Delivering an Integrated Collaborative Environment with HyperThought™

Matthew Jacobsen
Dr. Charles Ward
Materials & Manufacturing Directorate
Materials and Manufacturing
Organizational Profile

- 800+ scientists and engineers
- 108,000 sq ft lab space, 200 lab modules
- 750+ computers associated with research equipment
- 1000+ computers on desks: 2 separate networks
- 80+ scientific and engineering software packages
- Local computational clusters & remote HPC

And no supporting collaborative research environment
Desired Capabilities

• Enhance existing processes:
  • The coordination and management of research activities
  • The collection of research data (structured and unstructured)
  • Complete traceability of material evolution
  • Integration of data systems

• Introduce new capabilities:
  • Multi-scale, cross-domain materials research
  • Machine learning applications
  • Model-driven research enterprise
An Integrated Collaborative Environment

- The Materials Genome Initiative (MGI) calls for a Materials Innovation Infrastructure, in agreement with the goals of ICMSE
- The research community requires a configurable, federated infrastructure that is lightweight and low risk
- Success requires a joint effort between software and materials engineers to deliver game-changing functionality
The Integrated Collaborative Environment (ICE)

Integration of Capabilities

- Complete experimental and computational material pedigree
- Robust digital workspaces for collaboration
- Integrate with existing systems and equipment
- Offer data management for files, metadata, and datasets

Cross-Domain Applications

- Technology-agnostic integration layer
- Modular and scalable architecture
- Intra-organizational connectivity
- User-tailored to any domain – ICMSE, Digital Thread, maintenance, etc.
Two Pillars

• Integration layer – provides an interface for connecting new and existing systems in a federated manner, along with core collections of microservices.

• Tool suite – provides a user experience that includes content management, digital workspaces, equipment integration, workflow management, visualization tools, and a powerful search engine
Key Tenets

• API-first
• Re-usable Microservices
• Modular
• Deployable
Federated Concept

- Self-governance of connected systems
- COTS tools, in-house developed applications, or a mix
- Systems do not talk directly to each other - ICE “brokers” all transactions between connected systems
Putting the Pieces Together

• ICE Core - Common Service Bus, Group/Project Spaces and Apps: Django, Databases: MongoDB/Marklogic, advanced visualization (Plotly)

• ICE Extended - Material properties database (Granta), MTS Echo, Dream.3D

• Persistent identification, indexed and faceted triple-based metadata, data type registries, and SSO

• Graphical workflow design tools, equipment integration, item management, file management, advanced search tools
Data Retrieval via Search

- ICE repositories and databases
- Internal repositories and databases
- Materials Project
- Granta
- Materials Commons
- Aflowlib

<table>
<thead>
<tr>
<th>Object ID</th>
<th>Object Type</th>
<th>Date Created</th>
<th>Creator</th>
<th>Source</th>
<th>Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>Casting</td>
<td>1/21/2016, 10:21:10 AM</td>
<td>Workflow</td>
<td>ICE.forms</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Casting</td>
<td>1/21/2018, 12:30:24 PM</td>
<td>Workflow</td>
<td>ICE.forms</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Object ID</th>
<th>Object Type</th>
<th>Date Created</th>
<th>Creator</th>
<th>Source</th>
<th>Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>GeometryCondition</td>
<td>1/21/2016, 10:30:47 AM</td>
<td>Workflow</td>
<td>ICE.forms</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>GeometryCondition</td>
<td>1/21/2016, 10:30:52 AM</td>
<td>Workflow</td>
<td>ICE.forms</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>GeometryCondition</td>
<td>1/21/2016, 10:30:52 AM</td>
<td>Workflow</td>
<td>ICE.forms</td>
<td></td>
</tr>
</tbody>
</table>
Equipment Integration

• Connecting Experimental Equipment to workflow
• Set up equipment and connection parameters
• Define process on the canvas
• Stream data to monitoring dashboard
• Gather/parse data via staging/metatron
Success Stories

- Huge improvements in basic data/resource management
- Integration of numerous characterization labs and repositories (Granta, etc.)
- Automated ingestion of run data and metadata
- Deployment of ICE at several off-site locations
- Full material pedigree now available
Current Activities

• NoSQL Integration – Core services, replicative stores of attached systems (Granta, etc.), parsed run data and metadata

• Search Integration – reaching outside of the organization to reference databases and other in-house systems

• Lab and Equipment Monitoring – collaborations with vendors and developing open micro-service models

• Modeling and Simulation integration
Lessons Learned

• Developmental efficiency via API-first design
• Engage actual users (early and often)
• Start small, fail fast
Material Data Infrastructure Integration Summit (Hackathon)

2-weeks of design, development, and problem solving!

Objective: Produce working code that demonstrates meaningful additions to ICME use cases

- State of the art review
- System and API design patterns
- Making sense of objects and results (AKA schemas and semantics)
- Exploring Peer Data Management Systems
- Authentication protocols
- Modeling and simulation
- Enable exchange and reuse of data transformation tools
- Design and integration of microservices registry
- Develop a toolset for the creation of reusable schemas and object types
- Retrieving, Assembling, and Distributing pedigree data
- And much more…