuel cycle and associated issues of international diplomacy and economics
- ESE 597, section 2: Special Topics Course - “The Costs and Benefits of a Closed Nuclear Fuel Cycle” - 3 graduate credit hours
- Course is led by Andy Worrall, Fuel Cycle Technology R&D Leader, Reactor and Nuclear Systems Division, ORNL
- Andy is giving few lectures about nuclear fuel cycle and students are working on research on technical, political, and economic issues of this topic
  - The research includes assessment of the true costs and benefits associated with reprocessing and adoption of a closed nuclear fuel cycle – in terms of broad economic (long-term energy independence), political (mitigating risks to the overall nuclear sector), technological, and non-proliferation goals

BIG DATA IN PUBLIC DIPLOMACY - Spring 2017 Diplomacy Lab Course - develop a focused, prescriptive approach regarding whether and how big data can be used in providing actionable guidance for public diplomacy
Success from entrepreneurship track in ESE program - formation of small companies by our grad students

• **Nano Elements Source**
  – Evolved from entrepreneurship class last year - ORNL technology for nano-fermentation of quantum dots
  – License deal signed between small company from Chattanooga and ORNL for a laboratory technology - a bioreactor to make nano dots
  – Beth Papanek (finished after 4 years) and Patrick Caveney (3\(^{rd}\) year)

• **Grow Bioplastics LLC**
  – Produce home-grown high-performance plastic and rubber materials based on lignin
  – Tony Bova (3\(^{rd}\) year grad student) and Jeff Beegle (2\(^{nd}\) year)

• **SimPath LLC**
  – Combine DNA synthesis and assembly methods to construct multiple variations of a DNA pathway quicker and cheaper than normal
  – Rob Moseley (3\(^{rd}\) year grad student) and Ben Mohr (2\(^{nd}\) year)
Proposed Data Science and Engineering doctorate

- Plan for a student body size initially around half of Energy Science and Engineering degree program; grow to 20 to 25 new grad students per year
- Operate in parallel to ESE degree
- Leverage supercomputer at ORNL and *big data* expertise there
- Doctoral degrees through UTK with involvement of faculty and students at UT Health Science Center (Memphis) and UT Chattanooga
- Russell Zaretzki, Haslam College of Business, will be DSE director within Bredesen Center
- Should have big impact on Tennessee
Plan for DSE doctorate start

- October 2016 - DSE curriculum approved by Graduate Council
- November 2016 - submit DSE PhD proposal to Graduate School, then UT system, Board of Trustees, and Tennessee Higher Education Commission
- December 2016 - learn that governor has put $6M into his budget proposal
- January 2017 - accept grad school applications for ESE and DSE programs
- February 2017 - interview prospective grad students for both ESE and DSE
- March 2017 - make fellowship offers for ESE and for DSE, if it looks like the legislature will approve the DSE one-time money
- August 2017 - enroll first batch of DSE students - 10 to 15 in first class
Financial model for Bredesen Center

- We have no base funding from the university or state
- Annual operating expense of around $1.7M/year - few staff, operating, interviewing, first 1.5 year stipends ($30K/year) plus tuition for 25 new PhD students
- Require cost sharing when a grad student joins a group at UT or ORNL
- Receive overhead recovery on contracts from ORNL for partial or full support of a grad student - this year $540K of income
- 2010 - State of Tennessee provided $6.2M of one-time money for Energy Science and Engineering PhD program - this covered the differential for five years
- After five years, UTK returns to us the differential from the pool of tuition funds paid for the 137 doctoral students - 80% this year
- Data Science and Engineering PhD program - same model - hope to get $6M of one-time state funds for next fiscal year - this would provide the differential for five or six years