

CARMEN

ENHANCING COLLABORATION
IN NEUROSCIENCE

What can be learned from
the UK CARMEN project

Presenter: Leslie Smith
l.s.smith@cs.stir.ac.uk

Content

- Brief introduction to the CARMEN project
 - History, current status.
- CARMEN and the Future
 - What has the CARMEN group learned?



CARMEN

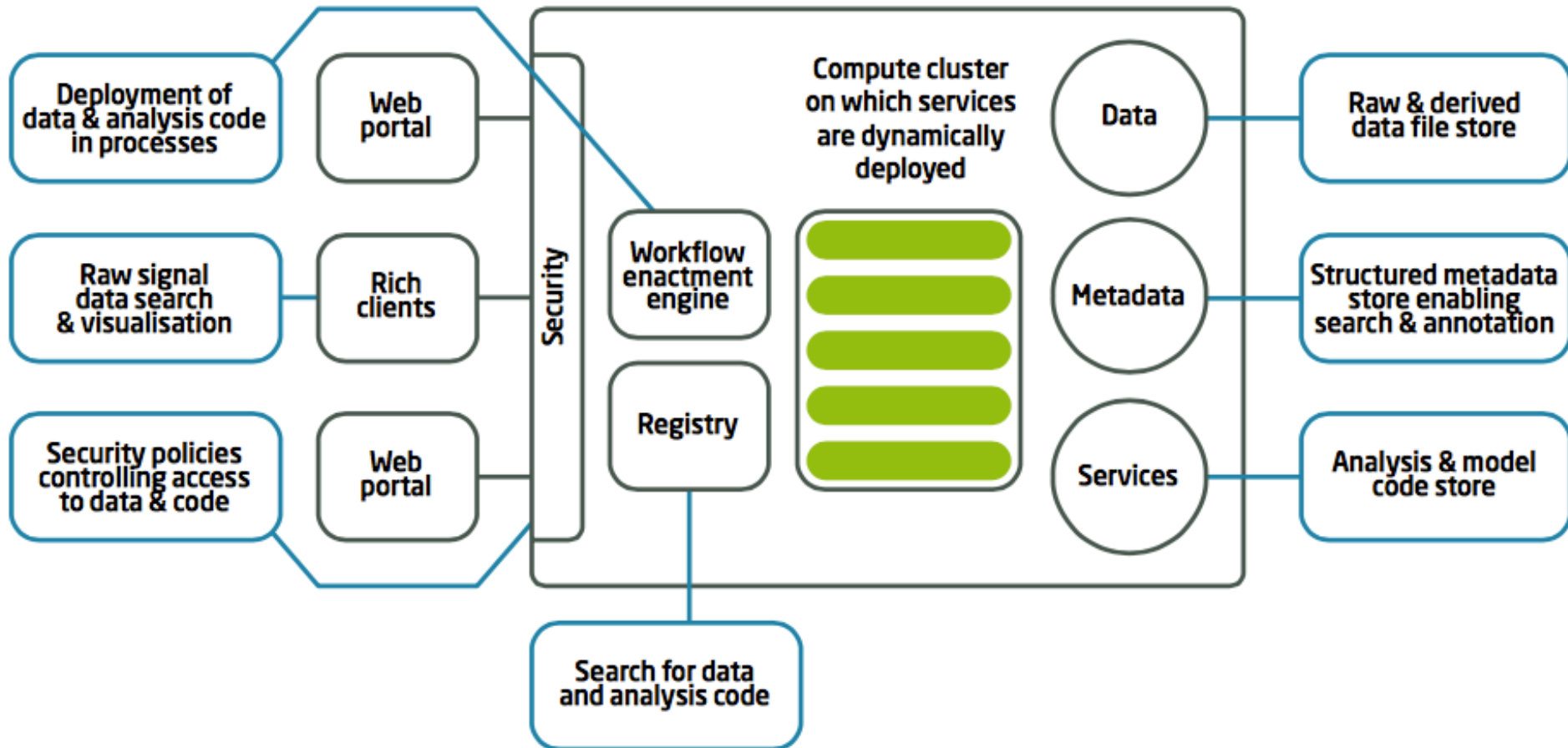
CODE ANALYSIS, REPOSITORY & MODELLING FOR E-NEUROSCIENCE

- CARMEN
 - Code analysis, repository and modelling for e-neuroscience
- ... is a project set up to provide a portal-based system allowing
 - Storage of time-series data (electrophysiological datasets)
 - Analysis of these datasets by services and workflows at the portal
 - Visualisation of datasets
- The concept was that researchers would use the portal to share data and analysis techniques
 - With collaborators (*a co-laboratory*)
 - With the research community in general (making them public)
- Funding for the project started in October 2006.

CARMEN history

- UK EPSRC funding 2006-2010
 - Jointly to 11 UK Universities
- UK BBSRC follow-on (tools and techniques) funding 2010-2015
- ...but it's run out of time and money now
 - Though there are ongoing related projects
- The portal is working.
 - <http://portal.carmen.org.uk>

CARMEN architecture



Contact support

Data Services Workflows Publications

Query

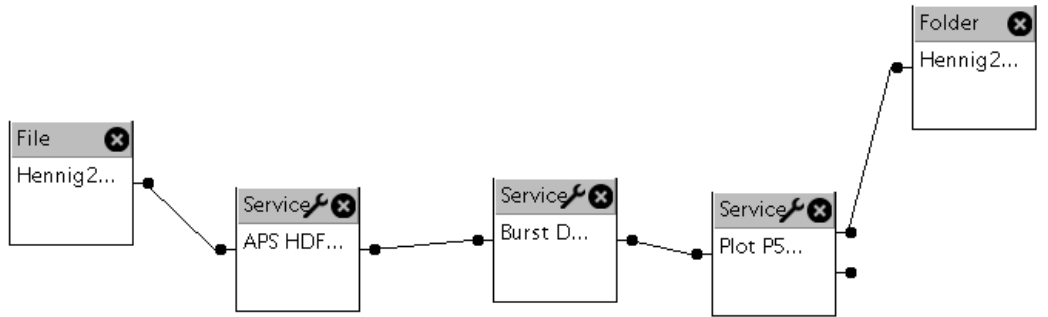
- Data
 - Public
 - Groups
 - L.s.smith
 - Shared

Upload

Annotation Metadata Visualise Sharing Download Run

Close Workflow Save Workflow

Workflow Name Quick Retinal Waves ana Workflow Description Burst duration for sponte



CARMEN: very much in use ...

NEWS

WARNING – UPLOADS AT RISK

🕒 MARCH 23, 2015

Storage on the CARMEN system is nearing it's full capacity so it would be greatly appreciated if users could limit their uploads to essential items and utilise their own local storage where available. Plans are in place to upgrade the hardware but there is a risk in the coming weeks that we may have to consider halting future uploads in order to maintain system stability. As soon as we know a date for the upgrades we will let users know. In the meantime where the storage service is business critical we would recommend alternative provisions be sourced in the unfortunate event that we do have to halt uploads.

Regards,
CARMEN team

Where now for CARMEN?

- It's up and running, and being used: it's almost full
 - But the portal is about to be transferred to a new system
 - Which includes more storage space
- But:
 - to expand it, or to continue it effectively needs a lot more work
 - Some of the equipment running it is getting old
 - Porting it is relatively straightforward, due to use of virtualization technology
 - Work is underway to port it to...
 - Some of the design is out of date
 - Internal parallelism is missing
 - Use of Java applets for uploading large data volumes

What to do with an old portal?

1. Give up on it

- It was a good idea at the time, but actually
 - It's too hard
 - (Neuro)-scientists don't really want this sort of thing sufficiently
 - Infrastructure is too hard to keep running, so don't try.

2. Mothball it

- Keep it running (get some more disk space first!), and hope that the University where it is sited will let it be looked after along with other servers

3. Look to expand it in an appropriate fashion

- Get more funding, to update it effectively
 - (In CARMEN's case: Jointly with HBP? Or CERN?)

CARMEN and the future: What has been learned?

Client pull and technology push?



Client pull: the users...

- Who are the target users?

- **Neuromodellers**



Human Brain Project

- **Data to constrain and test models**
-

- Clinical neurologists and neuroscientists

- Epilepsy, traumatic injury, Parkinsonism, ...

- Neuropharmacologists

- Assessing effectiveness of neuro-active pharmaceuticals

- Experimental neuroscientists

- In Universities and research laboratories.
-

- Educators

- Training the next generation of neuroscientists
(computational, clinical and experimental)

Making the system user-focussed

- What do prospective users want ?
 - 5 groups identified in last slide
- What do they need ?
 - What is the problem the system is trying to solve?
- What will they actually use
 - As opposed to what they say they might use?
 - Or what might be easy to provide (“low-hanging fruit”)
- How can the system be made attractive and straightforward enough for neuroscientists to use?
 - What are the issues that discourage users?

What do the users want:

- “Ease of use”
 - Not simply a question of good UI design!
 - May be different for different communities
 - E.g. different for modelers, for experimental neuroscientists, for educators
- Speed of response
 - All users are used to fast response from desktops, laptops etc:
 - Portals may be slower: User interface needs to manipulate users expectations
 - Particularly when large volumes of data are being processed
 - Must make it clear that something is happening!

For clinicians and experimenters:

- Data upload requirements - make it easy:
 - Single data entry
 - Preferably automated, from existing equipment
 - Single metadata entry
 - Preferably from same equipment and at same time as data
 - E-Lab book
 - Or at worst, entered into a single system once
- Data accessing:
 - Easy to find data, easy to manipulate data, easy to work with complex datasets
- Clear policies on crediting data & service creators when their data/metadata/services are re-used.

For neuroscience modellers

- Good data to constrain their models
 - And lots of data, particularly for complex models with many degrees of freedom
 - High density multi-electrode data (for example)
 - Perhaps as spike trains
 - Perhaps as LFPs
 - Must be easy to discover the right data through searchable metadata
- Different types of data for different types of model
 - Models may be at the membrane level, at the neuron level, at the multi-neuron level, ...
- Ability to set up analysis and modelling workflows
 - And use them on multiple datasets from many labs
 - Implies data format compatibility
- Data (and model output) visualisation
- How can CARMEN and its successors best provide this?

Technology push

- What technologies might be helpful?
 - Uploaders of data and computational modellers need systems that are modern...
 - But they must work across a range of target client systems
- Handling and remotely visualising large datasets
 - High bandwidth links
- Bringing the tools to the data
 - Because the data is too large to move quickly or easily
 - But making this effective and easy to use is difficult
 - Needs effective portal techniques
 - Designed for the users, providing high quality of user experience
 - Linked to other resources (e.g. NIF, ontologies, model repositories)

Technology Push continued

- Search technologies
 - Rapidly searching metadata, services, workflows (including modeling tools).
 - Direct data searching?
 - Searching for patterns in the data
- Parallelism
 - At the user level (multiple simultaneous users)
 - At the processing level
 - Parallel-ising models,
 - Working with Multiple datasets,
 - Parameter searching through large spaces
 - But not necessarily super-computers!

What organisations are relevant?

- Human Brain Project
 - Note also their relationship to the EU Horizon 2020 infrastructure project <http://ec.europa.eu/programmes/horizon2020/en/h2020-section/european-research-infrastructures-including-e-infrastructures>



- International Neuroinformatics Coordinating Facility (INCF)

Strategic Objectives

1. *Partner with international stakeholders to promote and prioritise neuroinformatics at global, national and local levels*

- Research Data Alliance

– <https://rd-alliance.org>



- GÉANT

– <http://www.geant.net/Pages/default.aspx>



- ELIXIR, EUDAT



- Link to major brain projects outside Europe



What to keep from CARMEN

- Concept of a co-laboratory
- Common data format
 - Internal, not necessarily external as well. Self-describing
- Effective privacy system
 - Small and large-grained privacy system with individuals, groups, etc.
- Importance of good user interface
 - Has to be easy to use
- Ability to eyeball data quickly

What to add to CARMEN

- Better user interfacing
 - Better more easily organised workflows
- Better eyeballing of data and results
 - Easier data manipulation
- Better/easier data manipulation
- As much open-ness as possible
 - Use of public code repositories, involvement of the community in code development
- Large volumes of open public data, services and workflows
 - For training both of prospective users, and of (e.g.) doctoral students.

Concluding remarks

- Modelling neural systems requires complex models with many parameters
- Constraining complex models requires large volumes of data
 - And so requires large-scale infrastructure projects
- Infrastructure projects
 - and that include model-platforms that use large volumes of data
- ... need to have a long lifetime
 - So sustainability is very important.
- Modelling platforms that use real data are important educating tools for the next generation in Neuroinformatics
 - Need to be have appropriate data and services, and good tutorial materials to make this happen.