ICM Centre for Neuroinformatics

Challenges and opportunities in the era of open science

Stanley Durrleman
<table>
<thead>
<tr>
<th>Molecular &amp; cellular biology</th>
<th>Neurophysiology</th>
<th>Cognitive neurosciences</th>
<th>Clinical and translational research</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 research teams</td>
<td>6 research teams</td>
<td>5 research teams</td>
<td>14 research teams</td>
</tr>
<tr>
<td>Multiple Sclerosis, brain tumors, Neurodegenerative diseases</td>
<td>Epilepsy, neuronal and muscular excitability</td>
<td>Motivation disorders, psychiatry</td>
<td>Alzheimer’s disease, Parkinson’s disease, prion disease, brain tumors</td>
</tr>
<tr>
<td>Sequencing and genotyping, cellular imaging, histology</td>
<td>Electrical recordings in humans, animal models, cellular culture</td>
<td>Behavioural and cognitive assessments, virtual reality platform, living lab</td>
<td>Clinical investigation center, functional exploration platform, therapeutics trials</td>
</tr>
</tbody>
</table>

...
Enable a digital transformation of neuroscience research

**Our Objectives**

- **Break down barriers between domains**
  - Share expertise, tools and data
  - Enable transversal multidisciplinary research programs

- **Increase reproducibility of our findings**

- **Value research on retrospective data**
  - Enable easy test of new ideas and hypotheses
  - Enable data-driven research
    - data mining, statistical learning, computational modeling
    - Technological development (decision support systems, IoT, etc.)

- **Make neuroinformatics a distinctive aspect in the international scene & ease collaboration**
Our strategy

- Creation of a Centre for Neuroinformatics
- Provide hardware and software infrastructure for scientific computing and data science
- Create a common data lake and data catalog
- Create a common portfolio of software tools
- Promote a culture of share (community tools, open data, etc..)
The ICM Centre for Neuroinformatics

- A virtual centre: a distributed model
  - Coordinate research & development effort in data management and analytics
  - Federate an open community of developers, engineers, data scientists

- A dedicated team and budget

- Event, training and communication:
  - Internal training sessions (10+ sessions/year: software engineering, High Performance Computing, statistics, etc.)
  - Monthly meet-ups
  - Scientific conferences (e.g. NOW)
  - Hackathons (e.g. Brain Hack Network)
Centre for Neuroinformatics
Coordinator - Stanley Durrleman

IT Department

Data Management & Analytics
iCONICS

Administration IT
Infrastructure
Scientific IT

Data Management
Database Development
Software Development
Biostatistics Analyses

Provide services from consulting to end-to-end studies

Contribute to a common portfolio of expertise, tools, methods, data

Domain-specific expertise
- Molecular & Cellular Biology
- Neurophysiology
- Cognitive Neurosciences
- Clinical and Translational Research

Core Facilities (data acquisition)
Teams (research projects)
A common **data lake** and **data catalog**

- **Built-on community tools in each scientific domain**

- **Multi-modal integration:**
  - Indexation (meta-data) for querying the data lake
  - Data warehouse
  - Use of common terminologies, ontologies
  - Common procedure for data de-identification
Unique de-identification procedure for ICM: technical and regulatory studies in progress
A common portfolio of software tools

- **An institutional gitlab repo**
- **Support in software engineering**
  - Optimization, parallelisation
  - GUI Development
- **Automation of data processing**
  - Built on community tools in each domain (e.g. Nypipe, DAX, Clinica in neuroimaging → tomorrow hands-on session)
Our challenges

• **Change management**
  • Work hands-in-hands with teams and core facilities
  • Invest lot of efforts in internal communication
  • Focus on projects with high value for the teams

• **Governance of the data lake**
  • Promote a culture of sharing with adherence of most ICM scientists
  • Motivation and rewards in the era of open science
  • Positive regulatory incentives (GDPR, FAIR, etc..)
Our challenges

• **Link with partners (AP-HP, networks, etc..)**
  - Heterogeneous level of maturity
  - Need a share strategy and technological choices

• **Ressources**
  - Core funding with institutional money: how to measure RoI?
  - Convince PIs to secure resources for neuroinformatics in grant applications
• Development of Open Science is still often done by individual researchers or teams, in fewer occasions by organisations

• Open Science is a powerful tool to transform research organization

• It needs a careful implementation to ensure the commitment of the teams and the executives