



IM-PLUS: Information Management with Privacy, Lineage, Uncertainty, and Security

Dr. Barbara Blaustein Dr. Len Seligman

703-983-5881

703-983-5975

bblaustein@mitre.org

seligman@mitre.org

MITRE Sponsored Research

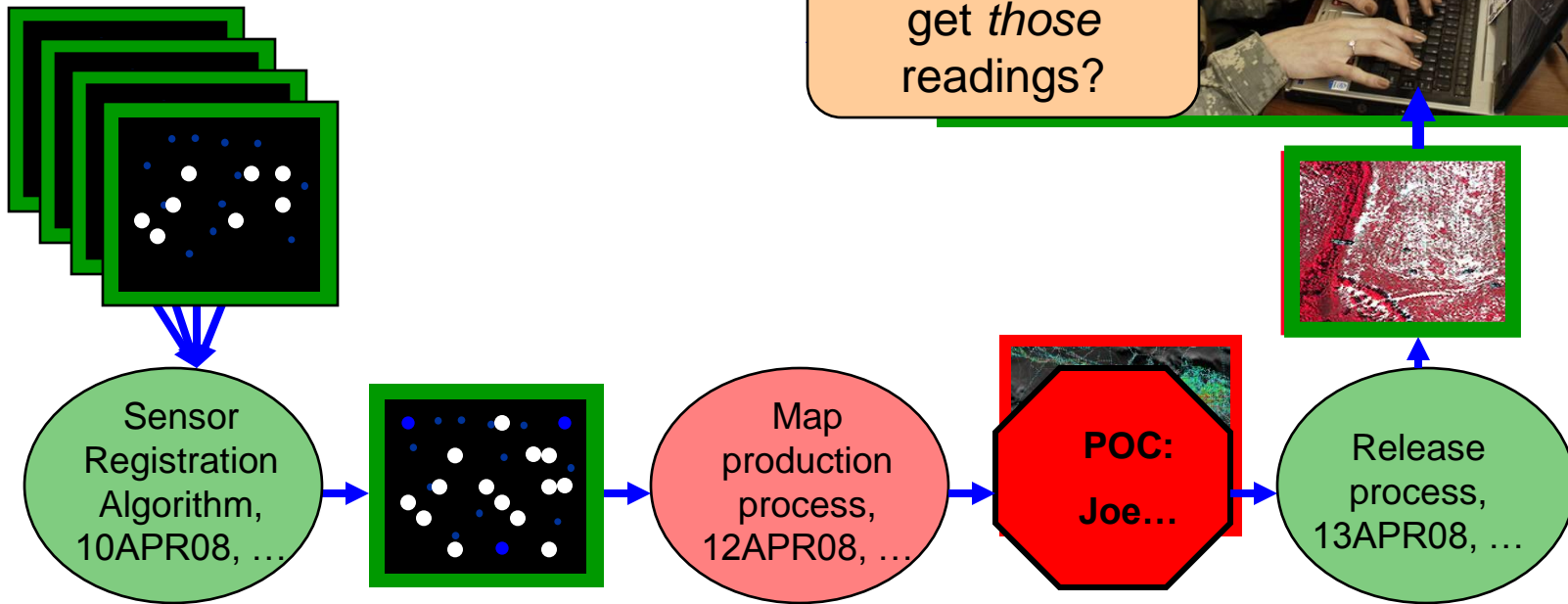
Problem



- **Lineage and uncertainty information improve**
 - **Users' understanding of data from diverse sources**
 - **Mission assurance despite bad data, whether caused by error or cyber attack**
- **How does protection of sensitive lineage and uncertainty annotations affect user understanding and data taint analysis?**

Background

As information is gathered and fused from diverse sources, lineage and uncertainty provide needed context.



Lineage information (processes and data) may have different sensitivities

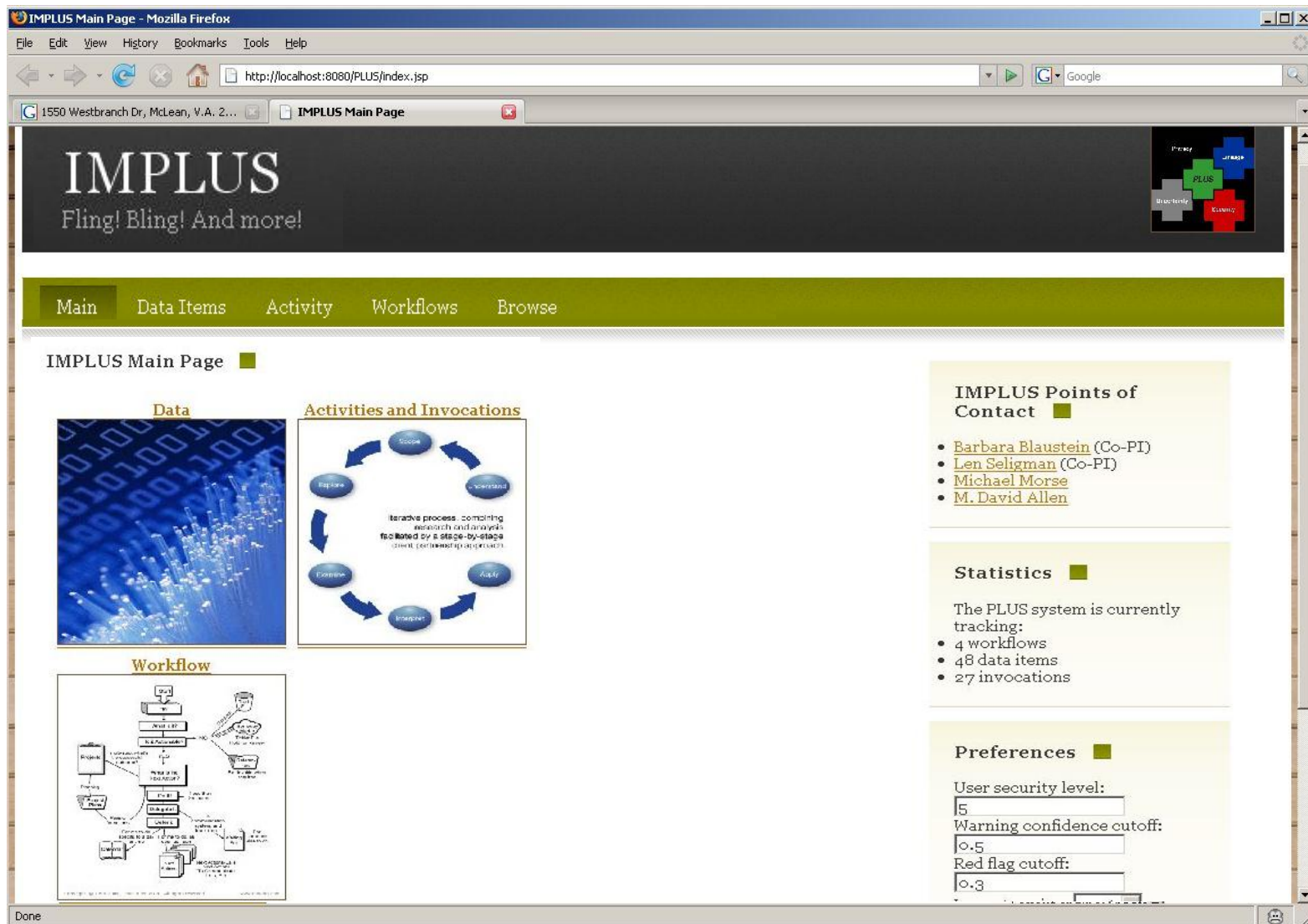
Objective



- **Develop integrated formal model of privacy, lineage, uncertainty, security**
- **Build prototype to understand trade-offs, analyze computational properties, and investigate sponsor requirements**

- **Synthesize and extend existing research to develop formal model:**
 - Lineage in databases and scientific workflows
 - Probabilistic XML databases
 - Developing concept of operations and correctness criteria for handling sensitive data, including lineage
- **Iteratively develop prototype and use cases**
 - Extended Stanford's Trio
 - Developing IM-PLUS prototype
 - Initial scenario inspired by Naval sensor fusion

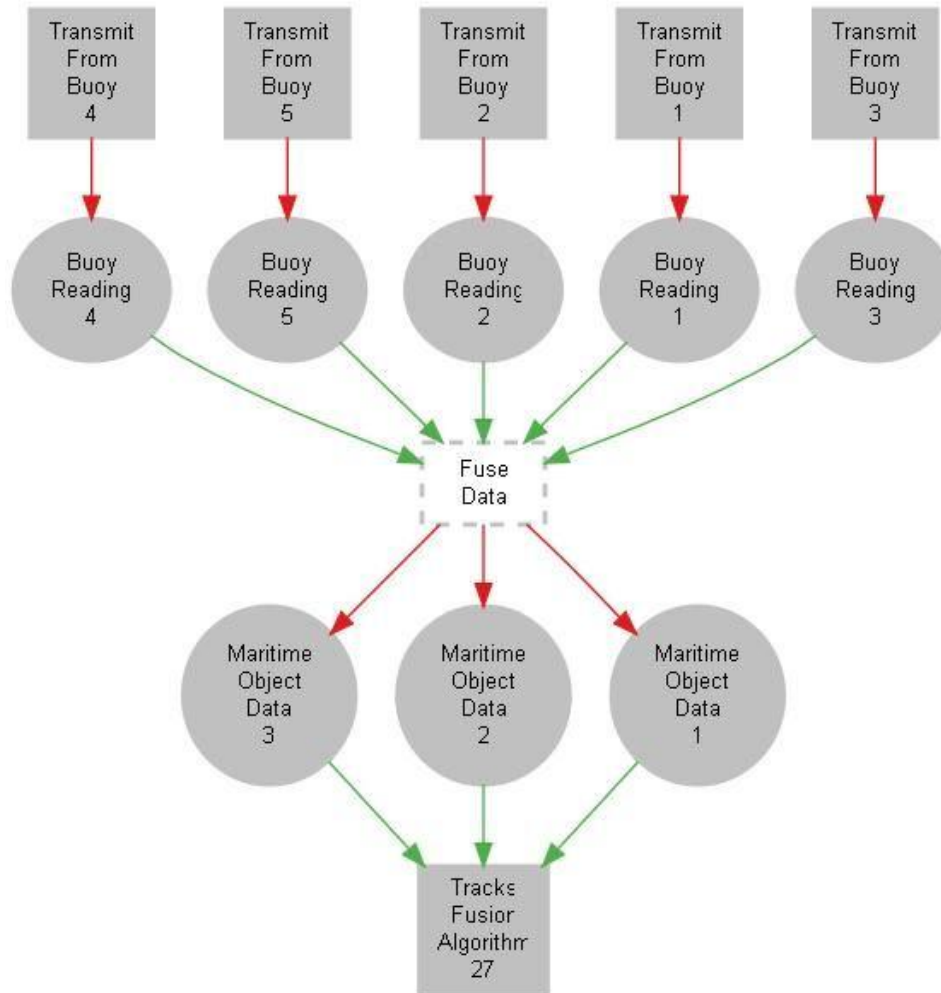
Highlight



The screenshot shows the IMPLUS Main Page in a Mozilla Firefox browser window. The browser's address bar displays the URL `http://localhost:8080/PLUS/index.jsp`. The page features a dark header with the text "IMPLUS" and "Fling! Bling! And more!". Below the header is a green navigation bar with tabs for "Main", "Data Items", "Activity", "Workflows", and "Browse". The main content area is titled "IMPLUS Main Page" and contains several sections:

- Data:** A photograph of fiber optic cables.
- Activities and Invocations:** A circular diagram with six nodes: Scope, Understand, Apply, Interpret, Explore, and Refine. The text in the center reads: "Iterative process, combining research and analysis facilitated by a step-by-step model (see below for description)".
- Workflow:** A complex flowchart diagram showing various steps and decision points.
- IMPLUS Points of Contact:** A list of names and roles:
 - Barbara Blaustein (Co-PI)
 - Len Seligman (Co-PI)
 - Michael Morse
 - M. David Allen
- Statistics:** A section stating "The PLUS system is currently tracking:" followed by a list:
 - 4 workflows
 - 48 data items
 - 27 invocations
- Preferences:** A section with three input fields:
 - User security level:
 - Warning confidence cutoff:
 - Red flag cutoff:

Demonstration



**Navigating lineage
information for
sonobuoy data**

Impacts



- **Advance state of art and influence research agenda in sponsor-critical area**
- **Develop design principles for PLUS systems**
- **Influence sponsor data strategies and acquisitions**

Future Plans

- **Develop additional customer scenarios**
- **Enhance prototype to support federated PLUS capabilities**
- **Publish papers on model and prototype**

