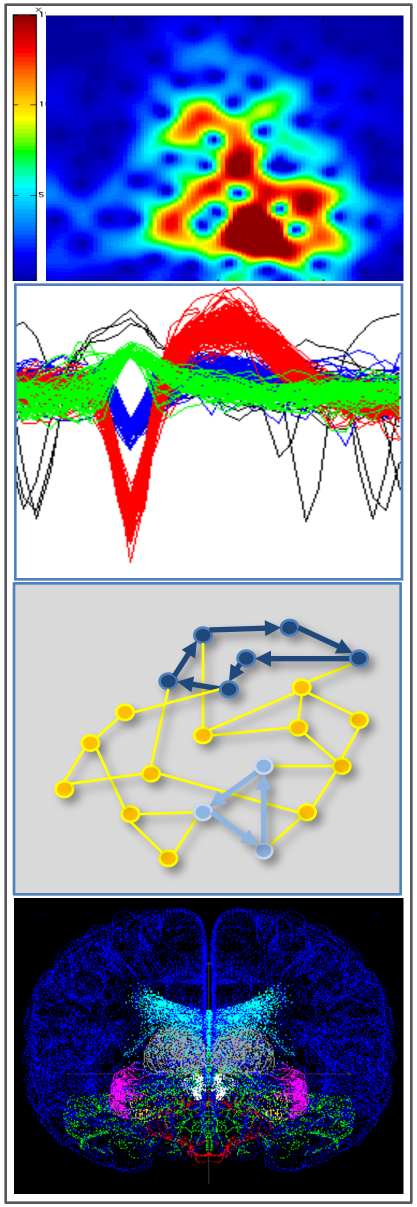


Neuroinformatics

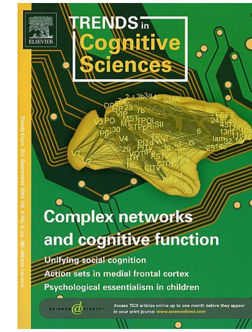
Marcus Kaiser

Week 1: Course Overview

www.dynamic-connectome.org

neuroinformatics.ncl.ac.uk

Brief introduction (Dr Marcus Kaiser)



2002 MSc Biology, Computer Science
Bochum University, Germany

2005 PhD Computational Neuroscience
Jacobs University Bremen, Germany

2007 Initiator and Co-Director of Wellcome Trust 4-year PhD Programme:
'Systems Neuroscience: From Networks to Behaviour'

2009-2012 Visiting Professor, Seoul National University, South Korea

2015 Professor of Neuroinformatics

2016 Editorial Board Member of Network Neuroscience (MIT Press)
and Royal Society Open Science;
Fellow of the Royal Society of Biology (FRSB)

2017 Leader of Neuroinformatics UK
Member of UK Computing Research Committee (UKCRC)

2018 Member of MRC neuroscience funding panel

Organisation of this course

Contact details: Prof. Marcus Kaiser (m.kaiser@ieee.org),
Course web http://www.dynamic-connectome.org/t/cneurosci/

Course components

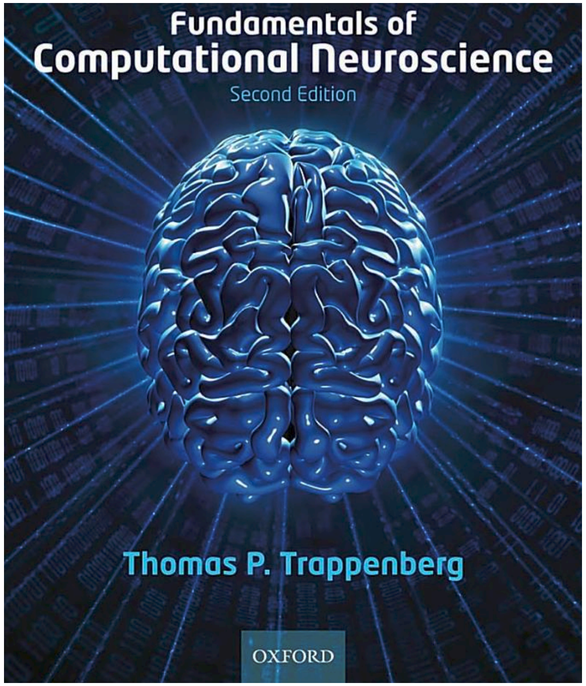
Practicals: Introduction to Matlab, Analyzing brain connectivity, Modelling Neural networks; Instructors: Dr Marcus Kaiser, Mrs Xue Chen, Ms Frances Hutchings, Mr Christopher Thornton.

Seminars (after the reading week): oral presentation of research articles (20 minutes plus 10 minutes discussion) worth 10% of the mark; list of articles will be provided within the next two weeks.

Small research project: worth 30% of the mark; list of projects will be available at the start of week 3 (choose before the end of week 4); submission of project: mid December.

Exam: January, worth 60% of the mark

Textbook



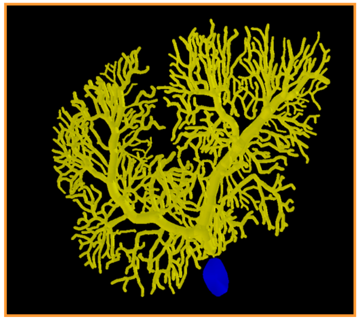
Thomas P. Trappenberg: Fundamentals of Computational Neuroscience, Oxford University Press, **Second Edition**, 2010, ISBN 0199568413.

Available in Robinson Library 4th floor, 612.82 TRA

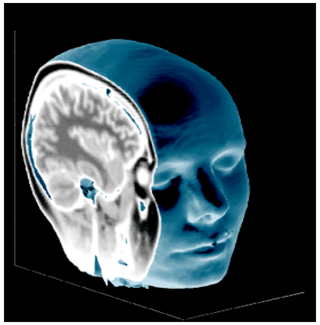
Course overview

1. *Introduction (chapter 1)*
2. Neurons and conductance-based models (chapter 2)
3. Simplified neuron and population models (chapter 3)
4. Information theory and neural coding (appendix D)
5. Associators and synaptic plasticity (chapter 4)
6. Brain connectivity analysis I
7. Brain connectivity analysis II
8. Feed-forward mapping networks (chapter 6)
9. Cortical maps (chapter 7)
10. The cognitive brain (chapter 10)

Neuroinformatics

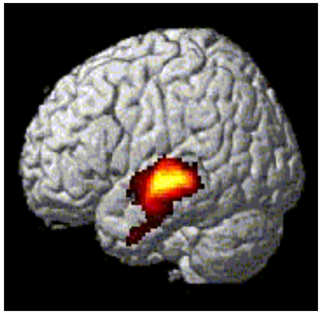


- The challenges



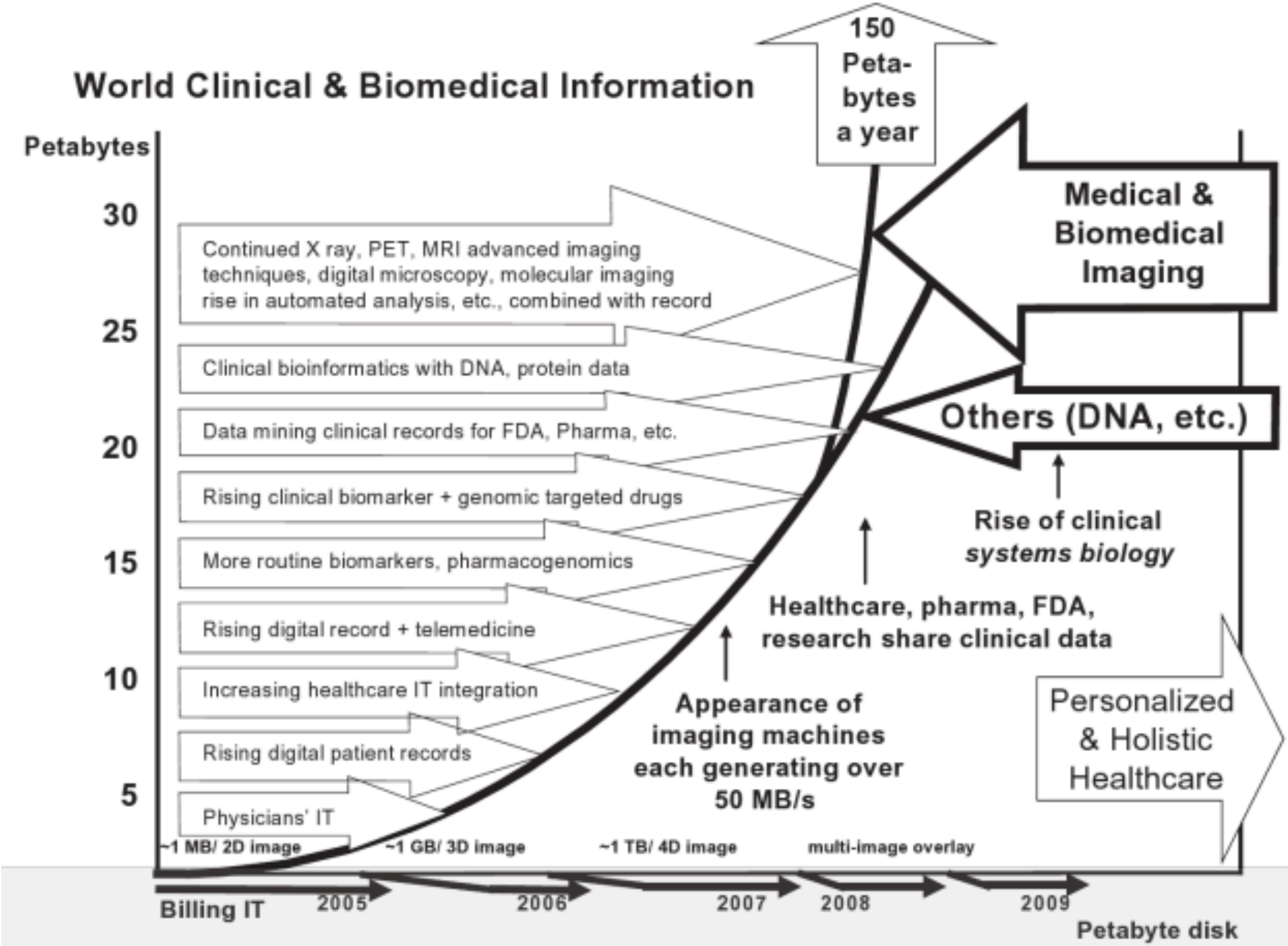
- The methods

- The solutions



- Neuroinformatics in Newcastle

A brief history



Barry Robson, O. K. Baek (2009) The engines of Hippocrates: from the dawn of medicine to medical and pharmaceutical informatics.

Information overload I

Example: Neuroimaging

2011 20,000 MRI scanners
-> 120,000,000 scans per year
(say 1 GB per scan)
-> 120 PB = 120 * 10^{15} Bytes = 120 million GB
(more than data storage at Google!)

Additional data from MEG, CT, EEG

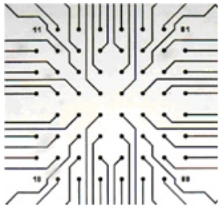
-> Data volume (SPACE): How to store and organize the data?

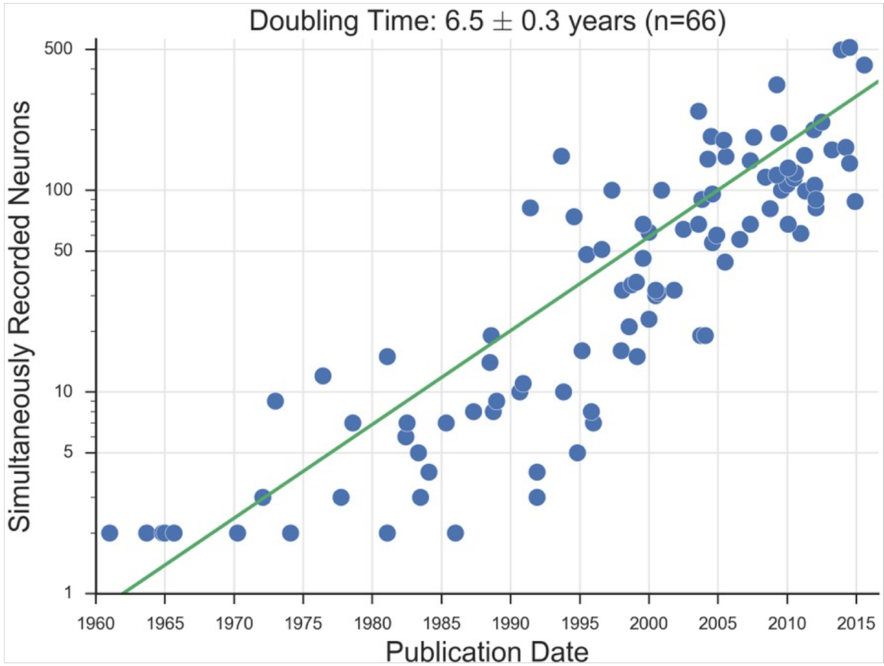
-> Long-term storage (DATA CURATION): neural diseases can last for a long time (e.g. epilepsy, schizophrenia, autism), patient data must be accessible for several decades!

Information overload II

<u>Example: electrophysiology</u>

1980 1 electrode
1990 10 electrodes
2000 100 electrodes
2010 10,000 electrodes





-> Data volume (SPACE): up to 100x as much as a decade ago
Problem: How do you share 5 TB of data with a colleague?

-> Calculations (TIME): up to 10,000 times as many as a decade ago (e.g. correlations between all pairs of *k* electrodes, $O(k^2)$)
Problem: Computer speed *only* increases by a factor of 32 in a decade (Moore's law)!

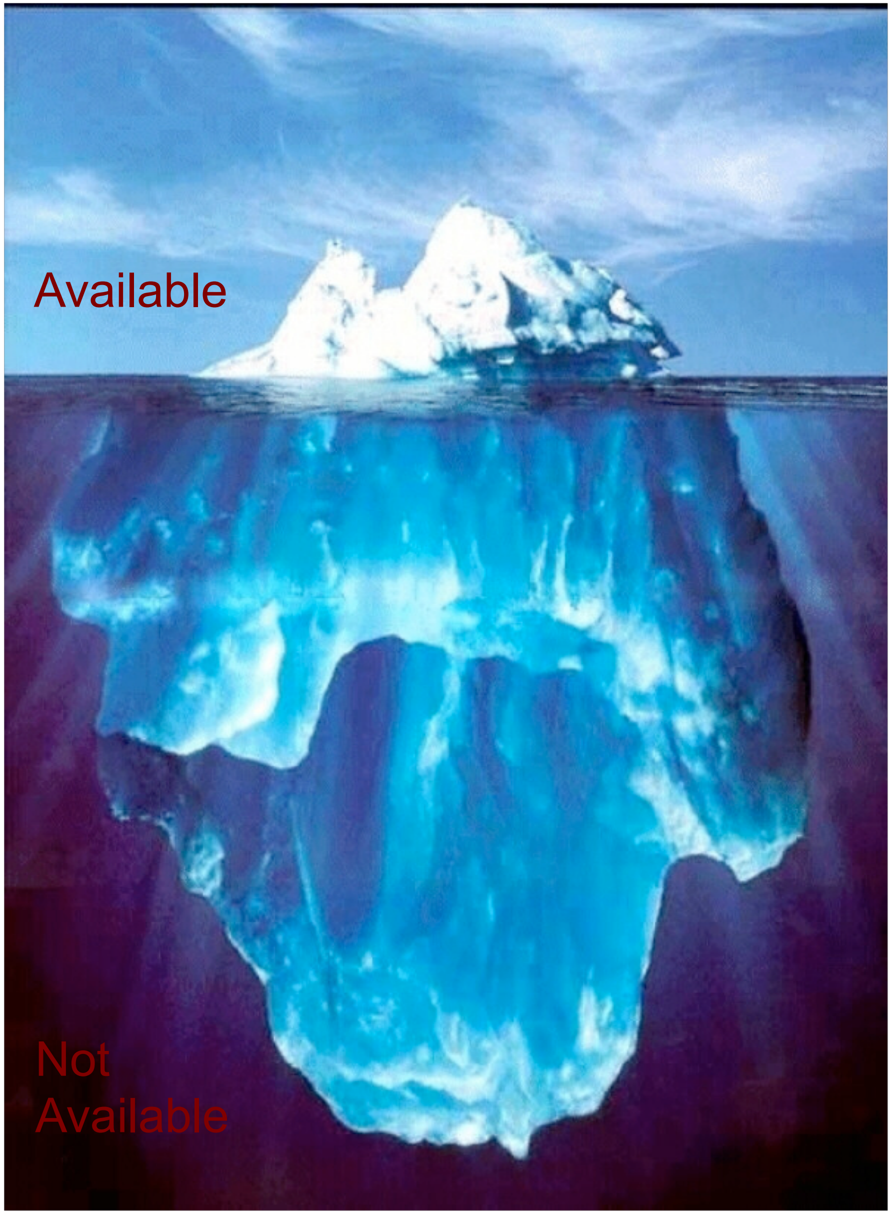
Data comparison

<u>Combining and comparing experimental results:</u>

More than 20 data formats for electrophysiology alone:
often each manufacturer of recording equipment has its own data format!
-> how to convert to a common format?

Experiments in different labs use different procedures: How much metadata must be provided to define an experiment (make it reproducible)?
-> annotate experimental data with metadata about the experiment (e.g. sampling rate, species, recording equipment).

The neuroscience data treasure



<u>Taxpayer value for money:</u>

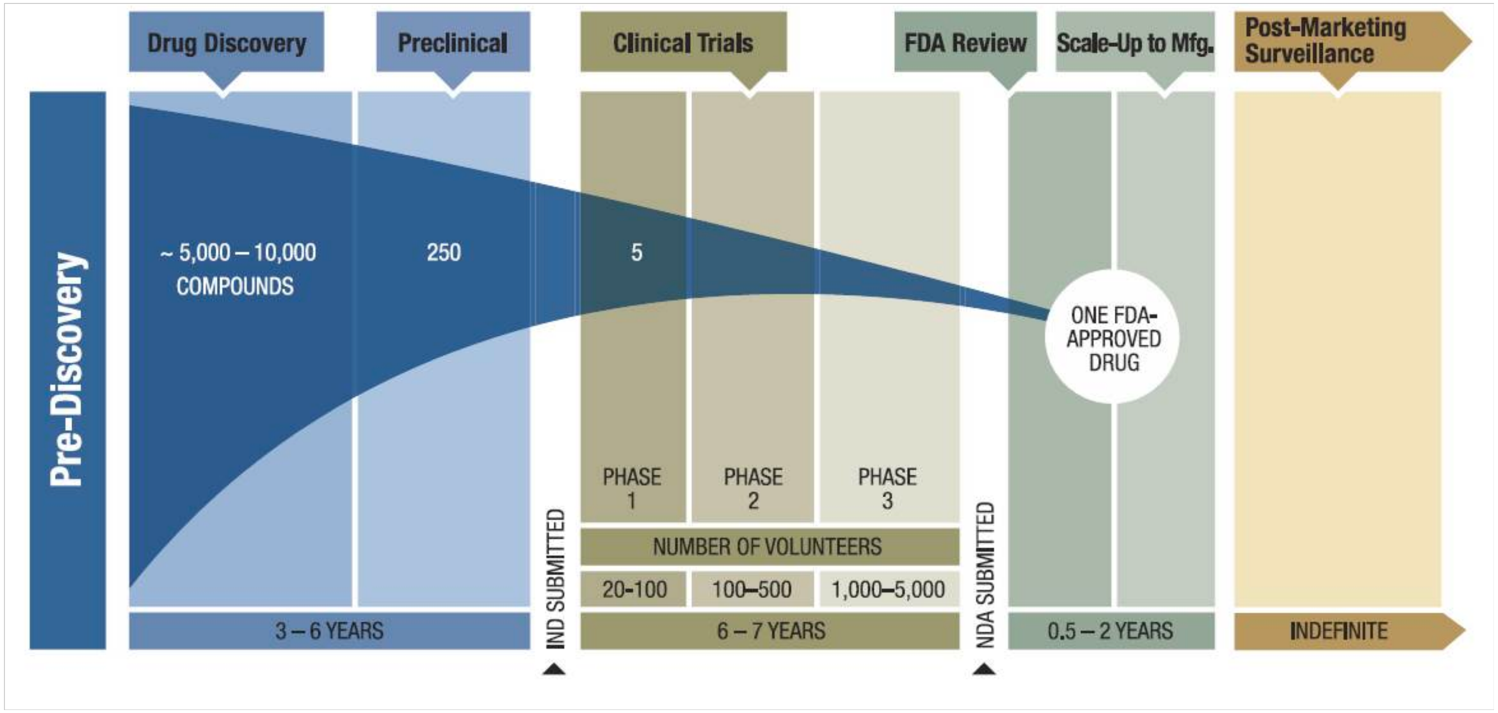
1) Getting data is expensive

2) Data from one lab might be useful for another lab (no need to re-do experiments)

3) Data should not be lost over the years (data/storage formats)

-> Databases are needed!

Bottleneck in drug discovery



(1) animal models are not human models
(more cell types in human brains)

(2) grown neural cell cultures differ from cortical tissue
(layer organisation, connections between cell types)

(3) tissue does not represent complex network interactions
(effect on activity between brain regions, individual differences)

Neuroinformatics around the world

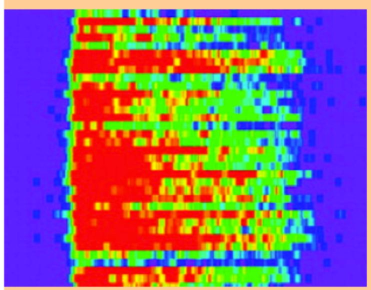
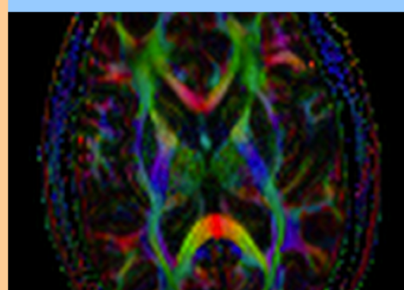
- In 2002 OECD Neuroinformatics Working Group identified the need to work cooperatively in order to achieve major advances -> formation of the International Neuroinformatics Coordinating Facility (www.INCF.org) is a consortium of 17 countries (Australia; Belgium; Czech Republic; Finland; France; Germany; India; Italy; Japan; Netherlands; Norway; Poland; South Korea; Sweden; Switzerland; United Kingdom; USA) established in 2005.
- Executive Office is based at the Karolinska Institute, Stockholm

Neuroinformatics UK

Home People Training Events Links Search

This is a network of researchers working in the emerging field of Neuroinformatics. Formed in 2004, it organises activities to strengthen and develop UK Neuroinformatics.

Sign up to one or more special interest groups related to experimental, clinical, or technological research. You can also join our mailing list or follow us on Twitter.

Experimental	**Clinical**	**Technological**
		
Neuroinformatics	Computational Neurology	Neuroinformatics and Neurotechnology
Learn more	Learn more	Learn more

www.neuroinformatics.org.uk

Follow us on @NeuroinfUK

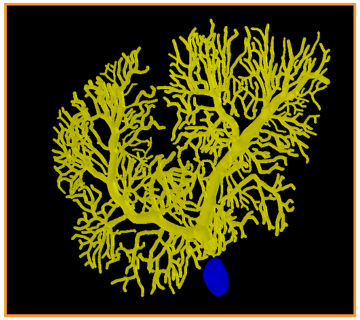
Neuroinformatics

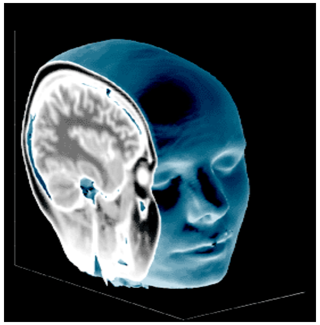
The challenges

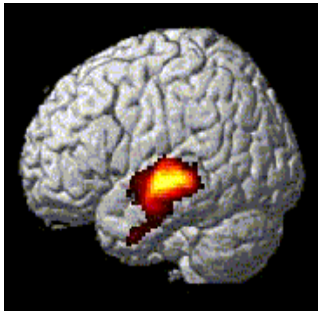
The methods

The solutions

Neuroinformatics in Newcastle







Metadata

- Metadata: means to describe data files providing information about a certain item's content (e.g., means of creation, purpose of the data, time and date of creation, creator or author of data)

- Often stored in XML (Extensible Markup Language) format. Do you know other markup languages?

- Minimum information: How much information is necessary to analyze a data set?

Nature Precedings : hdl:10101/npre.2008.1720.1 : Posted 25 Mar 2008

Minimum Information about a Neuroscience Investigation (MINI): Electrophysiology

Frank Gibson*[1], Paul G Overton[2], Tom V Smulders[3], Simon R Schultz[4], Stephen J Eglen[5], Colin D Ingram[6], Stefano Panzeri[7], Phil Bream[4], Evelyne Sernagor[6], Mark Cunningham[6], Christopher Adams[6], Christoph Echtermeyer[8], Jennifer Simonotto[1], Marcus Kaiser[1], Daniel C Swan[9], Martyn Fletcher[10], Phillip Lord[1]

Data curation

Digital curation involves organising and preserving digital information so that it may be available for future use. Effectively curated research data can be better shared among the wider research community, enhancing the long-term value of your work.

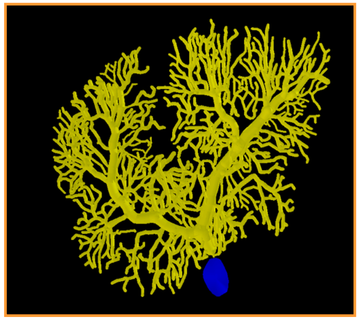
-> make sure people can still use your data in 10-50 years

Can you open a 1983 Word 1.0 file for MS-DOS?
How about a document file on a 5 ¼" floppy disk?

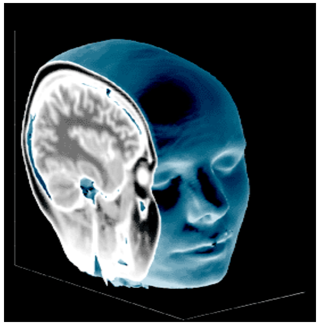
http://www.dcc.ac.uk/

Neuroinformatics

The challenges

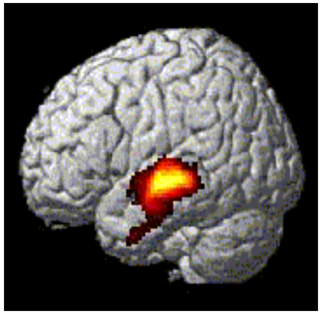


The methods



The solutions

The future



Neuroimaging Databases

Store data from several hospitals in one database.
Benefits:

- Consistency: Use the same scanning protocol for data comparison.
- Critical mass: easier to get enough patients for a clinical study (one hospital might not have enough patients).
- Time: the necessary number of patients can be reached earlier -> faster publication!

Human Connectome Project

HUMAN Connectome PROJECT

Mapping structural and functional connections in the human brain

Washington University in Saint Louis - University of Minnesota - Oxford University

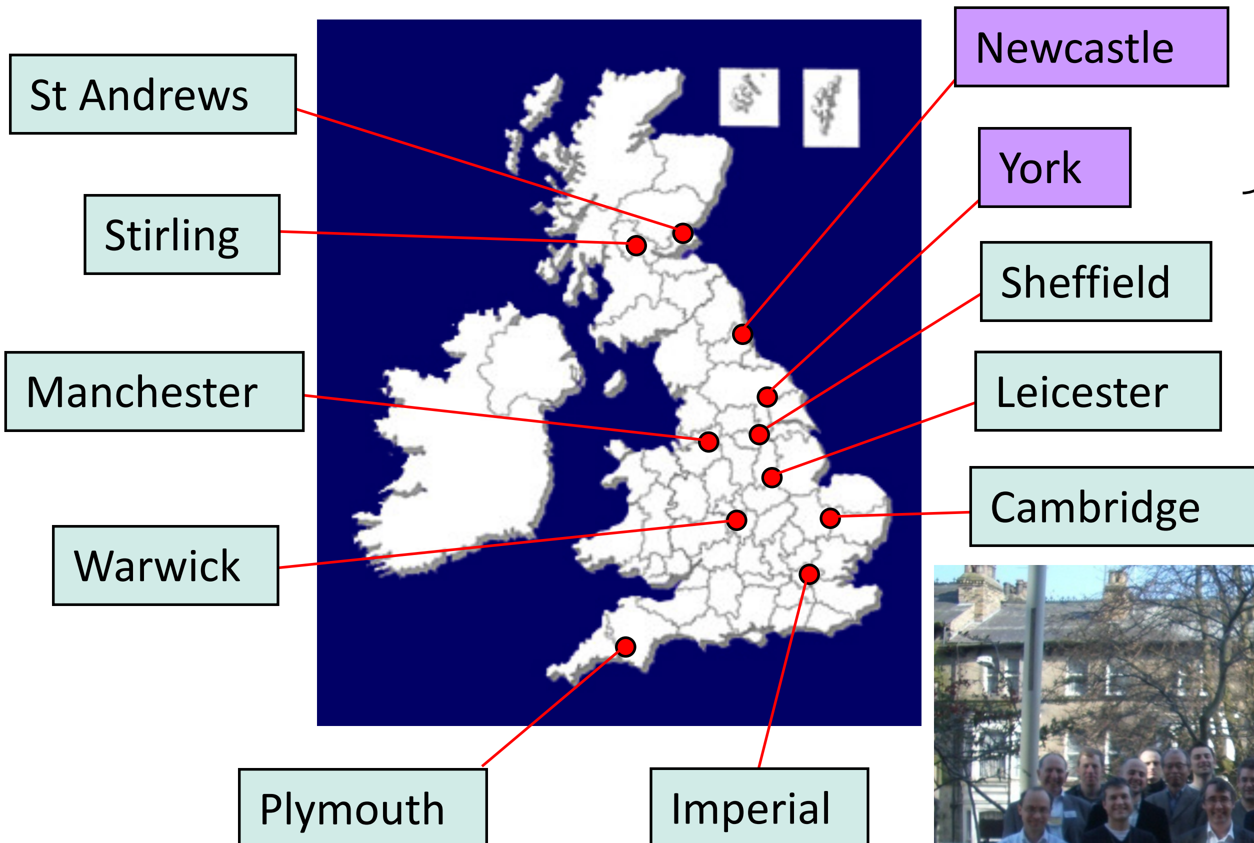
Saint Louis University - Indiana University - University d'Annunzio - Ernst Strungmann Institute

Warwick University - Radboud University Nijmegen - Duke University

- Running 2010-2015
- $30m NIH project

http://humanconnectome.org/

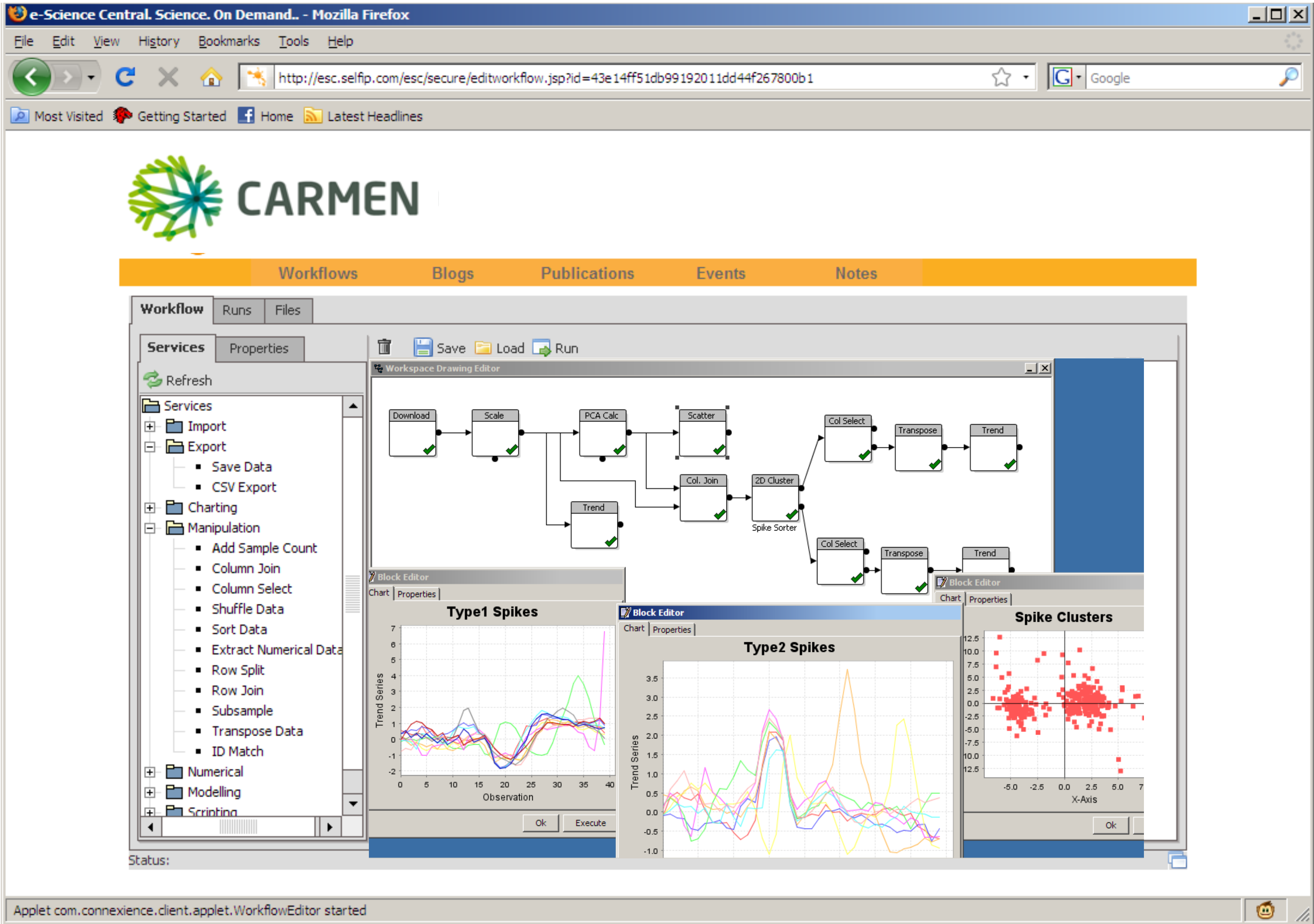
CARMEN electrophysiology DB



First two compute nodes

CARMEN

Workflow Software – e-Science Central



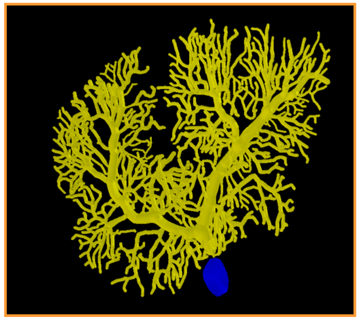
Neuroinformatics

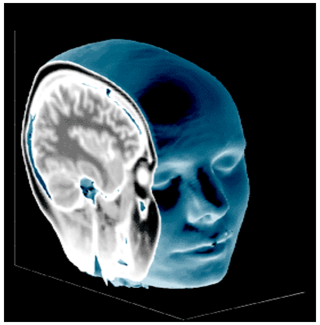
The challenges

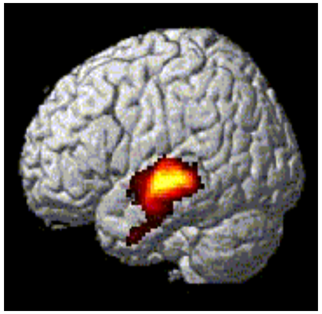
The methods

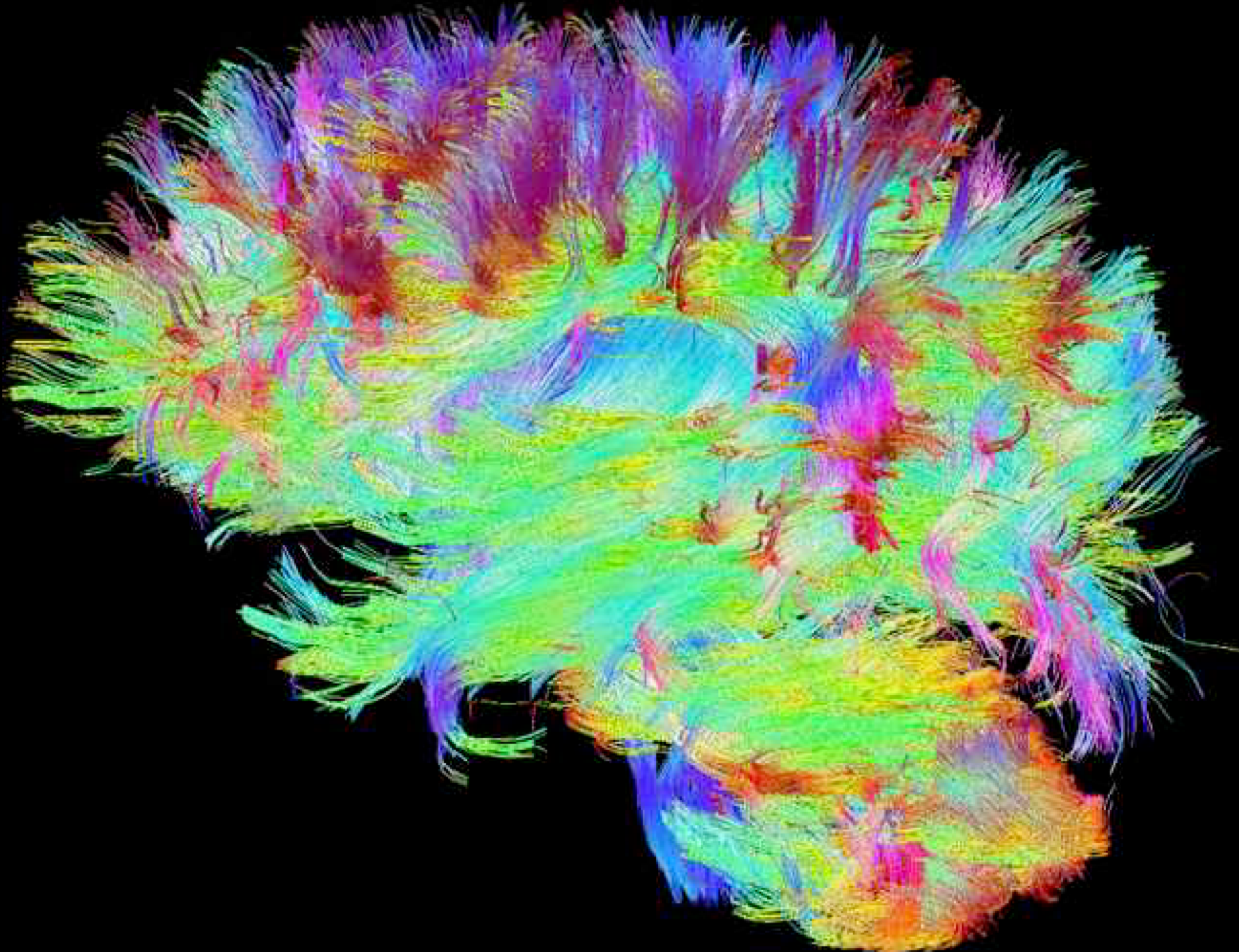
The solutions

Neuroinformatics in Newcastle

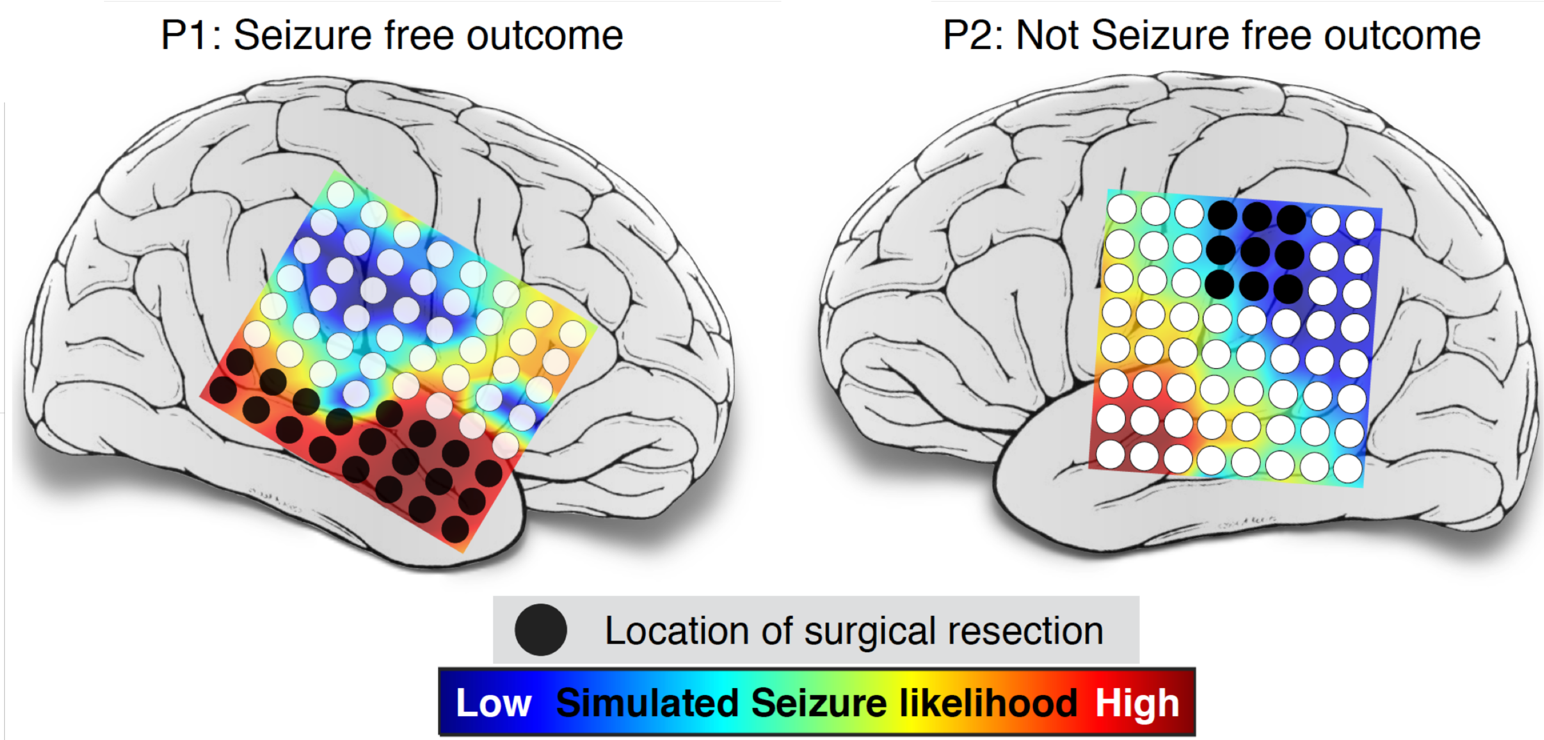


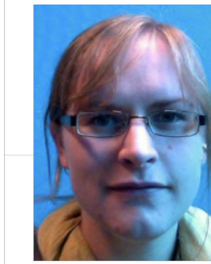


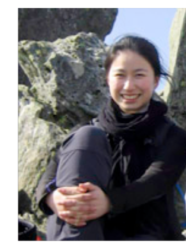




Computer simulations – predicting epilepsy surgery success







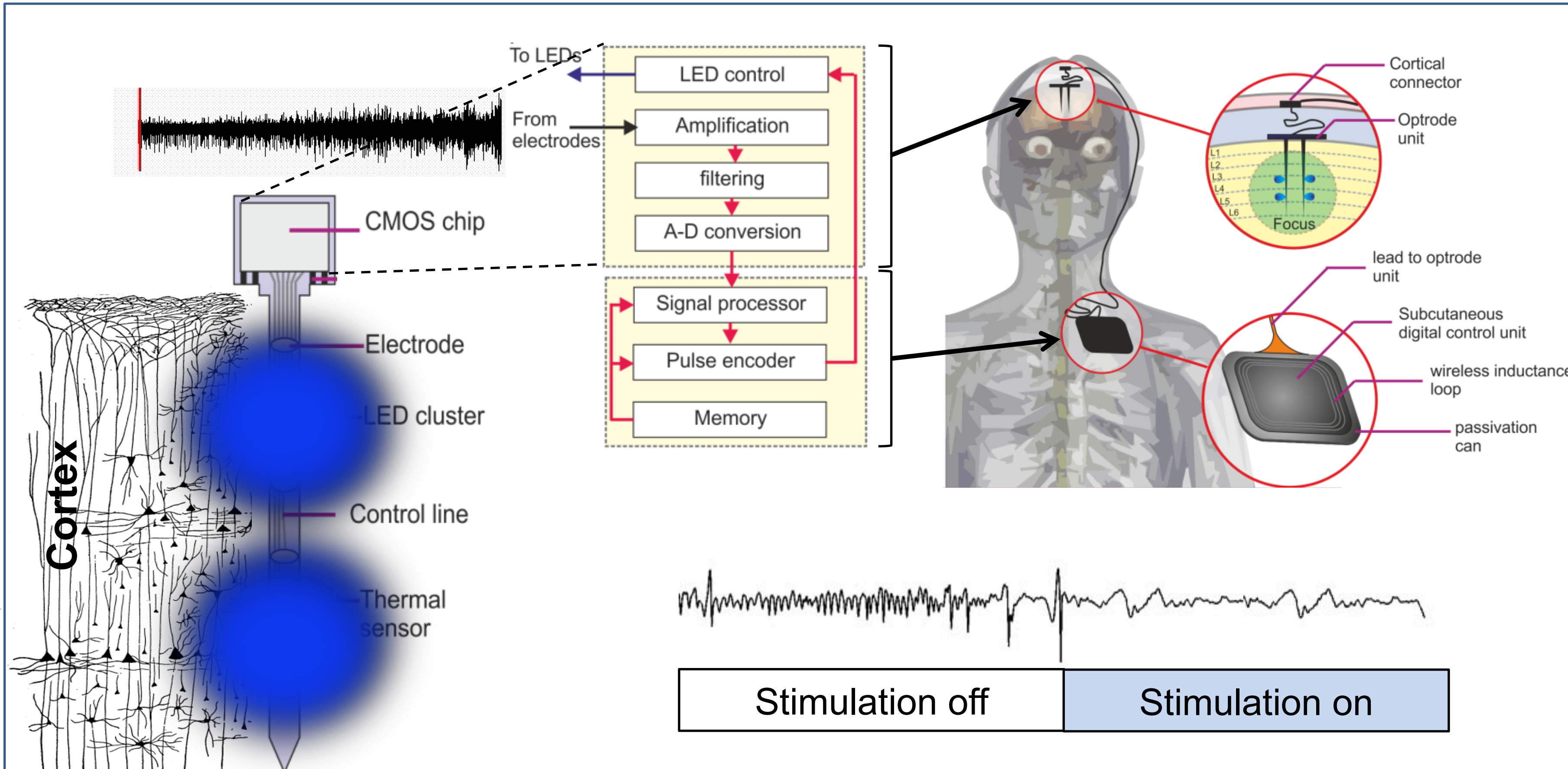
Prediction			Real outcome
seizure-free	→	100%	seizure-free
	↗	27%	
not seizure-free	→	*73%*	*not seizure-free*

Sinha et al. *Brain*, 2017
Hutchings et al. *PLOS Comput Biol*, 2015

Computer simulations – predicting the effect of optogenetic stimulation

Controlling Abnormal Network Dynamics with Optogenetics (CANDO)

7yrs (till 2021), £10m www.cando.ac.uk

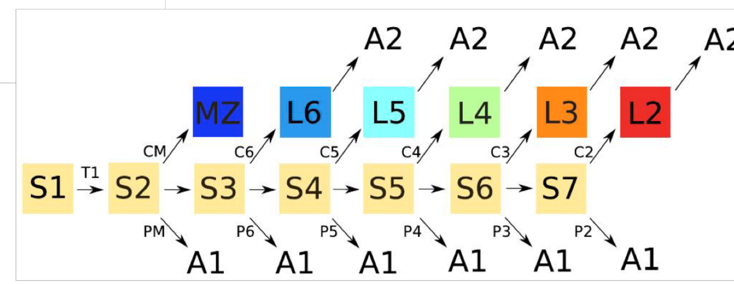


Models of tissue growth

Review

Mechanisms of Connectome Development

Marcus Kaiser[1,2,*,@]



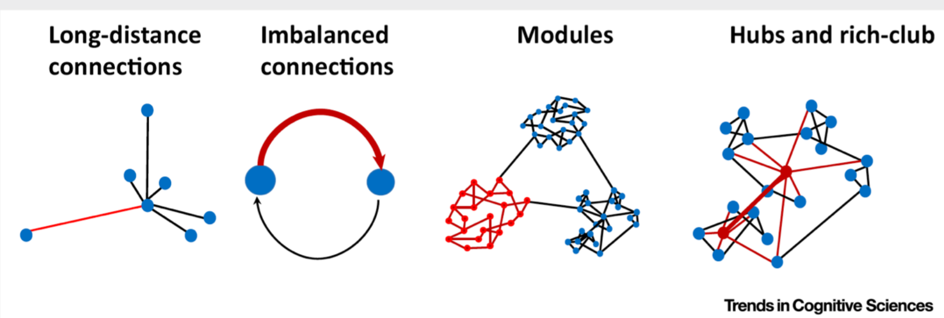


Figure I. Spatial and Topological Features of Brain Networks.

Kaiser (2017) *Trends in Cognitive Sciences*

https://biodynamo.web.cern.ch/

CERN openlab intel

www.dynamic-connectome.org

Neuroinformatics

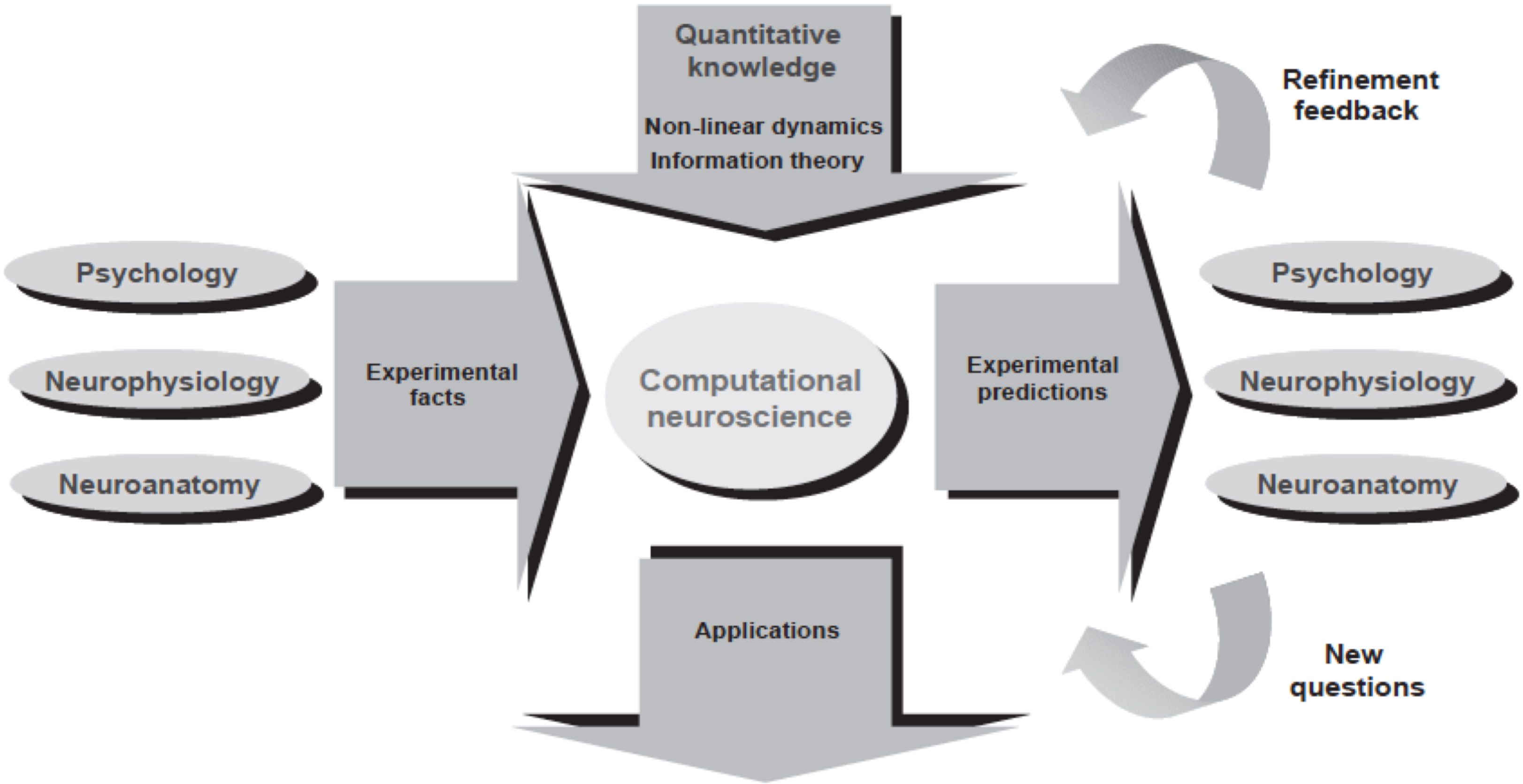
Marcus Kaiser

Week 1: Introduction (textbook chapter 1)

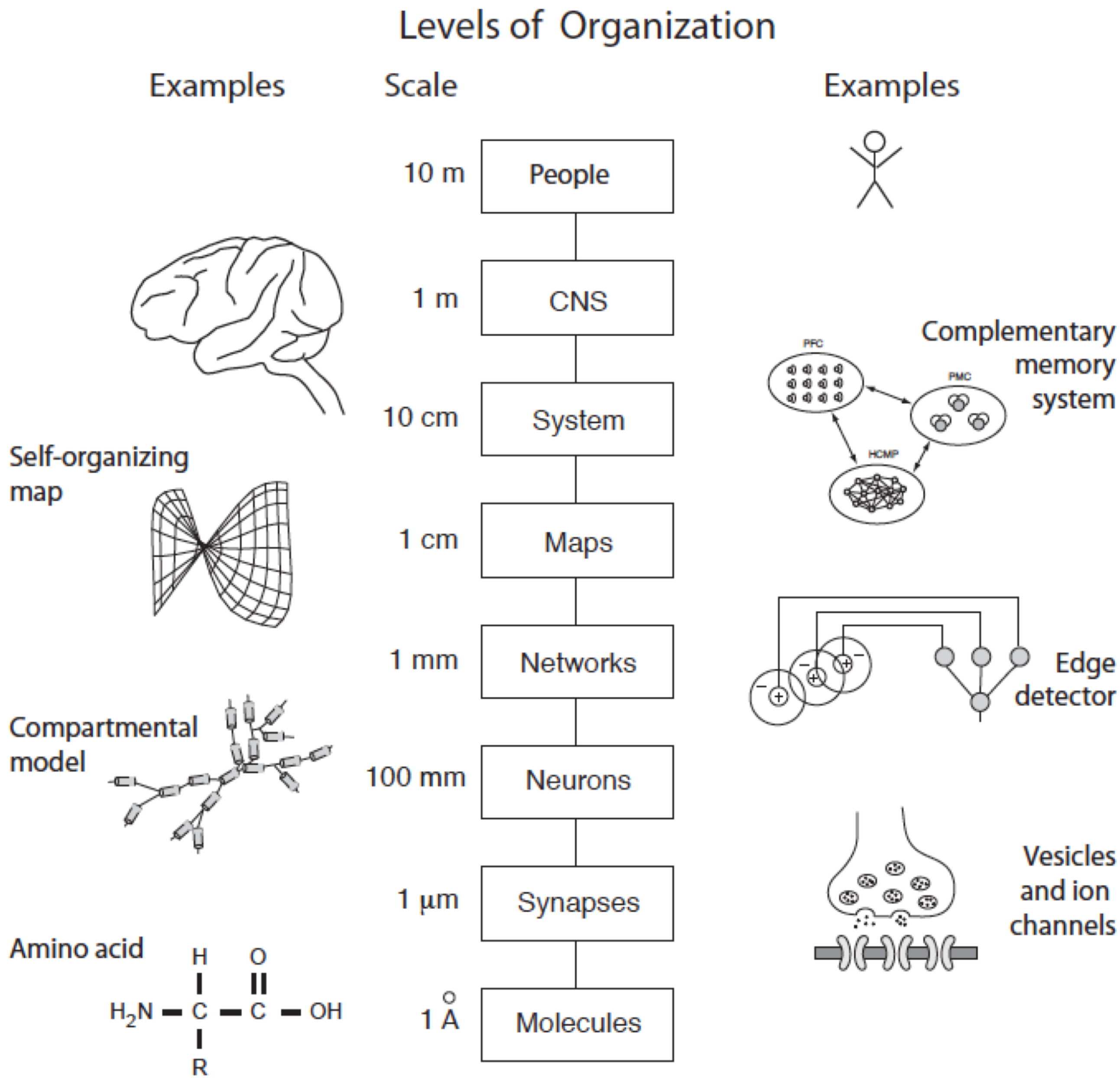
What is Computational Neuroscience?

Computational Neuroscience is the theoretical study of the brain to uncover the principles and mechanisms that guide the development, organization, information processing and mental abilities of the nervous system.

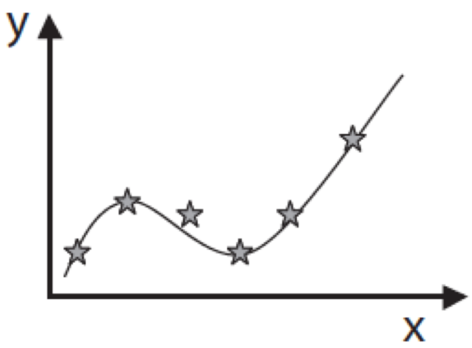
Computational/theoretical tools in context



Levels of organizations in the nervous system



What is a model?



Models are abstractions of real world systems or implementations of hypothesis to investigate particular questions about, or to demonstrate particular features of, a system or hypothesis.

Which model to choose?

1. **toy model**: abstract but easier to understand the role of different factors
2. **phenomenological model**: model with given input produces same output as the real system (black box approach)
3. **complex model**: model where many features of the real system are included (e.g. ion channels, neuron morphology)

Guideline: *Make a model as complex as is needed to explain a mechanism (but not more complex)! This is called 'principle of parsimony' or 'Occam's razor'.*

Is there a brain theory?

Availability of large-scale neuroscience data.

→Shift to **quantitative** hypotheses which can be tested experimentally

Marr's approach

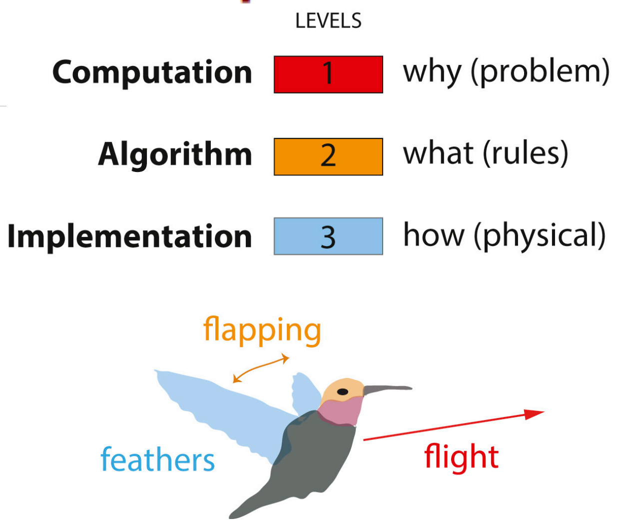
1. **Computational theory:** **What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?**

Marr's approach

1. **Computational theory:** What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?
2. **Representation and algorithm:** How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?

Marr's approach

1. **Computational theory:** What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?
2. **Representation and algorithm:** How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?
3. **Hardware implementation:** How can the representation and algorithm be realized physically?



Marr's approach

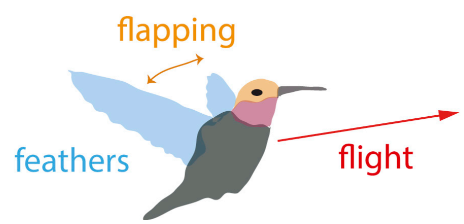
1. **Computational theory:** **What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?**
2. **Representation and algorithm:** **How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?**
3. **Hardware implementation:** **How can the representation and algorithm be realized physically?**

LEVELS

Computation	1	why (problem)
Algorithm	2	what (rules)
Implementation	3	how (physical)

Marr puts great importance on the first level:

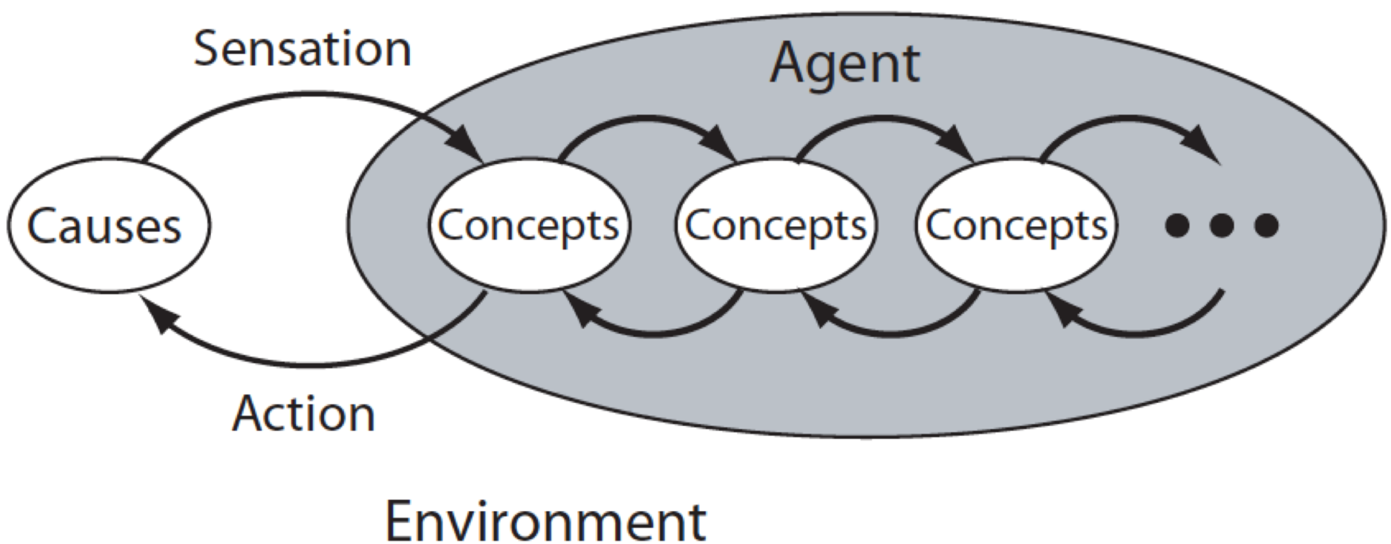
> *"To phrase the matter in another way, an algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanism (and hardware) in which it is embodied."*



A computational theory of the brain:

The anticipating brain

The brain is an anticipating memory system. It learns to represent the world, or more specifically, expectations of the world, which can be used to generate goal directed behavior.



Eye=camera? We only see **2%** of the visual field, the rest is "filled-in" (previous processing and prediction).

Why to use simulations of brain activity?

Theoretically feasible: many parameters for computer simulations can now be measured (e.g. brain connectivity or neuronal activity)

Computationally feasible: Increase in processing power

Economically desirable: *in silico* experiments are cheaper than wet lab experiments

-> **simulation-based science** (e.g. e-cell, Blue Brain)

Impact:
- The UK funder of biomedical research (BBSRC) plans to spent half of its money on "dry" research!
- Researchers like Henry Markram (1 bn Euro Human Brain Project) expect pharmacological research for brain disorders will be fully *in silico* in 10 years.
- Current calls for AI for biomedical applications and *in silico* clinical trials.

Why Neuroinformatics?

Neuroinformatics is a research field that encompasses the organization of neuroscience data and application of computational models and analytical tools. Neuroinformatics provides tools, creates databases and the possibilities for interoperability between and among databases, models, networks technologies and models for the clinical and research purposes in the neuroscience community and other fields.

1. **Data organization challenge:** organizing large-scale datasets (several TB) with thousands of files; how to transfer large files?
2. **Data integration challenge:** integrating data in different formats and from different sources
3. **Processing challenge:** finding faster algorithms for processing neuroscience data; use of parallel computing, grid computing, cloud computing, and high-performance computing (HPC)

Examples: CARMEN e-Science project for organizing and processing electrophysiological recordings (`http://www.carmen.org.uk/`); Human Connectome Project for collecting and analyzing human brain connectivity data

Further Readings in Computational Neuroscience

Patricia S. Churchland and Terrence J. Sejnowski, 1992, **The computational Brain**, MIT Press

Peter Dayan and Laurence F. Abbott 2001, **Theoretical Neuroscience**, MIT Press

Jeff Hawkins with Sandra Blakeslee 2004, **On Intelligence**, Henry Holt and Company

Journals: Frontiers in Computational Neuroscience, Journal of Computational Neuroscience, Network: Computation in Neural Systems, Neural Computation, Neural Networks, Neurocomputing, PLoS Computational Biology, Biological Cybernetics (see `http://www.neoxi.com/NNR/Neural_Network_Journals.php` for a larger list)

Further Readings in Neuroinformatics

Arbib & Grethe: Computing the Brain: A Guide to Neuroinformatics, Academic Press, 2001

Koslow & Subramaniam: Databasing the Brain: From Data to Knowledge, Wiley-Blackwell, 2005

Handbook of Brain Connectivity, Springer, 2009

Olaf Sporns: Networks of the Brain, MIT Press, 1st edition, November 2010 (ISBN: 978-0262014694)

Journals: Frontiers in Neuroinformatics, PLoS Computational Biology, Neuroinformatics

Questions

What are the different levels of brain organisation?
What is a model?

What are Marr's three levels of analysis?

What different types of models exist?

What are the benefits of using computer simulations?