From Human Brain Project to EBRAINS RI: coordinating the technical aspects of the transition

Amaryllis Raouzaiou, Evita Mailli
Athena RC

HBP Partnering Projects Meeting: Status quo & outlook
5-7 September 2022 | Nijmegen, The Netherlands
What is Technical Coordination (TC)

EBRAINS INFRASTRUCTURE PLANNING, COORDINATION AND IMPLEMENTATION

- planning, design, integration and delivery of the EBRAINS RI
- establishing common software quality principles and indicators
- monitoring development, integration, testing, and delivery of components
- ensuring the components’ efficient integration and sustainable operation
Before SGA3

- Many research teams, important – independent results
- At the beginning of SGA3 many components, tools, services, no connection between them

creation of a comprehensive, E BRAINS RI components catalogue
SGA3

- Working Groups
- Documentation
- Co-design process-collaboration with:
  - component owners
  - use cases
  - scientific WPs
- Roadmap created from the beginning, adapted periodically
- Collaboration with HBP bodies
TC weeklies

ISSUES

90 closed since last review
15 in progress
25+ open issues

WEBINARS

5 Presentations
Q&A section
Recordings
TC activity

- 12 TC-TF
  - TC planning
  - Integration guidelines
  - Workflows
- 200+ meetings
- Ad-hoc meetings and presentations

-EBRAINS events
  - Summit 2021, 2020
  - Codejam 2021, 2020
- Workshops and Educational events
  - TVB-EBRAINS workshop
  - Young Researchers Event

- TC Resources
  - Guidelines
  - Common tools and environments
  - Standardised processes
### SGA3 – Integration Phases

<table>
<thead>
<tr>
<th>PHASE 1</th>
<th>M4 — M12</th>
<th>output MS5.1 PoC EBRAINS RI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>• 4 components integrated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Guidelines for integration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Collaboratory/Not ebooks</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PHASE 2</th>
<th>M13 — M21</th>
<th>output MS5.2 Beta EBRAINS RI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>• ~50% components integrated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Updated guidelines</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Formal EBRAINS APIs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Workflows</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Provenance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PHASE 3</th>
<th>M22 — M33</th>
<th>output RC EBRAINS RI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>• All EBRAINS components</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Testing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PHASE 4</th>
<th>M34 — M42</th>
<th>output EBRAINS RI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>• Finalisation</td>
</tr>
</tbody>
</table>
Integration

EBRAINS Integration = technical integration

**INTEGRATION STEPS**

- Components’ assessment
- Components’ prioritization based on their maturity
- Checklist available at the beginning of every Phase
- Collaboration with component owners
At the end of Phase 2

### Meeting Phase 2 Integration Requirements

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code repositories</td>
<td>50%</td>
</tr>
<tr>
<td>CI runs</td>
<td>90%</td>
</tr>
<tr>
<td>TC Collab wiki page</td>
<td>60%</td>
</tr>
<tr>
<td>Deployment</td>
<td>80%</td>
</tr>
<tr>
<td>EBRAINS Docker registry</td>
<td>85%</td>
</tr>
<tr>
<td>Centralized Authentication/Authorization</td>
<td>85%</td>
</tr>
<tr>
<td>SGAs Acknowledgment</td>
<td>60%</td>
</tr>
<tr>
<td>Monitoring &amp; Observability</td>
<td>62%</td>
</tr>
<tr>
<td>Licensing</td>
<td>100%</td>
</tr>
<tr>
<td>Documentation</td>
<td>100%</td>
</tr>
<tr>
<td>Knowledge Graph entries</td>
<td>60%</td>
</tr>
<tr>
<td>HLST point of documentation</td>
<td>80%</td>
</tr>
</tbody>
</table>
Workflows in EBRAINS

Series of non-interactive EBRAINS tools and services, linked together to create graphs, loops or branches defined as structured, common recipes, executed via compatible software.

Environments

Offerings
- Collaboration
- Code sharing
- Data sharing
- Documentation
- Launch HPC jobs

Limitations
What is missing
- Reproducibility
- Shareability
- Portability
- Findability

Goal
- Standardised
- Automated
- FAIR
Standardisation

EBRAINS

FAIR Data KG + FAIR Tools KG TC + FAIR Workflows TC SLU = FAIR e-Infrastructure

EBRAINS users

- Reproducibility and Reusability
  - well-documented and structured recipes
- Flexibility
  - no need for familiarization with technical execution details
- Shareability and Association
  - via Knowledge Graph

TC: Technical Coordination
KG: Knowledge Graph
SLU: Scientific Liaison Unit
# Common Workflow Language (CWL)

Open, common, standard format to describe data analysis & simulation workflows as recipes

<table>
<thead>
<tr>
<th>EBRAINS tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>Defined</strong> via CWL</td>
</tr>
<tr>
<td>• <strong>Bundled</strong> via Containerization methods/Package managers with dependencies, libraries and binaries</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Workflows</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>Defined</strong> via CWL as recipes</td>
</tr>
<tr>
<td>• <strong>EBRAINS tools</strong> - graphs, loops and branches for scientific objectives</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>So far…</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Workshops, Summit, Codejam</td>
</tr>
<tr>
<td>• <strong>Familiarizing</strong> users with standardization methods and technologies</td>
</tr>
<tr>
<td>• Workflow management systems installation</td>
</tr>
<tr>
<td>• TC-SLU-Showcases co-design / co-planning</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ongoing</th>
</tr>
</thead>
<tbody>
<tr>
<td>• <strong>UNICORE – CWL</strong> compatibility</td>
</tr>
<tr>
<td>• <strong>Workflow Registry</strong> via Knowledge Graph</td>
</tr>
<tr>
<td>• <strong>Dashboard</strong> for submitting, monitoring, fetching results and logs</td>
</tr>
</tbody>
</table>
Workflows Lifecycle in EBRAINS

Create/Find/Compose

CREATE
- Graphical CWL editor (Rabix composer)
- Support for all code editors
- Workflow creation environment with visual and code editor

or FIND
- Knowledge Graph
- external Repositories

Execute

ON EBRAINS INFRASTRUCTURE
- HPC systems via UNICORE
- OKD cluster via standardized API for Task Execution
- GUI for submission, execution and monitoring

OUTSIDE EBRAINS
- Any system via compatible Workflow Engines
- Other Research Infrastructures via compatible Software

Share/Publish

WORKFLOW REGISTRY
- Workflows, results and metadata stored in the KG
- Findable, Accessible, Reproducible for EBRAINS users and scientific community
- Associated with research papers, publications
End of SGA3

- EBRAINS RI is operated by EBRAINS AISBL with the participation of National Nodes
- ESFRI
- European research projects
Main benefit from TC

- A platform, not just a catalogue of tools, with a single-entry point
- Collaboration between “scientific” and “technical” WPs
- Collaboration of Showcases and TC
What is next?

- We need more use cases, different types of users
- Show that research activity becomes easier, quicker, more expressive, more accessible on mature platform + ecosystem, compared to wrangling tools & configs manually.
- Difficult but crucial to keep distilling "user experience" (UX) feedback from scientific "target audience".
- Provide reusable scientific workflows as productivity shortcuts & even micro-workflows for fast composition of high-level workflows.
- Your input will tangibly influence creation of the platform you want.

Contact us: hbp-tc@athenarc.gr
Thank you

www.humanbrainproject.eu  www.ebrains.eu