

Data Infrastructure Building Blocks for Intelligence and Security Informatics (DIBBs for ISI) for Research and Community

CHALLENGES

In much the same way that the Open Source Initiative, Linguistic Data Consortium, and other organizations build and share resources to further research collaboration in their communities, so too the DIBBs for ISI project seeks to promote collaboration and community among the 4,000+ international members of the intelligence and security informatics community. Our project goals include:

- Building a repository of ISI data of considerable breadth and depth
- Providing analytical, visualization, and access tools
- Reaching out frequently to all potential users and contributors
- Evaluating the repository through regular feedback from users
- Promoting collaboration and community

Upcoming plans and challenges for Year 3 include objectives that continue to focus on **1)** expanding the data available through the repository and enabling more convenient access, **2)** engaging with the community in creative ways, and **3)** providing professional development opportunities for students.

For **Objective 1**, we will double our efforts to obtain data from outside researchers, and we will move our repository to a stable platform for additional functionality (the latter was originally a goal for Year 2). For **Objective 2**, we will continue to propose and develop novel data challenges for conferences such as ASONAM, ICDM, KDD, and IEEE, collectively attended by over 3,000 researchers and students. The challenges encourage innovative research and the sharing and dissemination of results to the wider community. We will also develop an introductory workshop for the Women in Cybersecurity Conference focused on the AZSecure-data portal, with a view toward introducing a new audience to data-driven research and the importance of sharing data. Finally, for **Objective 3**, we will continue to engage our students in research and in writing for publication (for example, our DIBBs project assistant participated in our DIBBs project panel at the recent IEEE Intelligence and Security Informatics conference), and in outreach to outside researchers and in system evaluation. This exposure to multiple project facets provides our students with broad exposure to the research process and with direct knowledge of our users.

The challenges we face in reaching our goals are those faced by nearly every repository project and by every organized effort attempting to build community. Some of those challenges, which vary in size and complexity, and potential solutions are summarized here.

GOAL	SPECIAL CHALLENGES	POTENTIAL SOLUTION
1	Establishing a repository can have a steep learning curve	Select a repository platform that meets users' needs for access, retrieval, and other functions, and hire an experienced developer
1	New challenge: handling sensitive data. Though our data is open-source, some of it does contain personally identifiable information (PID) when we receive it	Establish common-sense policies and best practices in line with other repositories
1	Obtaining data from busy and preoccupied researchers from outside our project	Short-term: Intensive contact over a period of time with selected targeted researchers, via multiple means Long-term: Effect a culture change for ISI researchers to incorporate the future sharing of their data into a conscious part of their research processes
2	Increase efforts to reach 100% of the ISI community	Send the quarterly update to additional conferences, meetings, and other events and listservs where information about security research can be presented
3	Continue to ensure all DIBBs students have professional development opportunities	With project Co-PIs and partners, identify all possible avenues of opportunities among our 4 schools