CARMEN: Code Analysis, Repository, and Modelling for e-Neuroscience

Case for Support Part 1: Previous Track Record

CARMEN Consortium:

The 19 investigators involved in the CARMEN consortium bring together complementary expertise in e-Science (Austin, Jackson, Lord, Watson), experimental neuroscience (Baker, Ingram, Schultz, Sernagor, Smulders, Whittington), computational neuroscience (Baker, Borisyuk, Feng, Quian Quoraga, Schultz, L. Smith, Whittington) and modelling (Eglen, Gurney, Kaiser, Panzeri, Schultz, V.A. Smith,). Furthermore, the consortium includes considerable experience in delivering large scale e-Science and neuroscience projects that are managed over geographically distributed sites. The different laboratories involved are located in 10 Universities which have strong connections to the wider neuroscience and e-Science communities, and the departments involved have been rated highly for research in RAE2001, notably: Computer Science (UoA25): York 5*, Newcastle 5, Warwick 5; Neuroscience: Newcastle (UoA02/03) 5; Applied Mathematics and Theoretical Physics: Cambridge 5*; Engineering: Imperial 5*, Leicester 5; Biological Sciences: Manchester/UMIST 5*/5, St Andrews 5; Psychology: Newcastle 5*; Sheffield 5.

1. Project Management Team:

The project will be implemented by a Project Management Team, the four lead members of which are representative of the major elements of the application. Austin and Watson both bring expertise in e-Science systems and extensive experience of delivering data intensive e-Science projects. Complementary to the focus of the project, Austin brings specific expertise in Grid systems for management and analysis of very large flat file time series datasets (DAME and BROADEN), while Watson brings an enviable track record in developing distributed database and e-Science security systems (OGSA-DAI, GOLD and ^{my}Grid). Ingram (Newcastle) represents the experimental neuroscience user community and his position as Director of the Institute of Neuroscience allows him to ensure that the benefits of the project will be delivered both nationally and internationally. L Smith's expertise in data extraction and neural computation provides a bridge between the computer science and neuroscience elements of the project.

Jim Austin (Project Management and WP0 Coordinator - Professor of Neural Computation, Department of Computer Science, University of York) directs the Advanced Computer Architectures Group at York. He has a background in Neurobiology (BSc, Sussex) and Artificial Neural Networks (PhD, Brunel). His research has focussed on the theory, implementation and application of binary neural networks for over 19 years, and has published over 200 reports, books and papers. He was the PI on the DAME e-Science pilot project which completed in October 2005 and is now involved in the DTI inter-enterprise project BROADEN that will take this technology into Rolls-Royce. He is a founder of the White Rose Grid, a metropolitan Grid that supports computing between the Universities of Leeds, Sheffield and York. He is a director of Cybula Ltd., a spin out set up to take the results of his group's research to industry.
(http://www.cs.york.ac.uk/arch/NeuralNetworks/index.htm)

Colin Ingram (Project Management - Director, Institute of Neuroscience, University of Newcastle) is a neurobiologist with an interest in affective state, including electrophysiological measurements from brain areas associated with stress and mood. He obtained his training in Animal Physiology (BSc, Bath) and Neuroendocrinology (PhD, Cambridge), before moving to the University of Bristol where he held positions as a Royal Society University Research Fellow and Reader. His major work during this time was on neuropeptide actions within the central nervous system. Since 2003 he has been Director of the Institute of Neuroscience at the University of Newcastle and has launched a number of major initiatives, including an interdisciplinary programme on neurotechnology aimed at developing devices which have application to nervous systems, such as multielectrode arrays.
(http://www.ncl.ac.uk/ion/staff/profile/c.d.ingram)

Leslie Smith (Project Management and WP1 Coordinator - Professor of Computing Science, Department of Computing Science and Mathematics, University of Stirling) has been working at the interface between computing and neurobiology for the last seven years. With a background in Mathematics (BSc, Glasgow) and Computing (PhD, Glasgow), he started work on neural network techniques in the 1980s before moving towards more biological systems, while retaining an interest in engineering biologically-inspired systems, particularly in the auditory domain. He ran the EPSRC network "Silicon and Neurobiology" (1998 - 2001) which resulted in two EPSRC grants, one of which has been on interfacing neural cultures and silicon technology and has enabled research on methods of spike detection and sorting. He is Chair of the IEEE UKRI Computational Intelligence Chapter and has published over 75 peer reviewed papers.
(http://www.cs.stir.ac.uk/~lss/research.html)

Paul Watson (Project Management and WP0 Coordinator - Professor of Computer Science and Director of the North East Regional e-Science Centre, University of Newcastle) has devoted much of his recent work on data-intensive e-Science projects, including information repositories and distributed query processing. His background is Computer Engineering (BSc, PhD, Manchester) and as a Lecturer at Manchester University he was a designer of the Alvey Flagship and Esprit EDS systems. From 1990-5 he worked for ICL as a designer of the Goldrush MegaServer parallel database server, which was released as a product in 1994. In August 1995 he moved to the University of Newcastle where he is/has been an investigator on EPSRC, BBSRC, MRC, HEFCE and DTI-funded research projects worth approximately £11M.
(http://www.cs.ncl.ac.uk/people/home.php?name=paul.watson@ncl.ac.uk)

2. Project Work Package Coordinators and Co-Investigators:

Stuart Baker (WP5 Coordinator - Wellcome Senior Research Fellow, Institute of Neuroscience, University of Newcastle) conducts research on the role of synchronous neural activity in the control of movement. He is a graduate of Physiology (MA, Cambridge) and Neuroscience (PhD, Cambridge). His research involves making multiple identified single unit recordings in cortical and sub-cortical areas (including the spinal cord) in order to investigate the function of synchronous 'beta band'

oscillations in proprioception, and the production and reduction of tremor. He developed a novel experimental system which distributed experimental tasks during multiple-electrode recordings over a local area network and multiple users, and is active in developing novel analytical methods, driven by the complex experimental data which arises from his experiments. He also conducts studies in human volunteers using MEG, EEG and EMG recordings, with carefully instrumented behavioural tasks.
(http://www.ncl.ac.uk/ion/staff/profile/stuart.baker)

Roman Borisyuk (WP5 Co-Investigator - Professor of Computational Neuroscience, Centre for Theoretical and Computational Neuroscience, University of Plymouth) has interests related to the dynamics of neural activity and understanding theoretical principles of information processing in the brain. He was trained in Mathematics (MSc equivalent, Moscow State University, USSR) and Biophysics and Mathematical Biology (PhD equivalent, Institute of Biological Physics, USSR), and has received a DSc from the Institute of Experimental and Theoretical Biophysics (Russian Academy of Sciences). Until 1998 he was Head of the Neural Networks Laboratory at the Institute of Mathematical Problems in Biology (Russian Academy of Sciences, Moscow). He has published over 75 books and refereed articles on analysis and modelling of neural signals. He is a member of the Editorial Board of Neural Networks and the International Journal of Integrative Neuroscience. He has been a recipient of many research grants, including five EPSRC grants.
(http://www.tech.plym.ac.uk/soc/staff/roman/home2.htm)

Stephen Eglen (WP6 Co-Investigator – Lecturer, Department of Applied Mathematics and Theoretical Physics, University of Cambridge) has a long-standing interest to understand the patterns of spontaneous activity in the developing nervous system, and the roles that activity might play. He has a background in Cognitive Sciences (BSc, Nottingham) and Neuroscientific Modelling (DPhil, Sussex) and currently works in the Computational Biology group at Cambridge. He has developed several analysis and visualization programs for studying spontaneous neural activity (waves2, waves3, sjemea). These tools have been used in several studies to investigate properties of the networks generating spontaneous activity and how they might change over time.
(http://www.damtp.cam.ac.uk/user/sje30/)

Jianfeng Feng (WP5 Co-Investigator - Professor of Computer Science, University of Warwick) has research interests that include novel variants of spiking neuron and neuron network models, and analysis of their properties at both biophysical and abstract levels. He received his academic training (BSc, MSc, PhD) from the Department of Probability and Statistics, Peking University, China, and has held posts in Germany, Italy, Cambridge and Sussex where he was Reader in the Department of Informatics (200-2004). He has EU collaboration with biologists at Cambridge, and Edinburgh and physicists in Rome to build a biologically plausible neural networks. He has published over 100 refereed publications and edited a book entitled "Computational Neuroscience: A Comprehensive Approach" (2003, Chapman and Hall (CRC Press).
(www.dcs.warwick.ac.uk/people/academic/Jian.Feng/).

Kevin Gurney (WP3 Coordinator - Reader, Department of Psychology, University of Sheffield) has formal training in Mathematical Physics (BSc, Sussex), Digital Systems Engineering (MSc, Brunel) and Neural Networks (PhD, Brunel). His research experience includes the development of neural network training algorithms, modelling of the human visual system, and computational neuroscience at several levels of description. He has published in all these areas in international journals (including TINS, Neuroscience and Neural Networks) and holds grants worth a total of £2.5M. Dr Gurney works in the interdisciplinary Adaptive Behaviour Research Group at Sheffield and collaborates extensively with researchers in computer science and experimental neuroscience.
(http://www.shef.ac.uk/~abrg/people.shtml)

Tom Jackson (WP0 Co-Investigator - Project Manager, Department of Computing Science, University of York) was project manager for a 3-year e-Science project, DAME, which developed the SDE system to be applied to CARMEN. This involved coordinating the research work of over 16 RAs, four academic sites and two companies. He is now the project manager for the four-year DTI funded BROADEN project, which is taking the DAME technology into use by Rolls Royce. He has a BSc in Electronics (Salford) and a PhD in Neural Networks (York), and prior to moving to York he worked for 5 years at the European Joint Research Centre in Italy, developing EU projects within the centre. His PhD was based on work that has become the basis for many projects in the Advanced Computer Architectures Group at York, involving the use of neural associative memories for intelligent systems.
(http://www-users.cs.york.ac.uk/~tomj/welcome.htm)

Marcus Kaiser (WP6 Co-Investigator - Academic Fellow for Complex Neural Systems, Institute of Neuroscience, University of Newcastle) has interests in network analysis both of cortical and neuronal networks, and the organization, development, and robustness of biological networks. He obtained his training in Biology (MSc, Ruhr-University-Bochum) and Computational Neuroscience (PhD, International University Bremen). Being trained in experimental neuroscience as well as in modelling, he has worked on Bayesian spike train analysis and structure and function of cortical and neuronal networks. His research on network analysis was featured as a cover article in Trends in Cognitive Sciences (Sept. 2004).
(http://www.ncl.ac.uk/ion/staff/profile/m.kaiser)

Phillip Lord (WP0 Co-Investigator - Lecturer, School of Computing Science, University of Newcastle) has extensive experience in the use of metadata for describing biological data. In particular, he has focussed on use of ontological technologies, including newer languages such as OWL, developed by the Semantic Web community. This work has been combined with work on Grid and other distributed technologies. As part of both the ^{my}Grid, and Grimoires projects, he has developed technologies for service discovery and description. Previously to this, he worked on adapting similarity measures for use over the Gene Ontology. Dr Lord's metadata and OWL expertise is internationally recognised. He is co-chair of the Bio-Ontologies workshop (bio-ontologies.man.ac.uk) and has contributed expertise to both BioPAX and the Genome Standards Consortium. His training was in Genetics (BA, Cambridge; PhD, Edinburgh).
(http://www.cs.ncl.ac.uk/people/home.php?name=Phillip.Lord@ncl.ac.uk)

Stefano Panzeri (WP2 Co-Investigator – Senior Lecturer, Department of Life Sciences, University of Manchester) conducts research at the interface between theoretical and experimental neurobiology aimed at understanding the principles of information processing in the cerebral cortex. He has a background in Physics (BSc, Turin) and Neuroinformatics (PhD, Trieste), and has acquired interdisciplinary research experience in both theoretical physics and computational neuroscience by means of various fellowships (including an MRC Research Fellowship in Neuroinformatics). He developed new and widely cited quantitative techniques, based on Information Theory, to analyze the time course of the electrical activity of populations of neurons. These techniques determine how, and how well, a "downstream" neuronal receiver decodes the messages contained in the activity of cortical neurons. This approach has provided major and novel insights into brain function, by demonstrating the cardinal importance of the millisecond-precise temporal structure of neuronal activity. He has published over 50 articles which have been cited more than 700 times to date. He is a member of the MRC Bioinformatics Panel.
(http://www.ls.manchester.ac.uk/people/profile/index.asp?tb=1)

Rodrigo Quian Quiroga (WP1 Co-Investigator - Lecturer in Bioengineering, University of Leicester) conducts research mainly focused on computational neuroscience and the analysis of electrophysiological data. His background is in Physics (MS Buenos Aires, Argentina) and Statistical Analysis of Neural Signals (PhD, Lübeck, Germany). He studies neural correlates of visual perception by analyzing single-cell recordings in epileptic patients. He is interested in the decoding of movement plans from single-cell recordings and, in collaboration with several clinical research centres, has studied brain evoked responses and their correlations with learning processes. His research involves the use and development of advanced methods of signal processing, such as wavelets, chaos theory and non-linear synchronization. He is a visiting associate at CalTech and a visiting researcher at the Department of Neurosurgery at the University of California Los Angeles.
(http://www.vis.caltech.edu/~rodri/home.html)

Simon Schultz (WP2 Coordinator - Lecturer in Bioengineering, Imperial College, London) has research interests that include modelling of neural circuitry, algorithms for analysis of neurophysiological data, and *in vivo* two-photon imaging of neural activity. He has a background in Physics and Electrical Engineering (BSc, Monash University, Melbourne) and Computational Neuroscience (DPhil, Oxford). His doctoral work involved modelling information transmission in the mammalian hippocampus, and developing a novel information-theoretic technique for analysing electrophysiological data. He has held fellowships at the Australian National University, New York University and University College London. In collaboration with researchers at UCL, he has been amongst the first to use new multi-photon imaging techniques to successfully record calcium transients in populations of individual dendrites. He is the author of over 20 peer-reviewed scientific publications.
(http://www.bg.ic.ac.uk/staff/schultz/)

Evelyne Sernagor (WP6 Coordinator - Senior Lecturer, Institute of Neuroscience, University of Newcastle) is interested in understanding how neurons wire together during development to form complex networks in the adult central nervous system. She received her BSc, MSc and PhD in Physiology and Neurobiology from the Hebrew University, Jerusalem and subsequently held posts at NIH, Bethesda and the Smith-Kettlewell Eye Research Institute, San Francisco. In recent years, she has concentrated her attention on the developing retina, studying the cellular mechanisms underlying oscillatory propagating activity patterns. She employs optical recordings (calcium and chloride imaging), which enable her to visualise and record synchronous activity from large neuronal populations. Through collaboration with Dr Knopfel (RIKEN Brain Science Institute, Japan) is studying neural networks using genetically modified mice with a variety of fluorescent probes.
(http://www.ncl.ac.uk/ion/staff/profile/evelyne.sernagor)

V. Anne Smith (WP6 Co-Investigator - Academic Fellow, School of Biology, University of St. Andrews) has expertise in both biology and computation. She has a BSc in Biology with Mathematics (College of William and Mary, Williamsburg, USA) and both her postgraduate (PhD, Indiana University, USA) and postdoctoral work combined biology with computation. Her current research is on the computational analysis of complex biological networks, building on research conducted in the Department of Neurobiology at Duke University (Durham, USA). There she applied a Bayesian network algorithm to infer information flow in the brains of songbirds processing auditory stimuli. Inferred networks were validated by matching known anatomy and biological properties of the auditory system. The networks showed significant differences between stimuli presented, showing the potential of this technique to analyse the relationship between behaviour and network flow.
(http://biology.st-andrews.ac.uk/vannesmithlab/)

Tom Smulders (WP6 Co-Investigator - Lecturer, Institute of Neuroscience, University of Newcastle) conducts research focused on the neural basis of spatial and episodic memory. He has a background in Biology (BSc equivalent, Antwerp), Zoology (MSc equivalent) and Biopsychology (PhD, Cornell University, USA). Dr Smulders held a post-doctoral positions at Wake Forest University (Winston-Salem, USA) where he recorded from ensembles of single neurons in the hippocampus of awake and behaving rats, and at Duke University (Durham, USA), where he continued recording with multi-electrode arrays in the auditory system of the zebra finch. Currently, Dr Smulders is developing methodology to make multi-electrode recordings from the hippocampus of birds.
(http://www.ncl.ac.uk/ion/staff/profile/tom.smulders)

Miles Whittington (WP6 Co-Investigator - Professor of Neuroscience, Institute of Neuroscience, Newcastle University) researches mechanisms underlying neuronal network behaviour. His background is in Pharmacology (BSc, PhD, Bristol) and post-doctoral work carried out at St. Mary's Hospital Medical School and Imperial College London focused on the generation and propagation of epileptiform activity in brain slices *in vitro*. Later work, at Imperial College London and Leeds University, focused on aspects of physiological network rhythms of cognitive relevance. He has been supported by project and equipment grants from the Wellcome Trust, programme and LINK grants from the MRC, a collaborative grant with the Volkswagenstifftung (Germany), and a programme grant with NIH (USA). He has published over 80 papers on these topics with seminal work including the design and characterization of *in vitro* models of EEG rhythms which allow greater understanding of network mechanisms underlying neuronal population behaviour associated with sensory processing.
(http://www.ncl.ac.uk/ion/staff/profile/m.a.whittington)

Case for Support Part 2

1. Project Background and Overview

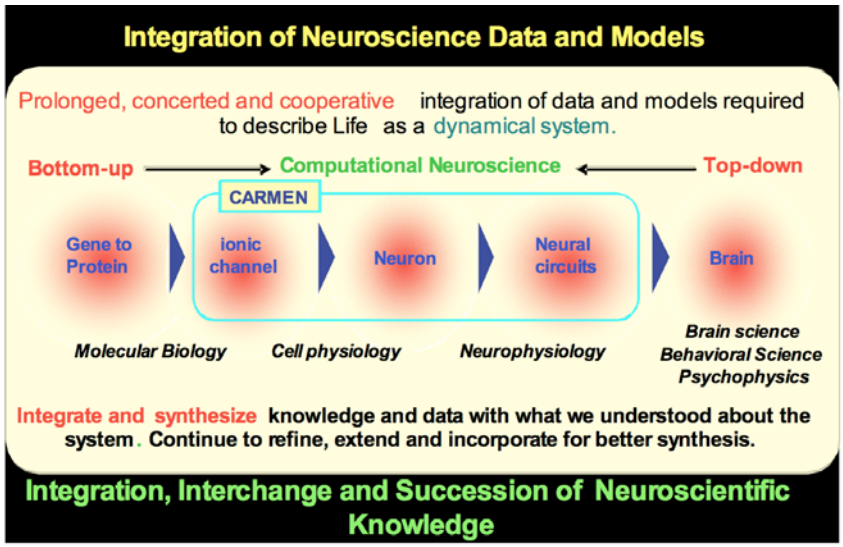
Understanding the brain is one of the major scientific challenges. The consequent challenge for bioinformatics is the capability to synthesize a detailed and applicable understanding of the way in which information is encoded, accessed, analysed, archived and decoded in neuronal networks. This is not only fundamental to the objectives of computational and experimental neuroscience, but has major application to computer science (e.g. *neuromorphic* and *neuromimetic* systems), nanotechnology (neuronal interfaces, neuroprostheses, and biosensors), electronic engineering and informatics (data handling and design strategies for multi-sensor systems), and pharmacology (*in silico* drug development).

Information flow in neuronal networks is made up of a series of component processes which can be investigated using a range of electrophysiological and imaging techniques. Traditionally, these techniques have addressed discrete levels of scale, e.g. patch-clamping to observe intracellular conductances and potentials, electrode and multi-electrode array (MEA) methods to record signals and identify signal patterns, optical and dye-based recordings to capture network dynamics, and behavioural observations. These data are difficult and expensive to produce, but are rarely shared and collaboratively exploited. A proliferation of techniques produces data that are voluminous, proprietary, locally described and curated, and therefore difficult to integrate. It is also uncommon for the recipient data analyst to apply analysis algorithms to more than one neuronal system.

As a result, there is: (a) a shortfall in analysis techniques that can be applied robustly and effectively across neuronal systems; (b) an absence of readily accessible data to support development of analysis techniques; (c) no structured curation and optimisation of data; (d) only sporadic interaction between disparate research centres with complementary expertise (e.g. data collection vs. data analysis), and; (e) very limited understanding of the informatics solutions that could be applied to broaden the science by integrating data across spatiotemporal scales. The challenge is thus to instigate a step change in the research methodology. This is internationally recognised (see the report of the OECD working group that led to the creation of International Neuroinformatics Coordinating Facility - INCF [1]). Recent developments in e-Science make it timely to address this challenge. CARMEN will leverage the UK e-Science infrastructure and capability to deliver a paradigm shift, supporting virtual integration of research teams and ambitious multi-modal experimentation. This will be achieved through provision of e-Science infrastructure to enact cross-modal data sharing and integration, supported by metadata and an expandable range of services accessible to users for raw, transformed and live experimental data. These innovations will create a **virtual neuroscience laboratory** that ties together experimental and computational neuroscientists.

1.1 International Context and Drivers

Neuroinformatics proposes that all data-resources (i.e. data, models and services) can be integrated, to 'database the brain' [24]. However, neuroscience is a vast, partitioned, *and* encompassing discipline, transcending the boundaries between physical, biological, medical and social sciences. Consequently, the INCF recognises the need for organic integration of scales and interests, as illustrated in **Fig. 1** [Adapted from Shiro Usui (RIKEN BSI & INCF)].



To position the UK at the forefront of international neuroinformatics it is essential to prove the benefit of organised data consolidation and integration across a neuroscience application that is tractable. CARMEN capitalises upon the UK strengths in experimental neurophysiology, computational neuroscience, and e-Science systems to deliver this. While this approach will demonstrate the validity of this data-resource integration, in the long-term it will be desirable to accord with the global neuroinformatics vision and strive for integrated implementation across all of neuroscience. Consequently, the e-Science challenge is to engineer an extensible Grid-based system that is: (a) generic to data-intensive neuroscience; and (b) supportive of very dynamic, flexible, integrated and secure deployment of domain semantics at the metadata, data and service levels.

1.2 Research Benefits and Deliverables

The focus of CARMEN will be neuronal activity, which encompasses: ionic conductances, synaptic and spiking activity and network dynamics. Therefore CARMEN will engineer, utilise and refine a scalable e-Science architecture to serve this research (**Fig. 1**). The CARMEN consortium is pioneering a rationalisation of their research practice, to embed organised data and software service resource sharing. Critical benefits that will be derived within the project (and beyond) include: (a) close collaboration between data collectors and analysts *during experimentation*; (b) cross-fertilisation and convergence of expertise between areas of specialisation; (c) diligent and long-term curation of data and tools, and; (d) optimal reuse and integration of data. The expedience of this virtual laboratory is illustrated in **Fig. 2** which contrasts the current neuroscience research process with the accelerated, integrating process that will be instigated by CARMEN. One benefit of the virtual laboratory is to allow the data collection and analysis components of the research to take place in parallel, saving time and money by eliminating the need to collect a significant volume of data before the processing and analysis begins. Cumulatively, this will make available curated data that will allow the focus of the research to shift from *in vivo* and *in vitro* methods, to *in silico* experimentation:

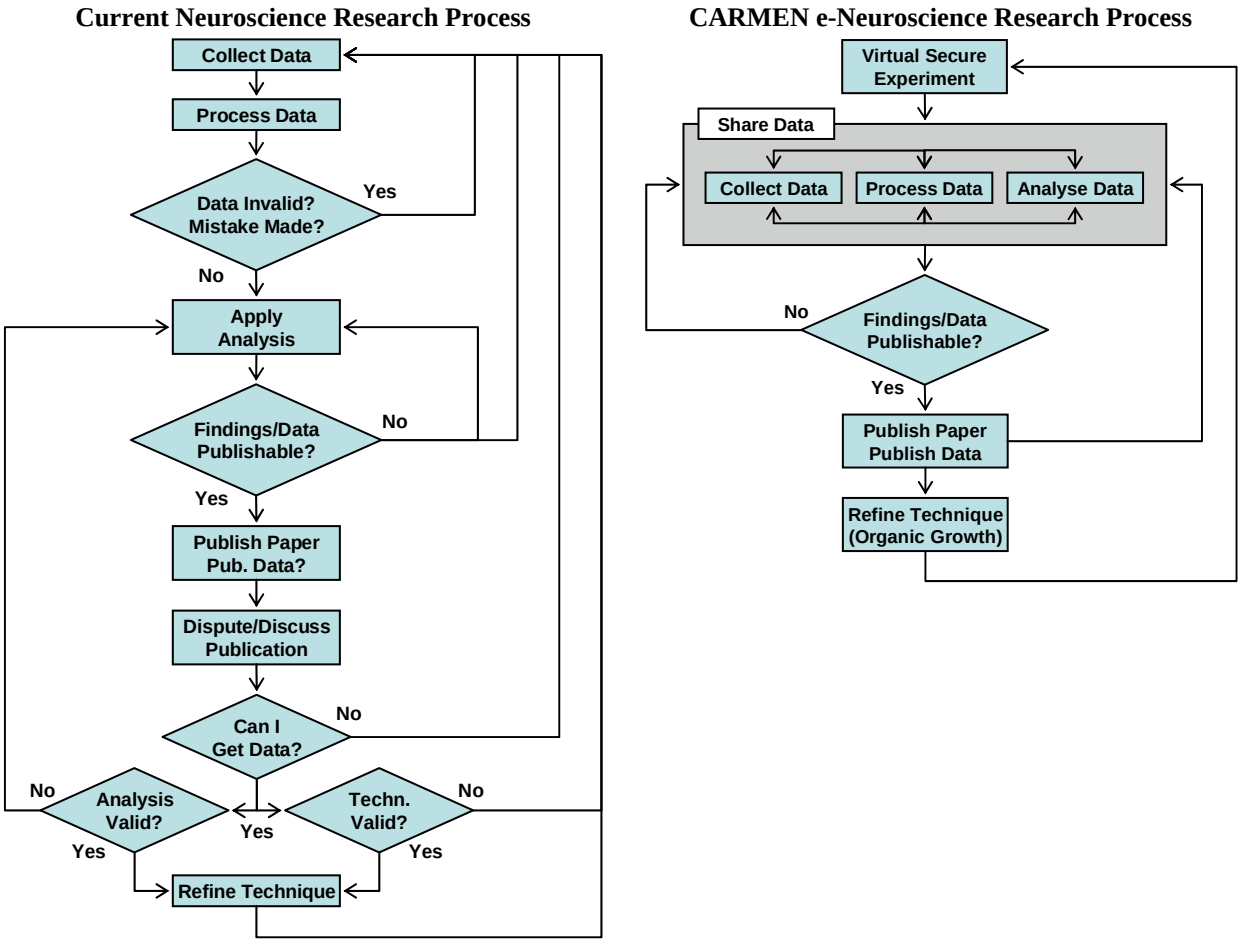


Figure 2: Current Neuroscience Research Process vs CARMEN Neuroscience Research Process

CARMEN will deliver:

a) A sustainable, extensible and distributed computing infrastructure, with integrated metadata and analysis services.

b) A series of pilot research applications, demonstrating cutting-edge, integrative approaches to neurophysiology on the Grid.

c) An understanding of the requirements and implications of (b), and demonstrable progress towards embedding, training and pump-priming a Grid-enabled user community.

d) The establishment and support of an organised and motivated user network, to ensure sustainability.

e) Exploratory collaboration with members of the INCF to pursue integration with other large-scale neuroscience databases, thus allowing UK researchers to be at the forefront of the international effort.

1.3 Summary of Technical, Research and Social Challenges

The distributed and data-intensive nature of the research application (multi-electrode electrophysiology data rates can easily exceed 20GB/h) requires a distributed data repository. Further, almost all of the time-series data have proprietary formats, demanding the application and extension of binary XML technologies to integrate and curate data resources. It will be necessary to utilise and refine spatiotemporal query tools to discover, analyse and annotate data. The partitioned nature of the user community dictates that diligent evaluation and extension of existing neuroscience schemas (BrainML [22], NeuroML [23] etc) will be necessary to formulate aggregated metadata, which can be interrogated. This, combined with the skills bias of the experimental community and sporadic curation of existing data, demands development of a highly assistive web portal (GUI) to support *get* and *put*, and deployment of (new and provided) services and workflows. It will also be desirable to integrate and share data in near real-time, at the experimental interface, so that data collectors can also benefit from the proposed step change by forming multidisciplinary, multi-site analysis teams to drive their experimental recording. Fine-grained security will be required to mitigate the early stage concerns of academic and commercial data providers and users. Further, all users will need to be trained. In spite of widespread realisation within the experimental and theoretical communities that more effective collaboration and value for money will be gained from the rationalised process, effort will be required to ensure community engagement; it is universally accepted that the application research challenge is difficult, and very long-term. Consequently, extensive business and community development will be needed to ensure that the service can be sustained (see section 4.3 and 4.5). This longevity is intrinsic to the scientific vision. Sections 2, 3 and 4 explore this range of challenges in detail, describing the innovations in e-Science and neuroscience that will be deployed and integrated, to realise the vision of CARMEN.

2. The CARMEN e-Science Infrastructure (WP0 – Grid Team)

Work Package Coordinators: Professor Paul Watson (Newcastle; 660h), Professor Jim Austin (York; 462h); **Co-Investigator:** Dr Phillip Lord (Newcastle; 655h); Professor Colin Ingram (Newcastle; 660h); Professor Leslie Smith (Stirling; 660h); **Research Staff:** RAi (months 1-48; Newcastle), RAj (months 1-36; Newcastle), SEa (months 6-42; Newcastle), RAk (Dr Mark Jessop; months 1-36; York); RAl (Dr Martyn Fletcher; months 1-36; York); **Other Resources:** Equipment/Consumables: £367,000; Travel: £34,850; Named Systems Administration (Chris Mountford; York; 4yrs x 25%FTE; £31,678); Unnamed Systems Administration (4yrs x 25%FTE; £30,481).

To fulfil the requirements identified in the previous section, an e-Science infrastructure will be built. This system will be comprised of the CARMEN Active Information Repository Nodes (CAIRNs), whose role it is to store and process the data. In order to understand the issues that will arise from the requirement for multiple CAIRNs to be integrated, two will be built in the first instance, one at York, and the other at Newcastle:

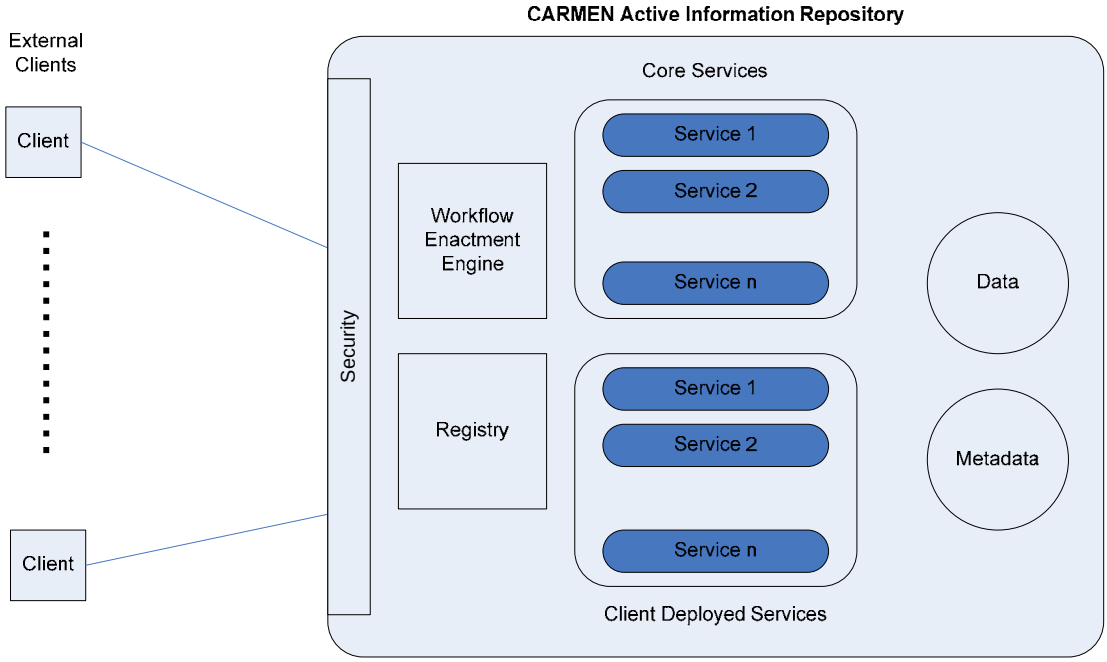


Figure 3: A CARMEN Active Information Repository Node (CAIRN)

Key components and technical features of a CAIRN are described below:

2.1 Storage and Management of Primary and Derived Data.

The repository will hold both raw voltage signal data (e.g. collected by patch-clamp and MEA recording) and image data (e.g. resulting from the internal chemical processes of the neuron and the activity dynamics of large neuronal networks). Services will be provided to import and export data in a variety of proprietary (raw) data formats (current neurophysiology instrumentation does not produce a uniform data format and is, in many cases, manufactured by small, autonomous companies). Therefore, in order to allow the neuroscience community to take full advantage of recent advances in multi-modal recording, CARMEN will develop a baseline data structure that is supportive of experimental co-registration. Due to the large volume of data that will be produced by the research work packages and new users, there is an initial requirement to hold in the region of 60TB of data. The relatively low and falling cost of modern storage makes this affordable, but raises challenges relating to the transfer, management, processing and archiving of such large volumes of scientific data. Whilst the primary data will be held in file storage, the derived data will be stored in a database in order to allow researchers to exploit the powerful functionality (especially indexing and querying). For example, CARMEN will investigate the utility of the spatiotemporal facilities provided by modern database systems in the analysis of neuronal spike trains. This innovation will raise new and exciting possibilities for the analysis of this data. The Storage Resource Broker (SRB – [21]) will be used to manage the primary data, as it can store petabytes of data robustly, and was used in the DAME [27] and BROADEN [44] projects. Databases will be exposed as services through the OGSA-DAI software [33].

2.2 Metadata.

Metadata is critical to ensure that the stored data can be discovered and understood. To allow CARMEN to provide rich metadata, it is first essential to ensure that enough information can be made available to provide an understanding of the experimental context in which the data were gathered. Therefore, the tools provided for importing data into the repository will enable the user to specify and reuse descriptions of the experimental context and conditions. Secondly, it is equally important that data derived by analysis can be accessed in the same way. Therefore, the analysis services provided by CARMEN will describe themselves in terms of their domain functionality. During workflow enactment this metadata will be used to generate automatic provenance traces.

Basic experimental metadata will be structured using the developing BrainML [22] standard where possible. BrainML is the current leading standard for neurophysiology data, and is endorsed by the Society for Neuroscience as the schema of choice for

new database development. Additionally, to establish coherence with other areas of systems biology that relate directly to neuroinformatics (e.g. genetics), we will investigate the applicability of other metadata representations from the biological community. For example, communities such as MIAME [25] and MIAPE (Minimal Information about a Microarray and Proteomics Experiment respectively) are starting to formulate generic scientific metadata. We will build on previous work from the myGrid [26] project, and use LSIDs (Life Science Identifiers – [28]) to draw relationships between data and metadata.

Support for service metadata will be based on previous work on 'Feta' in the myGrid project. Feta is a data model and implementation system, aimed at user-oriented service discovery for bioinformatics. It incorporates an ontological model of the domain, thus allowing services and their parameters to be described in domain terms. The Feta data model will be modified to allow it to describe neuroinformatics services. This will allow the OMII's Grimoires [29] registry to be used for storing and searching over both data and service descriptions. We will exploit the linkage between data and service descriptions to enable users to intelligently discover appropriate analysis services for both primary and derived data.

Both of these core components of the metadata will be used to record data provenance. We will extend the workflow enactment engine (see section 2.6) to automatically gather and store information about data derivation. Where possible, we will augment this with knowledge of the user's experimental context, ensuring that the purpose of the analysis is also stored. When combined with the basic experimental metadata, this will provide a very rich data discovery environment.

2.3 Data Analysis Services.
A key set of core services for data analysis will be provided (see section 3). These will include spike detection and spike sorting for the primary data, as well as services for derived data (e.g. cross-correlation). Initially, these will use well known algorithms, but novel algorithms will be developed within the project. Due to the high computational cost of running many of these services on large quantities of data (particularly in the guise of complex, parameter-constrained models), access to a compute Grid is required. Therefore, each CAIRN will contain a local compute Grid running on a cluster. Regional and National Grid resources, including the National Grid Service, the White Rose Grid, and the Newcastle Grid will also be integrated with the CAIRNs to provide additional resources. To achieve this, we will utilise the OMII GridSAM [30] to provide a standard interface to the various compute Grids. We will also integrate the Dynasoar [31] dynamic service deployment infrastructure with GridSAM so that services can be deployed (and un-deployed) on the underlying compute Grids as required, in order to meet changing demands.

2.4 Services for Annotation and Pattern Search in Time-Series Data.
Allowing users to efficiently locate patterns in the data is a crucial and novel requirement for the neuroscience that will be supported by CARMEN. Experiments will produce primary data that will be loaded into the CAIRN. Further processing will then generate derived data; for example spikes may be identified and stored. These data may then be reprocessed to find groups of spike events that represent some correlated, associated event. Thus, the repository will hold time-series data at a number of levels of abstraction. These data will only make sense if the events they contain (e.g. spikes) can be identified and annotated with 'what, where and when' metadata. Given the quantity and complexity of the data, neuroscientists will require tools to support this annotation process. For this, the project will build on the pattern search and matching technology developed by the DAME project. The Signal Data Explorer (SDE) [14, 32] can compare events in time-series based on: existing events, those stored in a file, or new events drawn by the user. It has the ability to match across multiple time-series (using a task planner tool) as well as allowing the user to tag the data with known events. To date, the main application of SDE has been searching vibration and performance data from aero engines. CARMEN will adapt this technology to detect events taking place within and across time-series data within given timing conditions. WP4 describes this work in detail (see section 3.4).

2.5 Facilities for the Dynamic Deployment of External Services.
It is unrealistic to consider providing a complete and comprehensive set of core analysis services for neurophysiology within the lifespan of CARMEN. Scientists will always have the need to invent and experiment with new analysis algorithms. Rather than force them to export the required data out of the CAIRNs to a client in order to achieve this (something that will be very, perhaps prohibitively, expensive when large datasets are involved), facilities will be provided to allow client-generated services to be deployed close to the data, within and beyond the CAIRN compute Grids. This will utilise the Dynasoar infrastructure.

2.6 Workflow Enactment Engine.
Workflows provide a way to represent and enact processes, such as complex data analyses, that involve data being processed by multiple services. The workflow enactment engine will allow clients to submit workflows to be enacted close to the data on which they will operate, removing the costly need to transfer data in and out of the CAIRN to an external enactor. We will investigate the suitability of two existing components to perform this task: OMII BPEL [34] and myGrid Taverna/Freefluo [35, 36].

2.7 Security Infrastructure.
To being with, the neuroscience community will have a pressing requirement to be able to control access to data and services. For example, researchers may wish some of the data they upload to be accessible only to themselves and their collaborators until the point at which they have completed and published their analyses. Further, there is a requirement to support commercial organisations that will have a legal requirement to control, secure and protect access to the data they store. The security infrastructure will, therefore, deal with the authentication and authorization of users, and will be configured so as to enforce a mutual IPR agreement defined by the user community. For this we will utilise the Gold [37] security infrastructure, which uses XACML [38] assertions for role and task based access control. To remove the need for users to manipulate XACML directly, we will allow data owners to specify their security preferences through a simple graphical interface.

2.8 Client Access.
CARMEN services will be made available as web services. A rich portal will be provided to allow users to interact with the CAIRN graphically (e.g. to upload data and annotate it with metadata; to locate and browse data; to create, run and monitor workflows etc.). This portal will play an equally important role in allowing the user community to manage their own interaction, by providing the means to establish virtual interest groups and secure collaborative environments. Training materials (e.g. online tutorials and troubleshooting materials) will be an integral part of the portal, as will news pages publicising events and research results.

2.9 CAIRN Integration.
It would be unrealistic to oblige the global neuroscience community to use one CAIRN-like node for all of their data. Instead, we anticipate that different communities, separated both by geography and focus, will create their own CAIRNs. CARMEN will, therefore, develop an integration layer that can present the data and services of a set of CAIRNs as if they were a single entity. This will be achieved by utilising and extending existing integration technologies. Distributed Query Processing software from the OGSA-DQP [39] project will be used to integrate data from distributed databases. The SRB allows universal access to data held in distributed datastores without knowledge of the location. At a higher level, globally unique identifiers and associated resolution services will be used to link data and metadata. For this we will build on the work on LSIDs in ^{my}Grid. The SDE system (see 2.4) already supports transparent searching across distributed data sets, though the project will look at enhancements to improve the intelligence of the search process, in particular to allow the system to reduce communication overheads. This will involve the transfer of knowledge and experience from the work on OGSA-DQP.

In summary, the CARMEN infrastructure will be built by adopting and, where necessary, extending existing e-Science infrastructure, developed or utilised by the investigators in previous projects (especially DAME, Gold, ^{my}Grid and the OMII). The first release of the system will largely be an integration of existing software with the aim of enabling the scientists in the spokes to begin uploading and analysing their data at an early stage in the project. The results of the work undertaken to address the research challenges will then begin to feed back into the infrastructure design, increasing the range and sophistication of the services that can be provided to the scientific community. These research challenges include: the design and exploitation of domain-specific metadata for neuroinformatics; a rich security model that can encompass multiple CAIRNs; locating patterns in time-series spike data across multiple levels of abstraction; dynamic service provisioning over compute grids; and seamlessly integrating data and services from multiple CAIRNs. Although the primary aim of the project is to support neuroscience, it is our assertion that the CARMEN infrastructure is sufficiently generic to support many other types of science.

3. Research Work Packages (WP1 – WP6)
The CARMEN e-Science infrastructure (WP0) will act to extend neuroscience by providing: (1) software services to process and analyse raw and derived neuronal activity data at a variety of levels (single cell to network); (2) a community-wide repository for neuronal activity data that may be interrogated by novel sequence recognition algorithms and which will permit various scales of data integration; and (3) a vehicle for creating a virtual laboratory, linking experimental neuroscientists and analysts at the experimental interface. The research work packages will develop and test the neuroscience specific functionality.

3.1 WP1 - Spike Detection and Sorting
Work Package Coordinator: Professor Leslie Smith (Stirling; 743h); **Co-Investigator:** Dr Rodrigo Quian Quiroga (Leicester; 495h); **Research Staff:** RAa (months 7-30; Stirling) and PHDa (months 1-36; Leicester); **Other Resources:** Equipment/Consumables: £10,000, Travel: £12,300.

Aims: Many CARMEN users will collect raw neurophysiological data in the form of electrical signals measured by extracellular electrodes (single or MEAs). The first critical steps in extracting accurate spike (action potential) sequences from these signals is the ability to resolve spikes generated by one neuron from the background noisy electrical signal generated by distant neurons or other sources (**spike detection**), and the ability to classify each spike as arising from one particular close-by neuron (**spike sorting**). Efficient application of these procedures enables users to recreate the original neuronal activity pattern with high fidelity. This is possible because the relative position of an electrode to different neurons means that recorded spikes can have distinguishable profiles. Furthermore, when electrodes are very close together (e.g. tetrodes or MEAs) one neuron may be recorded by several electrodes, allowing the application of coincidence detection to improve spike sorting. *L. Smith* has been applying spike detection algorithms to particularly noisy *in vitro* cultures [EPSRC Grant 'Talking to Nerve Cells'], and *Quiroga* has been working on spike sorting of *in vivo* signals. The aims of WP1 are to: (i) translate existing spike detection and sorting techniques into infrastructure services and (ii) develop improved spike detection and sorting techniques for the range of data types supported by CARMEN. Overall WP1 will provide a set of core spike detection and classification services that can be utilised in spike train analysis and modelling. Within the CARMEN repository many records will be stored as processed sequences leading to considerable data compression. However this has the disadvantages that important data may be lost and the quality of the spike sequences is dependent on the accuracy of the original detection and sorting. CARMEN will overcome this by allowing storage of raw records enabling users to have confidence in the spiking patterns by applying integral sorting and detection functions, which will be continuously updated. These software services may be applied to the users' own data or to stored data, and the Grid will enable this processing to run in (near-) real-time, allowing informative feedback to be provided during data acquisition.

Programme of Research: Initially, RAa (Stirling) will liaise with infrastructure developers to effect implementation of existing routines in CARMEN. This will provide early functionality which can be utilized by WP5 and WP6. This work will appraise the manufacturer-supported open software in Neuroshare [2], so it can be applied to raw data from many manufacturers' equipment. Newer techniques will require further development to allow them to be implemented as robust

services, in particular the de-noising and energy-based techniques [3, 5]. RAa will also develop existing test data generation software [4] (making it more realistic and able to emulate tetrode arrays) to make it more effective in assessing novel detection and sorting techniques. PHDa (Leicester) will develop the recent advances in spike sorting made by *Quiroga* [6, 7] to enable more effective interpretation of electrode signals. This method combines the wavelet transform which localizes distinctive spike features, with superparamagnetic clustering, allowing automatic classification of the data without assumptions such as low variance or Gaussian distributions. The current algorithm will be optimized for the different data sets, and implementations for on-line processing and for the interpretation of signals from tetrode arrays will be studied. In addition, the effect of pre-processing using de-noising techniques [5] will be investigated. *Quiroga* will also apply multivariate time-series analysis to measure synchronization within the resolved data [8]. Outputs will have strong links to WP2 and WP5 allowing the project to work towards convergence of methodology in order to apply generic solutions to both electrical signals and image-based recordings.

3.2 WP2 – Information Theoretic Analysis of Electrically- and Optically-Derived Signals

Work Package Coordinator: Dr Simon Schultz (Imperial College; 804h); **Co-Investigator**: Dr Stefano Panzeri (Manchester; 495h); **Collaborator:** Professor Michael Häusser (UCL); **Research Staff:** RAb (months 10-33; Imperial) and PHDb (months 13-48; Manchester); **Other Resources:** Equipment/Consumables: £11,000, Travel: £10,950.

Aims: The development of dye-based recording techniques enables the simultaneous detection of action potential-evoked signals from large numbers of neurons, both *in vitro* and *in vivo*, and CARMEN users are predicted to increasingly employ these types of data. Such imaging can provide localised information on ion transients resulting from spikes, down to the level of individual dendrites. In this respect, the spatial information inherent to imaging allows direct determination of the signal source, whilst analysis of the temporal variability allows determination of the underlying spiking activity. After extracting spiking data from images or electrical signals, one way to understand how neurons transmit information through a network is to concept-ualize the network as a communication channel. ***Information Theory***, the mathematical theory of communication, can then be applied to establish the 'neuronal code' by quantifying how much information is carried by different potential neuronal coding mechanisms. *Schultz* and *Panzeri* have been at the forefront of developing information theoretic algorithms for neuronal data. They have developed methods to alleviate the sampling problems of information measures and were the first to address in a systematic way the role of spike timing and spike correlations in neuronal population coding [9, 10]. To incorporate these processes into CARMEN, WP2 aims: (i) to extend spike detection functions to multimodal recording, including techniques that enable data to be captured optically, often simultaneously with electrophysiological methods; and (ii) to interface the signal data to information theoretic analysis. Overall WP2 will allow convergence of electrically- and optically-derived spike signals to enable CARMEN users to build large scale models of network activity (WP6).

Programme of Research: Initially, RAb (Imperial) will work with infrastructure developers to make the existing information theoretic tools for spike train analysis [9] available through CARMEN. WP2 will then adapt these algorithms for the analysis of optically-derived data. These analysis algorithms will either be applied to regions of interest (e.g. pixels in a given neuron) or on a pixel-by-pixel basis, and may be applied either after filtering using spike detection algorithms or directly to the optical signals to study non-spike-related aspects of the signal. To convert image data to spike occurrence times, spike detection methods for high resolution images will be developed. Code for interpreting the imaging data and techniques for coding the spatial information in the image will be determined. Additional imaging datasets will be required to develop the techniques for image interpretation and annotation, and for further evaluation of and application to the information theoretic analyses. Development work will be undertaken using records from two photon imaging of Ca^{2+} transients induced by complex spikes in large numbers of cerebellar neurons and captured at a temporal resolution of ~ 15 Hz [11]. However, the final spike detection routines will be generic, capable of handling data from dyes with different characteristics and adaptable to accommodate future development of novel voltage- and ion-sensitive dyes. They may also find application to other types of waveform data. Additional capabilities will include clustering algorithms capable of automatically choosing regions of interest for the information-theoretic analysis. In parallel, PHDb (Manchester) will work on novel methods for estimation of information conveyed by large neuronal populations, and will apply these methods to extracellular spike trains and multiple time series waveform data obtained from imaging. This analysis will be extended to graded neural signals, such as local field potentials, that reflect the average synaptic inputs to pyramidal neurons and, thus, will strongly link to the network modelling being conducted in WP6. The code developed will allow analysis and quantification of both how neurons encode information, and how neurons transmit information to one another. Both the datasets and the code for the analysis algorithms will be tested on CARMEN to ensure that these techniques work effectively on a wide diversity of data.

3.3 WP3 - Data-Driven Parameter Determination in Conductance-Based Models

Work Package Coordinator: Dr Kevin Gurney (Sheffield; 495h); **Co-Investigator:** Dr Paul Overton (Sheffield; 165h); **Collaborator:** Dr Ric Wood (Sheffield); **Research Staff:** RAc (months 13-36; Sheffield); **Other Resources:** Equipment/Consumables: £3,000, Travel: £7,000.

Aims: The analysis of patterns of spike activity and neuronal network properties (early phase deliverables of CARMEN) depend upon understanding the cellular properties that regulate spike timing. This may be investigated using biophysical neural models of membrane potential. However, a difficulty in developing such models is obtaining appropriate parameter values. This is because, whilst many parameters appear to remain constant across neurons, membrane ion channel densities (g_max.) vary adaptively to maintain specific target behaviour for the cell [12], their values depending on the morphology and past activity of individual neurons. Thus, estimates of g_max based on experimental measurement may be of little use in modelling. *Gurney* has developed a novel approach for obtaining g_max that is more efficient than existing techniques [13]. WP3 aims to adapt this technique to accommodate a variety of membrane dynamics (e.g. spiking behaviour) in order to: (a) build a canonical set of conductance-based models of varying complexity, and (b) obtain an analytic understanding of the search algorithm used. The power of existing search algorithms is that they can be applied to most systems of equations, but they typically require thousands of iterations to converge on a solution. The search technique to be integrated with CARMEN takes

advantage of the mathematical form of the system (noting the linear occurrence of g_max in the model equations) to deliver an efficient algorithm that rapidly converges on a solution. Using the Grid to develop these algorithms will facilitate handling of extremely large data sets and, by utilising a broad test bed of neurons from a range of sources, lead to generic modelling tools.

Programme of Research: *Gurney* and RAc (Sheffield) will develop the parameter search technique using several types of neurophysiological and morphological data. To construct a basic model requires both passive parameters and ion channel kinetics. Passive parameters will be obtained by fitting the response of a model to current-pulse data. These data will consist of a time-series of membrane potential values recorded intracellularly under current clamp after the application of neurotoxins. In general, the voltage responses reflect the activation, deactivation and inactivation kinetics and conductance characteristics of the ion channels in the neuronal membrane, and the analytic techniques will be directed at extracting parameter values from these responses. The process can be further constrained by morphological data to compartmentalise the model. Voltage-clamp data may also be required where the kinetics of a given ion channel are not well characterised. Finally, current-clamp data will be used in conjunction with the search to obtain the g_max. *Overton* will play a key role in interpreting these data. The search algorithm has previously been used in conjunction with simulations running under the neural simulation program GENESIS [13], and is limited to fitting the model output to non-spiking current injection data. While the existing code for the search is specific to the neural models employed, WP3 will develop a generic version for integration with CARMEN. This code will be developed in three major directions: (a) generalisation of existing code to allow it to interface with most major simulation languages; (b) addition of a graphical user interface that will enable researchers to take advantage of the search without the need to understand the code; and (c) extending the algorithm to enable searching on spiking data. CARMEN-facilitated collaborations will ensure these tools are designed to meet the needs of a diverse group of end users.

3.4 WP4 – Intelligent Database Querying

Work Package Coordinator: Dr Tom Jackson (York; 990h); **Collaborator**: Professor Jim Austin (York); **Research Staff:** RAd (Dr Bojian Liang; months 1-36; York); **Other Resources:** Equipment/Consumables: £4,500, Travel: £7,400.

Aims: The raw data to be stored by CARMEN consists of spatiotemporal signals expressed, either as time-series recordings collected by single electrodes or MEAs, or as image files, collected at regular intervals using various optical recording techniques. The data model to be implemented by CARMEN will allow these data structures to be integrated across space and time, so that experiments that make use of combined techniques can intuitively co-register data, and will ensure that secondary data (e.g. derived through analysis) are bound to both the raw data source and the experimental metadata. One way in which CARMEN can challenge and extend e-Science and neuroscience capabilities will be to allow users to query these raw and transformed data for abstracted (e.g. cut from existing data) or simulated (e.g. modelled or drawn) activity patterns that are of interest. This might, for example, allow the modeller to programmatically perform sweeping, serial iterations to improve the biological accuracy of their simulations, through interaction with real data, or allow neuroscientists to accurately correlate neuronal firing patterns with specific events. In tandem with the access to data that CARMEN will provide, this function will play a significant role in allowing theoreticians to accelerate the identification of flaws and limitations in hypotheses and algorithms. WP4 will implement this functionality in two ways: (a) at the high level by storing the secondary data in a database to utilise spatiotemporal pattern matching functions in the native query language and; (b) by applying and extending the utility of Signal Data Explorer (SDE) [32], originally developed by the DAME [27] project to search for patterns in temporal signals of vibration data across distributed repositories [14].

Programme of Research: Working with the infrastructure developers, RAd (York) will work on four main strands: (a) to improve the intelligence of the SDE search process. SDE is supported by two software components: Pattern Match Control (PMC) and Pattern Match Engine (PME). The neuroscience application will demand that the communication overhead between PMC nodes is reduced in order to allow the search tasks to be deployed across greater and more disparate volumes of data; (b) to extend the PME service to search for patterns in both raw signal and derived data (e.g. spike trains) discretely and simultaneously; (c) to evaluate how this flat file orientated tool should interact with the spatiotemporal query functionality that is provided by modern RDBMS systems (through the collaboration that will take place with Microsoft (see support letter), this may result in significant IP exploitation); and (d) to develop an improved user interface based on SDE's ability to allow users to draw the pattern that they wish to search for. For this final strand RAd will work with PHDc (WP5 - Plymouth) on high performance desktop visualisation in order to produce web services that allow the user to harness the SDE via the process inverse to the visualisation, e.g. by simulating or extrapolating from results that are displayed in the visualisation interface.

3.5 WP5. Measurement and Visualisation of Spike Synchronisation

Work Package Coordinator: Dr Stuart Baker (Newcastle; 743h); **Co-Investigators:** Professor Roman Borisyuk (Plymouth; 495h), Professor Jianfeng Feng (Warwick; 248h); **Collaborators:** Professor Ad Aertsen (Freiburg, Germany), Professor George Gerstein (Pennsylvania, USA), Dr Liz Stuart (Plymouth); **Research Staff:** RAe (Dr E Williams; months 13-48; Newcastle/Warwick), PHDc (months 13-48; Plymouth), RAf (months 19-42; Newcastle); **Other Resources:** Equipment/Consumables: £14,500, Travel: £44,375.

Aims: Central to CARMEN will be a software analysis toolkit which can be applied by users to their own or archived data using the large scale computational resources of the Grid. A range of analysis methods, adequate for single or small numbers of neurons, have been developed by members of the CARMEN consortium [e.g. 15, 16], and implementation of these will establish the initial toolkit. However, development of more advanced analytical techniques is required to handle large scale, simultaneous recordings arising from MEAs. In addition, methods to identify and extract information from patterns of spike trains are required in order to investigate the temporal coding employed by networks. WP5 aims to: (a) develop reliable and robust analysis techniques to address these issues, particularly sweeping statistical methods to test if measures show significant changes; (b) develop novel visualisation methods for displaying the results from these techniques, particularly those working in high-dimensional space; and (c) conduct real-time analysis of spike coding through a distributed Grid-enabled virtual laboratory.

Programme of Research: Working with *Baker* and *Feng*, RAe (Newcastle/Warwick) will develop novel measures of spike train synchronisation, including: (a) a reliable measure of instantaneous (single trial) neural firing rate based on inter-spike interval statistics. This will have advantage over current peri-stimulus time histograms in that it can be applied to rare stimulus events or to neurons which show adaptive responses to multiple stimuli. This will be an important precursor to measures of spike synchrony, as it allows construction of the null hypothesis (i.e. that cells code independently by firing rate alone); (b) Existing gravitational clustering methods for investigating synchrony between many neurons [17] will be developed further, particularly the methods for assigning significance to the findings and methods to examine how synchrony evolves during task performance or following a stimulus; (c) A related approach will compare recordings using MANOVA [18]. Although MANOVA can be applied to analysis of interspike intervals from small data sets, when applying this to data from large numbers of simultaneously recorded neurons, those neurons which show significance may be lost amongst those which do not. To overcome this, a version of MANOVA that detects the most significantly changed neurons in an array will be used to detect changes in the correlation between spike trains; (d) To detect nonlinear coherences between spike trains we will develop time-frequency analysis methods using a combination of short time Fourier transformation and continuous wavelet transformations to estimate the coherence in non-stationary signals; (e) To determine the causal relationship between signals is one of the challenging tasks in signal processing and we will develop algorithms for causality analysis and information flow. Using Granger causality, this method could demonstrate the relationship of neurons recorded in an electrode array [19]. Parallel to these analyses PHDc (Plymouth) will develop a multiple neuron analysis based on the correlation distance method of *Borisyuk* which assigns a distance metric relating one spike train to another. This will permit high level analysis of the network connectivity underlying an observed synchrony pattern. PHDc will also develop novel visualisation methods for displaying results from this clustering algorithm which works in high-dimensional space and hence presents visualisation challenges [20].

Finally, WP5 will conduct real-time analysis using the Grid-enabled virtual laboratory, thereby testing CARMEN for on-line distributed analysis and modelling. Such a real time interface could lead to a marked change in the way of working, particularly for complex and unique experiments which would benefit from the involvement of a distributed analysis team. Working with *Baker* and developers at Neuralynx, RAf (Newcastle) will develop NetCom software to facilitate near real-time, easy-to-implement data sharing and interaction with CARMEN services over the Grid and within partitioned networks (e.g. laboratory mini-Grids). RAf will also develop software methods to automate the manual element of multiple single-electrode operation (e.g. researchers continuously monitoring signals obtained from an array). This project will benefit from donation of Neuralynx data acquisition hardware for development (see letter of support) and will establish supportive tools and services (e.g. interfaces) that can render the technology more accessible for non-skilled programmers, simplifying integration and interoperation with Cheetah software. The proposed relationship will allow Neuralynx developers to receive training and develop expertise in Grid-based software authoring for neurophysiology, integrating commercial and academic beneficiaries of CARMEN.

3.6 WP6 - Multilevel Analysis and Modelling in Networks

Work Package Coordinator: Dr Evelyne Sernagor (Newcastle; 867h); **Co-Investigators:** Professor Miles Whittington (Newcastle; 248h), Dr V. Anne Smith (St. Andrews; 495h), Dr Tom Smulders (Newcastle; 248h), Dr Stephen Eglen (Cambridge; 248h), Dr Markus Kaiser (Newcastle; 248h); **Collaborators:** Dr Mark Cunningham (Newcastle), Dr Thomas Knopfel (RIKEN BSI, Japan); **Research Staff:** RAg (months 7-42; Cambridge/Newcastle), RAh (months 7-30; Newcastle/Cambridge), PHDd (months 1-36; St. Andrews/Newcastle); **Other Resources:** Equipment/Consumables: £49,500, Travel: £20,275.

Aims: Understanding activity dynamics within neuronal networks is a major challenge in neuroscience that requires simultaneous recording from large numbers of neurons. In addition, two major types of activity can contribute to understanding how neurons communicate within these networks: firstly, sub-threshold (non-spiking) synaptic activity that can be deciphered by examining changes in membrane potential elicited by neurotransmitter-driven modification in ionic conductances and, secondly, spiking activity. Using optical recording techniques or dense electrode arrays a comprehensive picture of the spatiotemporal dynamics of network activity can be obtained, particularly where this information is obtained concurrently. WP6 aims to integrate data from these different recording paradigms and to develop techniques both to resolve coordinate activity within large networks in cortex, cerebellum, hippocampus and retina, and to compare with the integrated 'field' signal. To this end WP6 will pursue three main activities: (a) integration of existing and novel network analysis techniques into CARMEN in order to build comprehensive models of network dynamics; (b) populating the CARMEN repository with data of exceptional quality and detailed provenance for analysis of network properties; and (c) development of new dynamic Bayesian network algorithms to trace paths of neural information flow in networks.

Programme of Research: In conjunction with infrastructure developers, RAg (Cambridge/Newcastle) will integrate existing network analysis methods into CARMEN. Code originally developed for application to specific systems will be made available for a wide variety of neural recording types. Contributions will include: dynamic Bayesian network algorithms for tracing paths of neural information flow (*V.A. Smith*); graph-theoretic tools for analysis of interconnected networks (*Kaiser*); and visualization and analysis tools for propagating activity within networks (*Eglen*). Close relations to WP2 will also ensure interfacing with information theoretic network analyses. RAg will also develop new methods for detecting motifs in large-scale networks (*Kaiser*); and apply information-theoretic measures to the analysis of spontaneous activity (*Eglen, Schultz*). In parallel, RAh (Newcastle/Cambridge) will populate the repository with data derived from a range of sources in order to evaluate the capabilities of CARMEN to analyse neuronal network activity. Population recordings arising from a variety of sources (e.g. intracellular responses, local/large-scale field potentials, optical methods) will make use of the resolved data format (see section 2.1), to enable an integrative analysis framework. The provision of aggregated compute resource will allow analysis of large data sets, rather than resorting to data reduction techniques (e.g. rasterisation) which compromise the spatiotemporal characteristics. Within the CARMEN consortium data will be derived from: multiple concurrent intracellular recordings from neurons in cortical networks; local field potential recordings (alone or with concurrent intracellular

recordings) to bridge between whole population activity (field potential) and single neuron contribution to on-going network activity (*Whittington/Cunningham*); large-scale multiple field potential recordings (>64 channels) for spatiotemporal analysis of network behaviour; and optical and MEA recordings of retinal waves (*Sernagor/Eglen*). In addition new data of detailed provenance will be collected in order to define the limits of generic applicability. For this, collaboration with *Knopfel* (RIKEN BSI) will provide access to transgenic models where spike activity can be combined with fast optical recordings. These data will be subject to analysis conducted by RAg and by PHDb (WP2). Finally, PHDd (St. Andrews) will develop new dynamic Bayesian network inference algorithms that can be applied to a variety of data sources for tracing the paths of information flow in networks. These algorithms are powerful computational tools for identifying putative causal interactions among variables, and have been used to capture linear, non-linear, combinatorial, and stochastic relationships among variables across multiple levels of organization (*V.A. Smith*). A combination of simulation and real data from CARMEN will be used to develop and evaluate these tools, particularly to address the influence of topology and sampling methodology. Initial development will be based on unit recordings obtained from multiple sites during specific behaviours (*Smulders*). The resulting information flow networks can serve as substrates for CARMEN's graph theoretic and motif detection tools (*Kaiser*) and activity propagation visualization and analysis techniques (*Eglen*). Analysis of structural properties of network flow will complement the information-coding and information-content properties of networks revealed in WP2. Overall WP6 will make extensive use of the Grid-enabled virtual laboratory environment and will work closely with infrastructure developers on the experimental framework and processes that will allow this to evolve (e.g. metadata, core services). This will add value by allowing a number of developers to have real-time input into the authoring of the new algorithms, which is currently uncommon.

4. Management and Business Case

4.1 Project Management

The Project Manager will have responsibility for maintenance and delivery of the project, working in direct consultation with the Lead Investigators (Austin, Ingram, L. Smith, Watson), and in partnership with the Research Work Package Coordinators (L. Smith, Schultz, Gurney, Jackson, Baker, Sernagor). The Project Management Team (Lead Investigators, Project Manager and WP Coordinators) will meet (virtually or in person) on a monthly basis to appraise the overall progress of the project against an integrated plan. This level 2 plan consists of the detailed workplan for WP0 (GANTT Chart) and the approved work plans for WP1-6 (milestones in GANTT chart). Approval of WP1-6 plans will be completed no less than one month in advance of work package commencement. The Project Manager will line manage the clerical support for the project. Line Management of infrastructure staff will be enacted by the expert teams at the infrastructure nodes (York – Austin, Jackson; Newcastle – Watson, Lord). Each of the Research Work Packages has a designated Coordinator, selected based on his/her knowledge of the relevant subject area. These Coordinators will formulate the detailed workplans and milestones for their own WP, which will be actively monitored by both the Coordinator and the Project Management Team. WP Coordinators will also have responsibility to ensure that suitable line management for research staff is deployed, delegating to the named student and RA supervisors as geographically appropriate.

To ensure that an agreed standard of operational service is provided, and to quickly address problems that could prevent delivery against WP milestones, a consultative service level agreement will be developed during the early stages of the project, and a Project Delivery Team comprising Grid Engineers, the Project Manager and WP Coordinators will meet (virtually) on a monthly basis to resolve issues. The Grid Engineers and the Project Manager will share the role of providing the Helpdesk Service, which will be deployed via telephone and email, as FAQs on the website, and through face-to-face consultation. It is an important facet of the project that the Project Manager and infrastructure staff will spend considerable time working in direct consultation with staff at each of the collaborating institutions. Further, all investigators and staff will be trained to make use of AccessGrids and WIKIs in order to communicate as they require.

To ensure that the work and the thrust of the project continue to be collaborative and cohesive, the full consortium will meet on an annual basis at a two day workshop. This workshop will allow direct training to be deployed economically, will allow consortium members to discuss new e-Neuroscience projects amongst themselves and with invited guests, and will provide inspiration by introducing CARMEN researchers to experts from the neuroinformatics and e-Science domains. The annual Steering Group Meeting will be held to coincide with the annual workshop. The remit of the Steering Group will be to ensure that CARMEN is complementary to, and can learn from the work of, established international neuroinformatics service providers. To ensure that this can be achieved, effort has been made to assemble a broad and influential advisory team of international experts and industrialists [Shiro Usui (RIKEN BSI, Japan - chair), Daniel Gardner (Cornell, USA), Goutham Edula (Astra Zeneca Lund, Sweden), Phil Boden (NeuroServe Ltd, Cambridge), others to be approached - see support letters].

4.2 Risk Management and Mitigation

Fig. 4 outlines major threats to the success of the CARMEN Project, and provisions for mitigation:

	Description	**Mitigation**	**Classification**
1.	Inability to recruit infrastructure/research staff within projected timescales.	Research WPs are phased. Work plans will be required in advance of research commencement for project management scrutiny.	Moderate
2.	Loss of key staff.	Research and Infrastructure WPs are phased to rationalise the burden on the infrastructure services.	Moderate
3.	Underperformance of spoke/s, or difficulty in managing against the agreed workplan.	Milestones will be defined and actively monitored in partnership with the Project Management Team.	Low
4.	Growth of customer-base hindered due to security concerns.	GOLD (XACML) security infrastructure will be used. Commercial standards for secure archive will be met.	Moderate

5.	Inability to establish CARMEN as a neuroinformatics system that is in harmony with the INCF.	Steps are being taken to establish CARMEN as part of a funded UK Neuroinformatics Node. Influential INCF members have agreed to join the Steering Group.	Moderate
6.	Lack of commercial/academic interest in using CARMEN, and hence inability to sustain the service.	A small grant fund to support early-stage customers is included in the project budget (£100,000). Prospective commercial users have joined the CARMEN Steering Group. The Project Manager will market and develop the CARMEN service, conducting demos and feedback sessions. The Project Management Team will lobby Research Councils and journals to recommend CARMEN for archiving neurophysiology data. These will generate sales channels.	Low
7.	Lack of confidence in the technical and ethical rigour of the service.	Business development and research will not commence until after beta testing (end year 1). Systems administration is embedded within existing university and Grid systems. User groups will be allowed to establish, and will drive improvements and refinements to the service.	Moderate
8.	Skills shortage amongst prospective users.	Training will be provided. User groups will be supported by the Project Manager and Helpdesk Service.	Moderate/High
9.	Inability to train new users after the project is completed.	Training materials will be published online, to address both practical use and practical reasons for use.	Moderate/High
10.	Inability to formulate detailed, extensible and appropriate semantic metadata.	Shiro Usui (Japan INCF) and Daniel Gardner (BrainML) to join the CARMEN Steering Group. Discrete and expert staff time is allocated to consult with users about metadata.	Moderate

4.3 Community Engagement and Business Plan

Sustaining and evolving the CARMEN service requires academic and commercial research groups to commit funds. For this reason, new CARMEN-supported projects will be fully cost-recovered as of the launch of operational service (end year one). To mitigate anticipated initial reluctance to pay, funds to support a small, competitive grants scheme are included in the project budget (£100,000). New project developers will be encouraged to bid into this fund to cover data storage and staff costs incurred by their project. This model is necessary to establish a stable, supportive, and *willing to pay* user community. **Fig. 5** illustrates the way in which pricing will be constructed. Possible sales targets have also been outlined.

PRICE ELEMENTS			**TYPICAL PROJECT**		**NEW PROJECT TARGETS**			
Cost Component	**Unit**	**Cost**	**Units (Avg)**	**Cost (Avg)**	**Volume**			**Revenue**
Hardware Storage	100GB	£750	1.00	£750	**2+3** {	Data Storage	1	**£10,000**
Hardware Depreciation	100GB	£750	1.00	£750		Data Access	4	**£40,000**
Account Management	1%FTE	£1,000	2.00	£2,000	**4** {	Data Storage	2	**£20,000**
Grid Engineering	1%FTE	£850	5.00	£4,250		Data Access	8	**£80,000**
Systems Administration	1%FTE	£350	3.00	£1,050	**5** {	Data Storage	4	**£40,000**
Administration	Per 1	£1,200	1.00	£1,200		Data Access	16	**£160,000**
			Price (Avg):	**£10,000**				

Figure 5: Projections for cost recovery in years 2, 3, 4 and 5 of CARMEN

In the medium-long term, the nature of the research dictates that it will be necessary for UKRCs to cover a large proportion of the cost (although some income will be obtained from commercial sources, e.g. pharmaceuticals), either in the form of support for strategic infrastructure, or by individual researchers including the cost of data publication in a curated repository in their research grant applications. For this reason, the CARMEN Project Management Team will work to develop two important sales channels: (a) to become the favoured repository for neurophysiology data by UKRC and journals, and; (b) to campaign for the development of the UK INCF node. These lobbying efforts are already in progress. A secondary income stream will also be explored in leased deployment of commercial neuroscience web-services (see supporting letter from Casey Stengel, Neuralynx). The Project Investigators and Project Manager will promote CARMEN by conducting demonstrations and making site visits.

Lastly, the CARMEN consortium will leverage its already prominent position within the UK neurophysiology community by hosting workshops, and providing on-line and off-line training for Grid-enabled neurophysiology. The latter will coincide with the annual workshop (open to the wider community). To encourage dialogue between researchers, the CARMEN portal will allow users to establish interest and self-support groups (see section 2.8). Evaluative user groups will emerge from this. An outline collaborative relationship has been developed with the NeuroGRID MRI system [42] (a repository for structural MRI data), which will explore future possibilities for integrative research through discussion to formulate the UK-INCF position.

4.4 International Engagement

Two important international links have been developed to ensure that CARMEN is in accordance with the INCF effort: (i) Supported by the MoU between RIKEN Brain Sciences Institute (Tokyo, Japan) and the Institute of Neuroscience (University of Newcastle), the CARMEN project will pursue functional integration with the Japanese INCF Node (Director: Shiro Usui, RIKEN). This collaboration will be in the form of four technical meetings (costs shared between RIKEN and CARMEN), to

engineer and implement an integration strategy. (ii) Professor Daniel Gardner (Cornell University), who has agreed to join the Steering Group, is Director of the US BrainML neurophysiology standards and data repository and will collaborate closely with the project through provision of BrainML software (see support letter). This link will be particularly important to the development of the CARMEN metadata. It is anticipated that these collaborations will allow CARMEN to be central to the formation of the UK INCF Node, and to be instrumental in the INCF metadata standards for neuroscience effort.

4.5 Commercial Engagement

The following organisations will contribute to the project:

(i) ***NeuroServe*** (Cambridge, UK) [41] has a critical requirement for improved analysis software and third party data archiving for neurophysiology. This was instrumental in initiating the CARMEN proposal. NeuroServe will contribute data from commercial experiments and staff consultancy to the value of £282,600.

(ii) ***Neuralynx*** (Tuscon, USA) [40] are one of the premier manufactures of MEA acquisition hardware and software, and have begun to establish corporate interest in neuroscience metadata in order to improve the intelligence their software. To extend existing software products for Grid use and establish a Grid-aware software team, Neuralynx will provide equipment, developer time and consultancy to the value of £90,000.

(iii) ***Astra Zeneca*** (R&D Lund, Sweden) employs experimental electrophysiology and related neuroscience recording techniques to validate pharmaceutical compounds and therapies across an international network of scientific research centres. To explore the suitability of the CARMEN system as an integral component of their research strategy, Astra Zeneca will donate consultancy to an estimated value of £20,000 (see support letter).

(iv) ***Microsoft Research*** (Redmond, USA) are interested in the CARMEN project as an opportunity to derive research challenges that will lead to enhancements to their database and software tools for bioscience (see support letter).

(iv) ***Cybula*** (York, UK) are expert in the development of pattern matching software for very large, time-series datasets, and wish to derive new insight into their product by applying Signal Data Explorer (SDE) [14,32] to raw and transformed neurophysiology data expressing multiple time-series signals. To allow the consortium members to use the SDE service, and to train the Grid engineering team that will work with it, Cybula will provide software and consulting (training) to the value of £30,000.

5. Cost Justification (Figures quoted are FEC):

1. Investigator and research staff effort has been indicated for each work package. Provision for project management has also been made within these figures. These figures have been costed subject to current salary rules which, in some cases, may alter following the implementation of fair measures.

2. The Project Manager will ensure the successful discharge of the project by conducting operational management. Clerical costs to the value of **£17,843** (Newcastle Clerical Grade 3; 25% FTE for 4 years) are included to provide administrative and financial support and, and maintain the web presence. The Project Manager will take responsibility for business development. Systems Administration for backup and maintenance of the Grid system is requested on the basis of 25% FTE over 4 years (**£30,481** - York, **£31,678** - Newcastle)

3. **£346,000** for purchasing and maintaining the Grid clusters (processors and disk storage; 60TB archive) installed at Newcastle and York to store the data: Newcastle [disk storage (£150,000), server cooling (£15,000)]; York [disk storage (£150,000), server cooling (£15,000), presence cards (£16,000)].

4. **£8,000** for purchasing new multielectrode array equipment that is required to conduct WP6 at Newcastle.

5. **£38,000** for purchase of laptops (14 x £1,500 + 1 x £2,000) for the new research staff and students and workstations (6 x £2,500) for the Grid teams at Newcastle and York.

6. **£10,000** for purchase of personal AccessGrid equipment to hold distributed meetings, for partners that do not have access to a purpose built facility (Stirling, Leicester, Warwick, Plymouth and St. Andrews; each £2,000).

7. **£19,000** for purchase of software licenses for new research and infrastructure staff, and students (19 x £1,000/person). This is necessary due to the present unwillingness of the industry providing the analysis software to supply web services.

8. **£18,730** for import and husbandry of transgenic mice (from RIKEN, Japan), and rats, required to conduct the multi-modal neurophysiological recording in WP6 at Newcastle.

9. **£15,000** for laboratory consumables that will be used during dye-based imaging experiments at Imperial College (£5,000 - RWP2), and during mutli-modal experiments at Newcastle (£10,000 - WP6).

10. **£12,500** for stationery and computing consumables: Newcastle [£5,000]; York [£3,000]; Stirling [£500]; Leicester [£500]; Manchester [£500]; Imperial College London [£500]; Cambridge [£500]; Warwick [£500]; Plymouth [£500]; Sheffield [£500]; St. Andrews [£500].

11. **£20,000** to host the annual workshops, to provide hospitality for steering group and other business meetings, to pay for the attendance of guest speakers at CARMEN events, and to cover the production costs for offline training materials.

12. **£25,000** to cover the cost of promotion and dissemination. This will allow a website and corporate stationery to be designed and produced, provide funds to support offline presence (e.g. posters, display stands and booths), at international meetings (e.g. Super Computing, GGF, Society for Neuroscience), and to service new relationships created by business development activities with regular meetings.

13. **£100,000** to provide a small grants fund to support CARMEN business development. Prospective users will submit proposals to this fund to pay for the CARMEN effort cost that will be incurred by their project.

14. **£20,000** to cover the cost of staff recruitment and interviewing across the project.

15. **£93,025** to cover the cost of travel. This will allow WP members to meet face-to-face at appropriate intervals, will service associative collaborations identified in individual WPs, will allow the consortia to convene at the annual workshop, will

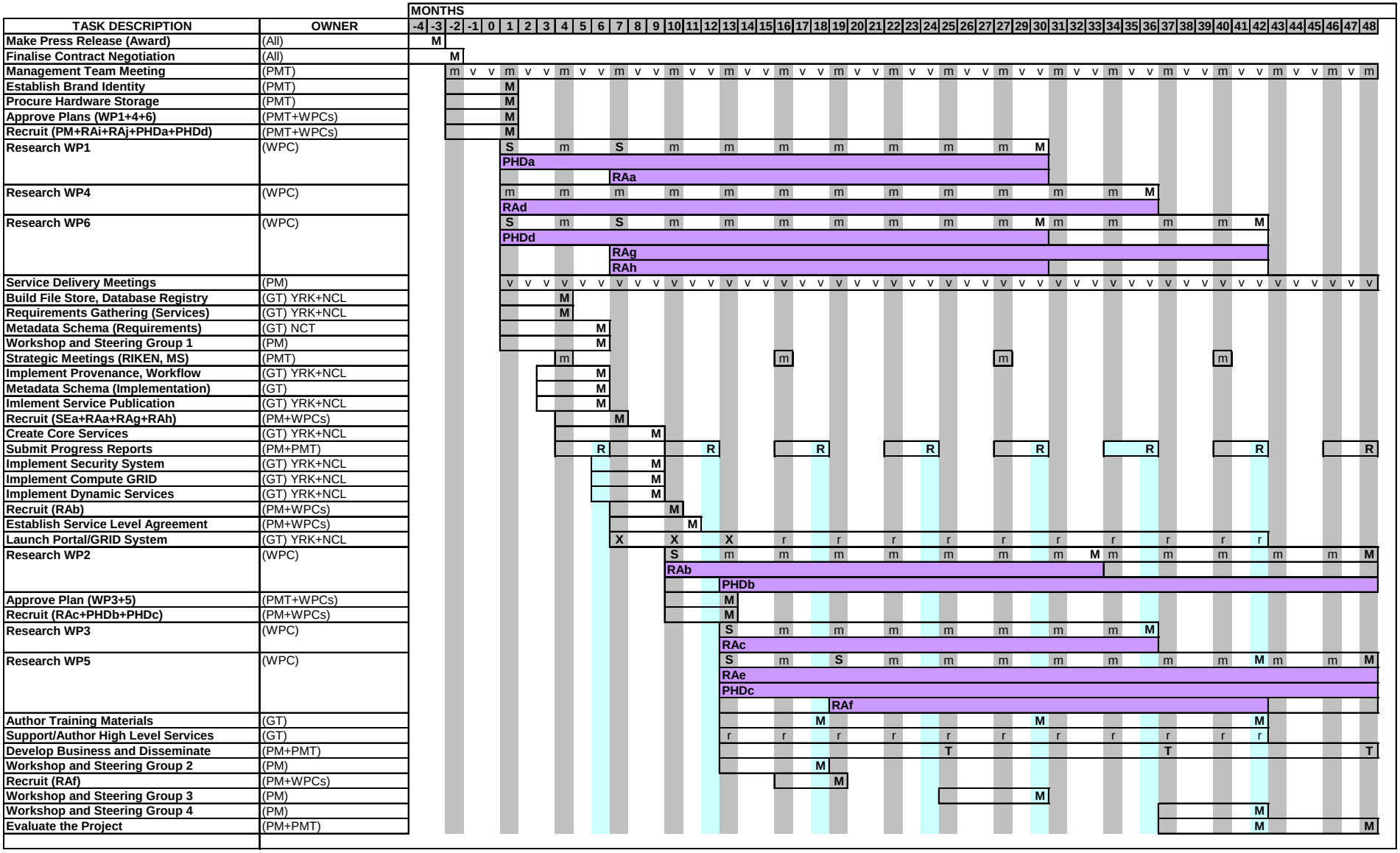
cover travelling expenses for steering group members to attend meetings, will allow the proposed strategic relationships with Astra Zeneca (Sweden), Microsoft (US), Neuralynx (US) and RIKEN (Japan) to be serviced with research visits, and will allow investigators and research staff to promote the project by making conference presentations and visiting new collaborators: Newcastle/CARMEN [£44,850]; York [£11,400]; Stirling [£5,500]; Leicester [£2,900]; Manchester [£2,725]; Imperial College London [£2,350]; Cambridge [£2,400]; Warwick [£3,250]; Plymouth [£11,250]; Sheffield [£2,850]; St. Andrews [£3,550].

16. **£45,375** to cover the cost of accommodation for the above journeys (**16**): Newcastle/CARMEN [£23,850]; York [£6,075]; Stirling [£3,525]; Leicester [£2,025]; Manchester [£1,125]; Imperial College London [£1,050]; Cambridge [£1,725]; Warwick [£900]; Plymouth [£2,325]; Sheffield [£1,650]; St. Andrews [£1,125].
17. **£24,775** to cover the cost of subsistence for the above journeys (**16**): Newcastle/CARMEN [£12,550]; York [£3,100]; Stirling [£1,500]; Leicester [£775]; Manchester [£1,000]; Imperial College London [£700]; Cambridge [£625]; Warwick [£850]; Plymouth [£1,525]; Sheffield [£1,000]; St. Andrews [£1,150].
18. **£23,250** to cover the cost of conference fees for the above journeys (**16**): Newcastle/CARMEN [£8,750]; York [£4,250]; Stirling [£2,500]; Leicester [£1,000]; Manchester [£1,000]; Imperial College London [£1,000]; Cambridge [£500]; Warwick [£750]; Plymouth [£1,000]; Sheffield [£1,500]; St. Andrews [£1,000].
19. **£6,495** to cover the cost of the levies that some institutions impose to cover their costs for technical and laboratory support staff: Stirling [£5,298]; Leicester [£597]; St. Andrews [£600].
20. **£362,471** to pay the estates support charges under FEC: Newcastle/CARMEN [£203,034]; York [£78,590]; Stirling [£32,674]; Leicester [£3,139]; Manchester [£3,245]; Imperial College London [£20,267]; Cambridge [£512]; Warwick [£1,971]; Plymouth [£2,886]; Sheffield [£14,078]; St. Andrews [£2,075].
21. **£1,334,058** to pay the indirect support charges under FEC: Newcastle/CARMEN [£637,210]; York [£348,072]; Stirling [£109,479]; Leicester [£8,513]; Manchester [£10,320]; Imperial College London [£105,540]; Cambridge [£5,590]; Warwick [£4,569]; Plymouth [£8,680]; Sheffield [£87,540]; St. Andrews [£8,545].

References:

[1] Amari S-I, and 29 others (2002), Neuroinformatics: the integration of shared databases and tools towards integrative neuroscience, J Integr Neurosci 1:117–28; **[2]** http://neuroshare.sourceforge.net/index.shtml; **[3]** Mtetwa N, **Smith LS** (2006) Smoothing and thresholding in neuronal spike detection. Neurocomputing (in press); **[4]** Mtetwa N, **Smith LS** (2006) An analysis of neuronal spike detection techniques which use smoothing and thresholding. J Neurosci Meth (submitted); **[5] Smith LS** - http://www.cs.stir.ac.uk/~lss/noisyspikes/index.html; **[6] Quiroga RQ**, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurones in the human brain, Nature 435:1102-07; **[7] Quiroga RQ**, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comp 16:1661-87; **[8]** Pereda E, **Quiroga RQ**, Bhattacharya J (2005) Nonlinear multivariate analysis of neurophysiological signals. Prog Neurobiol 77:1-37; **[9]** Pola G, Thiele A, Hoffmann K-P, **Panzeri S** (2003) An exact method to quantify the information transmitted by different mechanisms of correlational coding. Network 14:35-60; **[10] Schultz SR**, **Panzeri S** (2001) Temporal correlations and neural spike train entropy. Physical Rev Lett 86:5823-6; **[11] Schultz SR**, Kitamura K, Häusser M (2005) Two-photon imaging of sensory-evoked population activity in cerebellar Purkinje cell dendrites in vivo. Soc Neurosci Abstr 35:178.1; **[12]** Golowasch J, Abbott LF, Marder E (1999) Activity-dependent regulation of potassium currents in an identified neurone of the stomatogastric ganglion of the crab *Cancer borealis*. J Neurosci 19:RC33, 1-5 **[13]** Wood R, **Gurney K**, Wilson C (2004) A novel parameter optimisation technique for compartmental models applied to a model of a striatal medium spiny neurone. Neurocomputing 58-60C:1109-16; **[14] Jackson T**, **Austin J**, Jessop M, Linag B, Pasley A, Ong M, Allan G, Kadirkamanathan V, Thompson H, Fleming PJ (2005) Distributed Health Monitoring for Aero-Engines on the Grid: DAME. Proceedings of IEEE Aerospace, Montana, March 2005; **[15]** Rossoni E, **Feng JF** (2005) Non-parametric approach to extract information from interspike intervals J. Neurosci. Methods (in press); **[16] Baker SN**, Spinks R, Jackson A, Lemon RN (2001) Synchronization in monkey motor cortex during a precision grip task. I. Task-dependent modulation in single-unit synchrony. J Neurophysiol 85:869-85; **[17] Baker SN**, Gerstein GL (2000) Improvements to the sensitivity of gravitational clustering for multiple neurone recordings. Neural Comp 12:2597-620; **[18]** Bonny L, Nicol A, Kendrick K, **Feng JF** (2005) Applications of multi-variate analysis of variances (MANOVA) to multi-electrode array data. J Neurosci Meth 146:22-41; **[19]** Cheng YH, Rangarajan G, **Feng JF**, Ding MZ (2004) Analyzing multiple nonlinear time series with extended Granger causality, *Phys. Lett A*, 234 26-35; **[20]** Stuart L, Walter M, **Borisyuk R** (2002) Visualisation of synchronous firing in multi-dimensional spike trains. Biosystems 67:265-79; **[21]** Storage Resource Broker http://www.sdsc.edu/srb/); **[22]** BrainML http://brainml.org/goto.do?page=.home; **[23]** NeuroML http://www.neuroml.org/; **[24]** Databasing the Brain. Featured article, Nature: http://www.nature.com/nature/journal/v406/n6798/full/406822a0_fs.html; **[25]** Minimum Information about a Microarray Experiment MAIME: http://www.mged.org/Workgroups/MIAME/miame.html; **[26]** myGrid Project: http://www.mygrid.org.uk/; **[27]** Distributed Aircraft Maintenance Environment Project (DAME): http://www.cs.york.ac.uk/dame/; **[28]** Life Science Identifiers (LSIDs) Project: http://lsid.sourceforge.net/; **[29]** OMII Grimoires Registry: http://www.omii.ac.uk/mp/mp_grimoires.jsp; **[30]** OMII Gridsam Job Submission and Monitoring Web Service: http://www.omii.ac.uk/downloads/product.jsp?productid=44; **[31]** Dynosoar Project: http://www.neresc.ac.uk/projects/dynasoar/; **[32]** Signal Data Explorer: http://www.cybula.com/flyers/SignalData.pdf; **[33]** OGSA-DAI: http://www.ogsadai.org.uk/; **[34]** OMII BPEL: http://www.omii.ac.uk/mp/mp_bpel.jsp; **[35]** Taverna: http://www.mygrid.org.uk/index.php?module=pagemaster&PAGE_user_op=view_page&PAGE_id=44; **[36]** Freefluo: http://www.mygrid.org.uk/index.php?module=pagemaster&PAGE_user_op=view_page&PAGE_id=54; **[37]** Gold Project: http://www.goldproject.ac.uk/; **[38]** XACML: eXtensible Access Control Markup Language, Version 2.0, ed. T. Moses, OASIS Standard, Feb 2005; **[39]** OGSA-DQP: http://www.ogsadai.org.uk/about/ogsa-dqp/; **[40]** Neuralynx: http://www.neuralynx.com/; **[41]** NeuroServe: http://www.neuroserve.co.uk/; **[42]** NeuroGRID: http://www.neurogrid.ac.uk/; **[43]** Bio-ontologies Workshop: http://www.bio-ontologies.man.ac.uk; **[44]** BROADEN: www.cs.york.ac.uk/dame/broadenposter.pdf.

Project Plan (GANNT Chart)



KEY: Milestone: **M**; Report: **R**; Release: **X**; Target: **T**; Staff Start: **S**; Meeting: m; Virtual Meeting: v; Refinement: r.

OWNER KEY: Project Management Team: PMT; Work Package Coordinator: WPC; Project Manager: PM; Grid Team (RAi+RAj+RAk+RAl+SEa): GT.