



• **2015** •
TECHNOLOGY
exchange

OCTOBER 4-7
CLEVELAND OH

**DISTRIBUTED BIG DATA & ANALYTICS (DBDA):
INNOVATION WORKING GROUP MEETING**

FLORENCE HUDSON

Senior Vice President & Chief Innovation Officer

EMILY NICHOLS

Innovation Program Manager

DBDA: INNOVATION WORKING GROUP MEETING

AGENDA

- Welcome and Introductions
 - Review of the Collaborative Innovation Community
- Status of Current Plans & Next Steps
- Other DBDA Issues & Innovations
- Closing

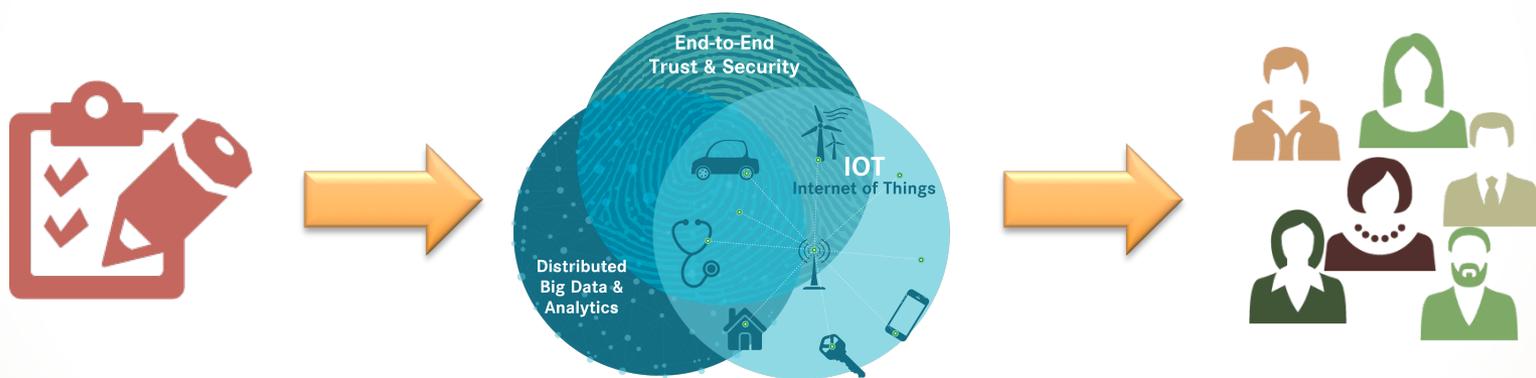


OCTOBER 4-7 CLEVELAND OH



Collaborative Innovation Program

Established three new Collaborative Innovation Working Groups based on March 2015 Member Survey



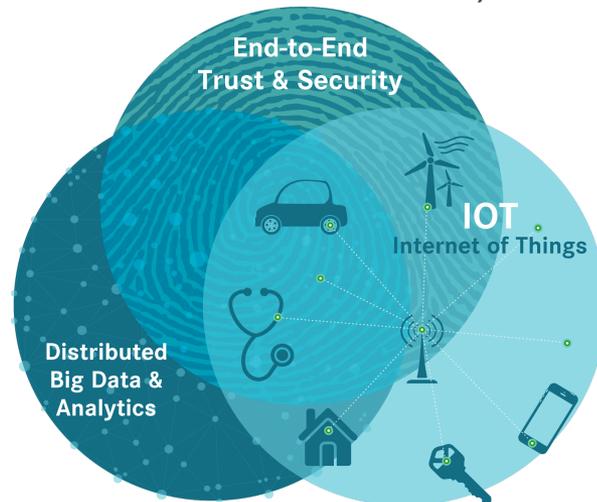
Collaborative Innovation Program Current Focus Areas

E2E Trust & Security:

- End to End Trust and Security for IOT
- TIPS – Trust, Identity, Privacy & Security
- SDP (Software Defined Perimeter), Network Segmentation

Distributed Big Data & Analytics:

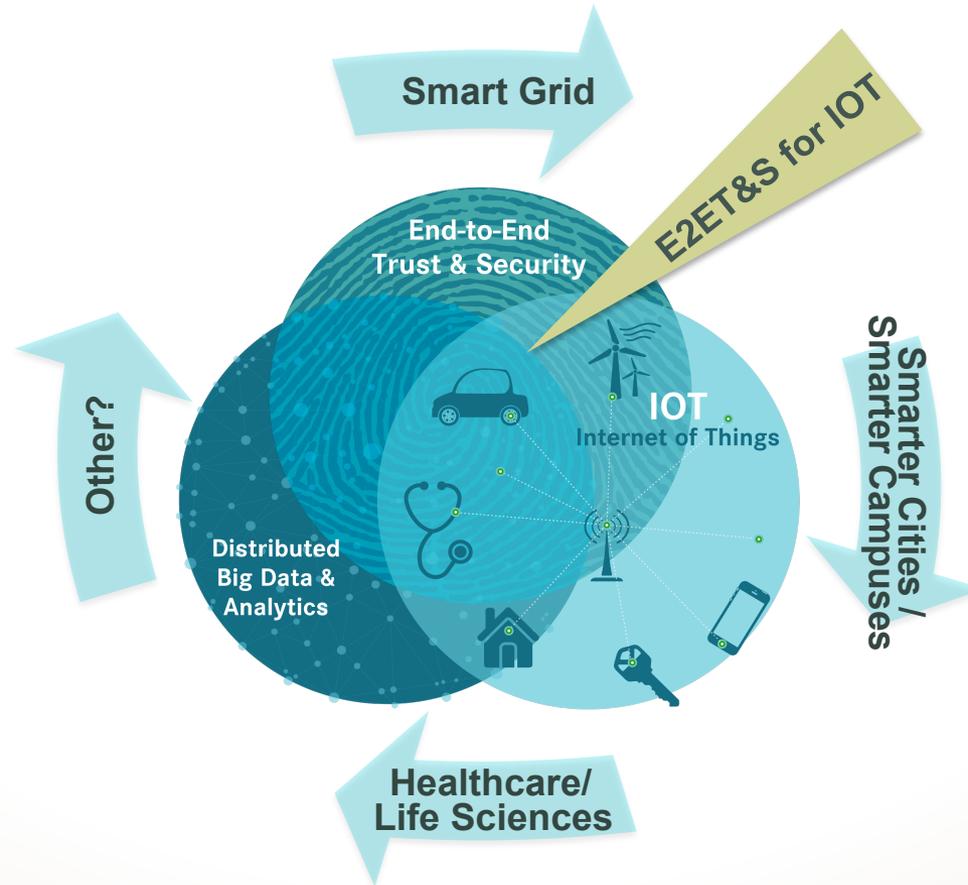
- Genomics
- Smarter Cities / Smarter Campuses
- Digital Humanities



Internet of Things:

- IOT Sandbox
- Smarter Cities / Smarter Campuses
- Smart Grid Testbed

Members Can Participate in Collaboration Opportunities Across the Collaborative Innovation Community Working Groups



Smarter Cities and Healthcare/Life Sciences are beacons of the future economy, and will provide the use cases that bring new applications and technologies to life

Smarter Cities

- Grid
- Campus



Healthcare/ Life Sciences

Collaborative Innovation Program Working Group: Distributed Big Data & Analytics (DBDA)



Co-Chairs

- Alex Feltus, Clemson University
- Sam Gustman, University of Southern California
- Marc Hoit, North Carolina State

75+ Members Representing Universities, Industry, Affiliates, Regional & International R&E Networks

Scope:

- *Enable Internet2 members to better manage distributed big data and analytics through partnerships, collaborations, shared services and leveraging of Internet2 capabilities in advanced networking, cloud services, and federated identity*

DBDA Top Focus Areas

- Thank you to our member universities for submitting use cases – Clemson University, University of Cincinnati, University of Southern California

Focus Area	Description	Plan
Genomics	Confluence of data from genomics, clinical research, clinical care for precision medicine; Includes medical and agricultural genomics, bioinformatics	<ul style="list-style-type: none"> • Develop workshop – perhaps via NIH or NSF grant – to better understand data, analytics, and network needs and potential innovations • Determine workshop participants
Digital Humanities	Video/image libraries (e.g., Shoah Foundation), Global sharing of data and content	<ul style="list-style-type: none"> • Develop workshop to increase understanding of data, analytics, network needs, and potential innovations
Smart Cities	Sharing, security, privacy, and analytics of data generated in a Smart Grid environment, leveraging the Internet2 network as a testbed, to extend to Smart Cities	<ul style="list-style-type: none"> • Align and collaborate with the IOT innovation working group’s Smart Grid efforts

Pain Points in DBDA: Where Internet2 and Its Members Can Develop New Innovations (Page 1 of 3)

- Agreement on Top 3: **Genomics, Digital Humanities, Smart Cities/Campuses**
- Data management: who owns it & defines it, normalizing it; lack of data schemas
 - Lack of standards on the new ways of collecting data: preservation & exportation
- Movement of big data environments: more collaboration (norm) “data movement” Indiana University
 - Collaboration between various universities on the same data set as it changes
- Moving the data: transparent E2E performance tuning; good network engineers are key
- Pittsburgh SC: (1) 10G; run by NSF; IADS standard – how to adopt by various operating standards; improve the experience for the researchers
 - NSF helping (Kevin at NSF), big needs but who will manage (from an HR/bandwidth stand point)
 - Training: More engineers, grad students getting involved/running it
 - Internet2 Broadening the Reach program; high demand (share the 10/5 PPT w/folks)



Pain Points in DBDA: Where Internet2 and Its Members Can Develop New Innovations (Page 2 of 3)

- Researchers/HPC collaboration: COMMUNICATION
 - Governance models needed: Research & IT collaboration, building the trust
 - Regular touch points between researchers & network – staffing needs are great: IT & Librarians
 - Who owns what data: applied physics, digital humanities, engineering, life sciences, librarians
 - Understanding of researcher needs vary & are quite different between physics & digital humanities
 - Penn State: Science DMZ onboarding to better understand data needs & support needs
 - Technology staff not taking time to understand the researcher needs
 - NIST risk management framework standards adoption (can be re-engineered towards HIPAA)
 - Researchers need to commit – most grants are short term
 - Send SC15 survey in conjunction with SuftNet: ask where data is being sent (which SC center being used)
- Alex: Cross Connects Workshop Series: collaboration Internet2 & ESNET
 - Joint researcher engagement

Pain Points in DBDA: Where Internet2 and Its Members Can Develop New Innovations (Page 3 of 3)

- DBDA & Big Data Hub Initiatives grant – not awarded just yet (late Oct)
 - Reference to infrastructure was lacking
 - How to bring the two groups together
 - Researcher focused, not infrastructure focused initiative
- Ed A: Use case: how to best analyze & utilize social media data (digital humanities)
- Internet2 can help build the collaboration with the network as the center
 - Know the data is moving – Analytics of the data transfer (know what’s moving)
 - Network flows to match for partnering opportunities on how data is managed on campus
 - NSF funding opportunities that seek the collaboration on campus
- Workshop participants to get right people at the table for collaboration/partnering around the data management on campus
 - IT that supports research
 - Researchers
- Measurements to determine best practices in data flows: use as tool to get researcher understanding/engagement
 - Intra campus: Internet2; Inter campus: campus IT (not happening regularly)
- Resources connected to Internet2 & ensuring E2E performance tuning
 - Determines problem spots & then can improve data transfers

Closing: How You Can Get Involved



- **Participate in DBDA workshops: Genomics, Digital Humanities, & Smart Cities**
 - Email CINO@Internet2.edu
- **Help your campus be on the leading edge, participate in our new Smarter Campus focus group**
 - Email CINO@Internet2.edu
- **Provide feedback on the Smart Grid white paper**
 - <http://bit.ly/1iJ0N5V>
 - Email CINO@Internet2.edu
- **Join the DBDA Working Group**
 - Email CINO@Internet2.edu
- **Check out our Wiki for more detailed DBDA information:**
 - <http://bit.ly/1iEOojq>

Collaborative Innovation Program



DBDA Co-Chairs

dbda-chairs@Internet2.edu

Florence Hudson

CINO@Internet2.edu

fhudson@Internet2.edu

@FIoInternet2

Emily Nichols

CINO@Internet2.edu

enichols@Internet2.edu





• **2015** •
TECHNOLOGY
exchange

OCTOBER 4-7
CLEVELAND OH

**DISTRIBUTED BIG DATA & ANALYTICS (DBDA):
INNOVATION WORKING GROUP MEETING**

FLORENCE HUDSON

Senior Vice President & Chief Innovation Officer

EMILY NICHOLS

Innovation Program Manager

INTERNET2