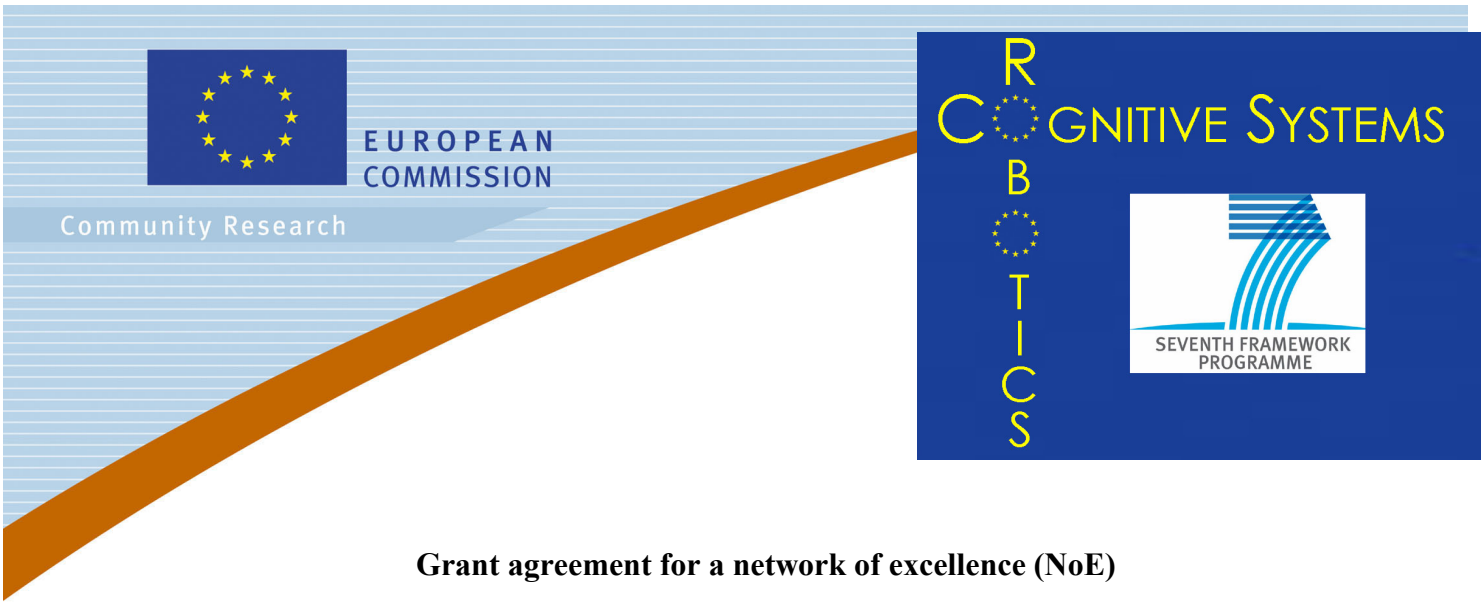

Grant agreement for a network of excellence (NoE)

Annex I – Description of Work

Project acronym: *PASCAL2*
Project full title: Pattern Analysis, Statistical Modelling and Computational Learning 2
Grant agreement No: 216886

Beneficiary Number	Beneficiary name	Beneficiary short name	Country	Date enter project	Date exit project
1 (Coordinator)	University of Southampton	SOTON-ECS	UK	Month 1	Month 60
2	University College London	UCL	UK	Month 1	Month 60
3	University of Edinburgh	UEDIN	UK	Month 1	Month 60
4	Centre National de Recherche Scientifique	CNRS	FR	Month 1	Month 60
5	XEROX SAS	XEROX SAS	FR	Month 1	Month 60
6	Jozef Stefan Institute	JSI	SI	Month 1	Month 60
7	Università degli Studi di Milano	UNIMI	IT	Month 1	Month 60
8	University of Bristol	UOB	UK	Month 1	Month 60
9	University of Manchester	UNIMAN	UK	Month 1	Month 60
10	Helsingin yliopisto	UH	FI	Month 1	Month 60
11	Fondation de l'Institut Dalle Molle d'Intelligence Artificielle Perceptive	IDIAP	CH	Month 1	Month 60
12	Stichting Centrum voor Wiskunde en Informatica	CWI	NL	Month 1	Month 60
13	Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.	FRAUNHOFER	DE	Month 1	Month 60
14	Max-Planck-Gesellschaft zur Förderung der Wissenschaften	MPG	DE	Month 1	Month 60
15	Helsinki University of Technology	TKK	FI	Month 1	Month 60
16	Bar-Ilan University	BIU	IL	Month 1	Month 60
17	Universite Pierre et Marie Curie-Paris 6	UPMC	FR	Month 1	Month 60

Table of Contents

A1. Budget breakdown and project summary ... 3
A.1.1 Overall budget breakdown for the project ... 3
A.1.2 Project summary ... 4
B1. Concept and objectives, long term integration, Joint Programme of Activities ... 5
B.1.1 Concept and project objective(s) ... 5
B.1.1.1 Main Challenges ... 5
B.1.1.2 General methodology ... 6
B.1.1.3 Milestones ... 7
B.1.1.4 Measurable results ... 7
B.1.2 Long term integration ... 7
B.1.3 Joint Programme of Activities ... 7
B.1.3.1 Overall strategy and general description ... 7
B.1.3.2 Timing of work packages and their components ... 8
B.1.3.3 Work package list /overview ... 9
B.1.3.4 Deliverables list ... 10
B.1.3.5 Work package descriptions ... 11
B.1.3.6 Efforts for the full duration of the project ... 19
B.1.3.7 List of milestones and planning of reviews ... 21
B2. Implementation ... 22
B.2.1 Management structure and procedures ... 22
B.2.1.1 Management Principles ... 22
B.2.1.2 Management Structure ... 22
B.2.1.3 Management Process ... 22
B.2.1.4 Budgeting ... 23
B.2.1.5 Internal bidding ... 24
B.2.2 Beneficiaries ... 25
B.2.2.1 University of Southampton ... 25
B.2.2.2 University College London ... 26
B.2.2.3 University of Edinburgh ... 27
B.2.2.4 Centre National de Recherche Scientifique ... 28
B.2.2.5 XEROX SAS ... 30
B.2.2.6 Jožef Stefan Institute ... 31
B.2.2.7 Università degli Studi di Milano ... 32
B.2.2.8 University of Bristol ... 33
B.2.2.9 University of Manchester ... 34
B.2.2.10 University of Helsinki ... 35
B.2.2.11 Fondation de l'Institut Dalle Molle d'Intelligence Artificielle Perceptive ... 36
B.2.2.12 Stichting Centrum voor Wiskunde en Informatica ... 37
B.2.2.13 Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. ... 38
B.2.2.14 Max-Planck-Gesellschaft zur Förderung der Wissenschaften ... 39
B.2.2.15 Helsinki University of Technology ... 40
B.2.2.16 Bar-Ilan University ... 41
B.2.2.17 Universite Pierre et Marie Curie-Paris 6 ... 42
B.2.3 Consortium as a whole ... 43
B.2.3.1 Sub-contracting ... 43
B.2.3.2 Other countries ... 43
B.2.3.3 Additional beneficiaries ... 44
B.2.3.4 Third Parties ... 44
B.2.4 Resources to be committed ... 44
B3. Potential impact ... 45
B.3.1 Strategic impact ... 45
B.3.2 Spreading excellence, exploiting results, disseminating knowledge ... 45
B4. Ethical issues ... 46
B5. Consideration of gender aspects ... 46

Part A

A1. Budget breakdown and project summary

A.1.1 Overall budget breakdown for the project

Table 1 illustrates the budget breakdown for the overall project. Note that due to the responsive nature of parts of the project it is not possible at this time to provide a precise breakdown between all beneficiaries. However, an indicative breakdown is given which assumes equal contributions from all beneficiaries, plus the additional management budgets at the coordinating sites.

Table 1 Budget breakdown

Beneficiary Number *	Beneficiary short name	Estimated eligible costs (whole duration of the project)			Total A + B + C	Total receipts	Requested EC contribution
		RTD (A)	MGT (B)	OTHER (C)			
1	SOTON-ECS	426,383.69	417,304.99	0.00	843,688.68	0.00	737,092.76
2	UCL	3,354,608.00	2,851.24	0.00	3,357,459.24	0.00	2,518,807.24
3	UEDIN	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
4	CNRS	806,400.00	3,000.00	0.00	809,400.00	0.00	606,300.00
5	XEROX SAS	302,400.00	1,500.00	0.00	303,900.00	0.00	152,700.00
6	JSI	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
7	UNIMI	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
8	UOB	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
9	UNIMAN	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
10	UH	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
11	IDIAP	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
12	CWI	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
13	FRAUNHOFER	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
14	MPG	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
15	TKK	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
16	BIU	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
17	UPMC	201,600.00	1,500.00	0.00	203,100.00	0.00	152,700.00
TOTAL		**7,510,591.69**	**444,156.23**	**0.00**	**7,954,747.92**	**0.00**	**6,000,000.00**

A.1.2 Project summary

PASCAL2 builds on the FP6 PASCAL Network of Excellence that has created a distributed institute pioneering principled methods of pattern analysis, statistical modeling, and computational learning (see http://www.pascal-network.org/). While retaining some of the structuring elements and mechanisms (such as the semi-annual Themes, and the Pump-Priming and Challenges programmes) of its predecessor NoE, PASCAL2 refocuses the institute towards the emerging challenges created by the ever expanding applications of adaptive systems technology and their central role in the development of artificial cognitive systems of different scales. Learning technology is key to, for instance, making robots more versatile, effective and autonomous, and to endowing machines with advanced interaction capabilities. The PASCAL2 Joint Programme of Activities (JPA) responds to these challenges not only through the research topics it addresses but also by engaging in technology transfer through an Industrial Club to effect rapid deployment of the developed technologies into a wide variety of applications. In addition, its Harvest sub-programme provides opportunities for close collaboration between academic and industry researchers. Other noteworthy outreach activities include curriculum development, brokerage of expertise, public outreach, and liaison with relevant R&D projects. Furthermore, PASCAL2 has adopted an open membership policy allowing for active inclusion in Network activities, of researchers working at non-beneficiary institutions.

PART B

B1. Concept and objectives, long term integration, Joint Programme of Activities

B.1.1 Concept and project objective(s)

During the last decade very significant advances have been made in the development and application of adaptive systems. Statistical analysis and algorithmic development have been combined to produce a new generation of well-founded techniques that have proved effective in a wide range of applications, from bioinformatics through document analysis to computer vision. We can summarise the current state of play:

- For well understood domains classification, ranking and regression tasks are being widely adopted using off-the-shelf techniques such as kernel methods;
- Significant progress has been made in our understanding of how to approach multi-class/complex output learning, though here the methods cannot deliver standardised solutions;
- Significant progress has been made in developing techniques for complex datatypes and domains, though again standardised solutions for multimodal and complex data are not available;
- Significant progress has been made for some tasks that involve partial or delayed feedback, but there is no comprehensive methodology for these types of task;
- Significant progress has been made in our understanding of the effects of representation and architecture on the learning task, however, a key challenge is to learn about their interface and interaction to support cognitive rich models.

In this development the main application areas that have been considered are text, images, expression arrays and sequence data in bioinformatics, time series data and speech.

The rest of this section is organised as follows: We first highlight the *main challenges* facing the research area, showing their direct relevance to the work programme objectives as well as their potential impact; we then discuss the *general methodology* to be adopted in the research finishing with a list of measurable outcomes that can be used to assess the progress of the network.

B.1.1.1 Main Challenges

There are four main fundamental challenges that adaptive systems research is currently facing together with issues concerning the scalability and roll-out of the technology. We first consider the fundamental challenges.

The first challenge is best characterised by complexity in prior knowledge. An example of this type of situation is that promoted by the PASCAL challenge in Textual Entailment. This widely cited development pioneered by PASCAL requires systems to classify pairs of sentences as positive if the second sentence is entailed by the first. Even though this is an example of a well-understood classification task, the data is highly structured and requires detailed domain knowledge for the development of effective solutions. The related task of machine translation is the subject of a STREP project SMART involving PASCAL researchers. Language, grammar and semantics are but one example of complex prior knowledge, another which is being investigated is the fitting of complex stochastic differential equation models to observations, be it in climate or gene network modelling. The objective is to develop a toolkit of techniques that can be applied across a broad swathe of applications involving complex and heterogeneous data types.

The second challenge is learning in multi-component systems. Computer systems seldom operate in isolation and the outcome of learning tasks on one component may affect a related task on another. For example learning how best to redirect network traffic will once implemented affect the solution that should be adopted at an adjacent node. Cognitive systems composed of multiple agents are another example in which different components may be adapting their behaviour to achieve certain goals, the effects of which will influence the operating environment of other components. The design and analysis of systems involving interacting learning systems is still in its infancy, particularly when we consider theoretical analysis that can be used to guide their design, and if we include self-organisation as a design principle. A related set of challenges arise when we consider integrating information from diverse sources as for example in distributed sensor networks. Once again learning must be used to decide how to filter the data to ensure the network can provide informed responses to a range of different queries. Learning at one node of the network will influence the optimisations at other nodes. The key objective that can enable solutions in all of these applications is to build a well-founded theoretical framework analysing learning in a game theoretic setting. The learning approach can deliver the flexibility, robustness and scalability that are properties required for many applications of cognitive systems, for example in robotics. Such a framework can then provide the criteria that can be used to design and optimise multi-component systems for a wide range of applications.

The third challenge is dealing with data where there is only partial or delayed feedback. This will include semi-supervised learning where large quantities of unlabelled data can be exploited to ensure that only small amounts of expensive labelled data are needed to deliver good performance. The more extreme case of clustering and analysing unlabelled data is still relatively poorly understood theoretically. PASCAL has initiated research into a theoretical foundation of clustering and semantic analysis, but there are still significant gaps in our understanding of how best to quantify applications desiderata and hence translate them into optimisation criteria. Extracting structural knowledge from unlabelled data is undoubtedly one of the keys to enabling learning to be applied in large scale data analysis. The other aspect of this challenge is the situation in which feedback is delayed. Here, the system must make informed guesses that trade-off between the need to perform well (exploit) and obtain useful feedback (explore). Again, the well-founded study of this theme has been initiated in PASCAL, where its importance in applications ranging from web portal design to software testing has been explored. The objective is to develop a framework within which the design of systems that work with partial or delayed feedback can be analysed and optimised.

The fourth challenge is concerned with two of the central themes of traditional cognitive science research: representation (i.e., what representational formats and systems are used to encode knowledge?) and architecture (what are the computational principles with which this information is deployed?). These questions are, of course, tightly entwined---thus, symbolic representations appear most naturally processed by formal, symbol-processing operations; high-dimensional feature-based representations may more readily be processed using extensions of regression, such as connectionist architectures or support vector machines. Moreover, questions of representation and architecture are closely related to problems of learning, the central focus of this project. The learning problem itself is crucially shaped by the choice of representation, and the computational architecture within which learning takes place; moreover, the generalisability and flexibility of knowledge derived from learning crucially depends on representational and architectural issues. Thus, a symbolically represented regularity may, in the appropriate inferential architecture, freely combine with other symbolically represented information, via processes of logical or probabilistic inference. A regularity that is represented "implicitly" in a regression-based architecture may, by contrast, not combine so readily with background knowledge, but may support relatively local generalisations. Both types of process are likely to play an important role in modelling cognition; and understanding how different architectures and representations can interface and interact to support cognitive rich models is a key challenge. Domains in which this work would be conducted would range from explicit reasoning, natural language processing, and the representation of high-level perceptual information.

The issues concerning the scalability and roll-out of adaptive systems technology arise from the recognition that even for tasks such as classification that have been extensively analysed, practical applications frequently require scaling to large scale data, 1000's of classes. Furthermore, for learning to become the standardised technology that we believe it can, and should be, a dependable and robust set of algorithms are required that can handle non-stationary distributions, out-of stream events, sensitivity analysis, and messy real-world data, while giving guarantees for learning machines embedded in larger systems. We do not expect that embedded software enabling all of these features will be delivered by PASCAL2, but these are the long term goals that will guide the strategy and focus of the network's activities.

We now consider the relevance of the proposed research directions. Learning has been identified as a key enabling technology for the engineering and operation of artificial cognitive systems. It is, for instance, key to making robots more versatile, effective and autonomous, and to endowing machines with advanced interaction capabilities. As indicated above learning in cognitive systems will necessarily be multi-component, furthermore it will involve fusion of information from complex data-streams. These features indicate the direct relevance of the first two challenges identified, while operating with partial or delayed feedback is an essential requirement of any adaptive cognitive system. Indeed, we hypothesise that cognition will arise as an emergent property of an interacting network of learning agents able to leverage knowledge from each other, while trading exploration and exploitation in the pursuit of their local goals. Other topics from the work programme for which we aim to develop enabling technologies are sensor networks, where learning in a multi-component system can provide the key to extracting the right mix of useful information for a variety of users; this again can be achieved with the help of trading of exploration and exploitation strategies. Finally, for the area of interfaces, learning can be a vital ingredient for more effective operation of different modes, for their integration, and for adaptation to the users' needs and interests, particularly when combined with associated information access and retrieval, together with cross-lingual tools.

B.1.1.2 General methodology

The leitmotif of the network is to develop principled methods for adaptive systems' design and implementation. By principled we mean that they are supported by an appropriate theoretical model, be it Bayesian, frequentist or statistical. Furthermore, the analysis should extend to the algorithmic complexity and hence be able to gauge the ability of the approach to scale to large-scale applications. The analysis may also require the development of novel models within which to analyse the more complex scenarios outlined above and included in the work programme. The track record of recent developments in machine learning has borne out the contention that well-founded analysis and algorithmic design ensure that new techniques can be applied reliably across a wide variety of domains with very little adaptation.

We maintain that upholding this approach in the more challenging tasks being addressed in this proposal will minimise the risks of adopting ad-hoc and suboptimal solutions.

B.1.1.3 Milestones

Milestones of the project are set at 9 months when the initial adaptation of the current PASCAL website to the new structures (MS2) and to the new consortium arrangements (MS3) will have been completed. There will also be a milestone after each 6 months of the network (MS1, 4-12) at which the Thematic Programme just completed will be reviewed and proposals for new programmes will be considered for adoption in response to the developing needs that have been identified and results that have been obtained in the preceding period. These 6 monthly milestones will also provide a suitable point for approving Harvest and pump-priming projects.

B.1.1.4 Measurable results

The measurable outputs of the network will be:

- The repository of technical reports and publications initiated in PASCAL will continue to expand as an important resource for groups, working in, or applying machine learning technologies. We expect in the region of 500 reports to be produced per year.
- Challenges[1] organized by the network will continue to provide a benchmark for research on particular tasks of relevance to the overall research goals as well as providing a link to application domains.
- A programme of workshops together with video recordings of all main events will provide an unequalled resource for groups studying in this area or attempting to make use of the developed tools.
- Infrastructure and pump-priming projects will develop tools and focused research that will enable novel ideas to be brought to maturity in major research initiatives or in freely downloadable software with a minimum of delay.
- Harvest projects will provide a forum for the development of exemplar applications of direct relevance to industry that can demonstrate the power and range of applicability of the developed approaches.

B.1.2 Long term integration

Principled machine learning research has developed in Europe in a highly distributed way. The EU funded network project PASCAL has succeeded in welding these diverse groups into a remarkably coherent ‘distributed institute’ of researchers, empowered by the web portal that has enabled a virtual centre of excellence to be created.

The advantages of a distributed organisation are the biodiversity of research that is thereby encouraged together with a far greater ability to reach out to different application domains and influence the wide-scale take-up of good practice throughout Europe. PASCAL2 will continue to recreate in the distributed organisation the corresponding advantages of a physically unified institute:

- Easy interchange of questions and ideas among researchers and students
- Opportunity to develop deeper and longer term collaborations
- Workshop, conference and summer school organisation devoted to current challenges and developments
- Exposure of young researchers to different approaches and applications
- Ease of collaboration on papers and proposals
- Pump-priming of novel research directions
- Sharing of resources and creation of a centre for an archive of publications, downloadable software, data sets and technical summary reports
- Opportunity to influence the development of educational programmes through curriculum development, freely downloadable videos of summer schools, etc.
- One-stop location for those seeking assistance with the roll out of the technology into industrial applications

PASCAL2 will recreate each of these features through the establishment of an appropriate *infrastructure* and the launching of a series of *programmes* each tailored to deliver some component of the distributed Institute.

B.1.3 Joint Programme of Activities

B.1.3.1 Overall strategy and general description

As with PASCAL the organisation of the activities will be handled through a series of Programmes that are responsible for the different aspects of the work, termed programme tasks. These programme tasks are grouped into five work packages and are listed in The Network allows for a non-beneficiary membership status that will entitle to participation in Network activities which do not involve payments for salaries. Details are given in section B.2.3 and further elaborated in the Consortium Agreement. In the following we also use the term "site" to denote not only a party to this

[1] Here, the term "Challenge" refers to an element of the JPA, as described in WP3.

grant (beneficiary) but also a non-beneficiary member. Where necessary to make the distinction we shall also use the terms "core site" (beneficiary) and "member site" (non-beneficiary member).

Table 2. The list of programme tasks is based on those that evolved through the PASCAL Network though there are a number of important variations, such as the Harvest and Industrial Club concepts. The work packages (WP1-WP5) will contribute to the nth Year Summary deliverables (D5, 8, 11, 14, 17).

The majority of the Network's research activity will be undertaken under WP3 Joint Research and WP4 Mobility work packages and is organised around **Thematic Programmes**. The first four themes will each last for six months and occupy the first two years of the Network. The remaining themes for years 3 – 5 of the project will be determined by the theme management process as part of WP2. As indicated in section B 1.1.1, they will centre around issues and pursue directions of relevance to the objectives of Challenge 2 of the ICT Work Programme 2007-2008. The Thematic Programmes will each have their own deliverable (D1, 4, 6, 7, 9, 10, 12, 13, 15, 16) in the form of a download of the relevant sections of the project website, with the activity of the relevant tasks such as challenges, pump-priming and Harvest being captured into these deliverables.

The Outreach WP and in particular the Liaison task, will also ensure impact on and cross-fertilisation with parallel activities under Challenge 2 of the ICT Work Programme 2007-2008.

The Network allows for a non-beneficiary membership status that will entitle to participation in Network activities which do not involve payments for salaries. Details are given in section B.2.3 and further elaborated in the Consortium Agreement. In the following we also use the term "site" to denote not only a party to this grant (beneficiary) but also a non-beneficiary member. Where necessary to make the distinction we shall also use the terms "core site" (beneficiary) and "member site" (non-beneficiary member).

Table 2 Breakdown of Programme Tasks by Work Package

WP	Organisational Programme	Programme Tasks
WP1	Operational & Knowledge Management	Operational Management Knowledge Management
WP2	Scientific Integration & Co-ordination	Scientific Management Balance & Integration Site Activity monitoring Theme Management
WP3	Joint Research	Pump-Priming Harvest Challenges Conference & Workshop Organisation
WP4	Mobility	Internal Visiting Conference & Workshop Attendance External Visitor
WP5	Outreach	Curriculum Development Liaison Brokerage of Expertise & Public Outreach Industrial Club, IPR & Exploitation

B.1.3.2 Timing of work packages and their components

All work packages will last for the full 60 month duration of the project.

B.1.3.3 Work package list /overview

Table 3 Work package list

Work package No	Work package title	Type of activity	Lead beneficiary No	Person-months[2]	Start month	End month
WP1	Operational & Knowledge Management	MGT & RTD	1	75	1	60
WP2	Scientific Integration & Co-ordination	RTD	2	35	1	60
WP3	Joint Research	RTD	4	374	1	60
WP4	Mobility	RTD	7	352	1	60
WP5	Outreach	RTD	8	22	1	60
	TOTAL			858		

[2] These figures are indicative of the total research effort on network-related themes by the participating researchers, but they include both "own effort" person months funded by the participants internally or via other R&D projects, and person months funded directly by the network.

B.1.3.4 Deliverables list

Table 4 List of Deliverables – to be submitted for review to EC

Del. no.	Deliverable name	WP no.	Lead beneficiary	Estimated indicative person-months[2]	Nature	Dissemination level	Delivery date (proj. month)
D1	Leveraging complex prior knowledge for learning	WP3,4	UNIMAN	80	O	PU	7
D2	Structures consolidation	WP1	SOTON-ECS	9	R	CO	9
D3	Procedures definition	WP2	UCL	4	R	CO	9
D4	Multi-component learning	WP3,4	UOB	80	O	PU	13
D5	First Year Summary	WP2	UCL	9	R	PU	13
D6	Partial or delayed feedback	WP3,4	UNIMI	80	O	PU	19
D7	Cognitive Architecture and Representation	WP3,4	UCL	80	O	PU	25
D8	Second Year Summary	WP2	UCL	9	R	PU	25
D9	Thematic Programme 5	WP3,4	UCL[3]	80	O	PU	31
D10	Thematic Programme 6	WP3,4	UCL[3]	80	O	PU	37
D11	Third Year Summary	WP2	UCL	9	R	PU	37
D12	Thematic Programme 7	WP3,4	UCL[3]	80	O	PU	43
D13	Thematic Programme 8	WP3,4	UCL[3]	80	O	PU	49
D14	Fourth Year Summary	WP2	UCL	9	R	PU	49
D15	Thematic Programme 9	WP3,4	UCL[3]	80	O	PU	55
D16	Thematic Programme 10	WP3,4	UCL[3]	80	O	PU	60
D17	Fifth Year Summary	WP2	UCL	9	R	PU	60
			TOTAL	**858**			

[3] The lead beneficiary of later thematic programmes will be determined through a call for proposals, and at this time we simply enter UCL as a default.

B.1.3.5 Work package descriptions

Work package number	WP1	**Start date or starting event:**		Month 1		
Work package title	Operational & Knowledge Management					
Activity type	MGT & RTD					
Participant number	1	2	6			
Participant short name	*SOTON-ECS*	*UCL*	*JSI*			
Person-months per participant[2]	56	14	5			

Objectives

- To oversee the smooth running of the entire Network, by assisting in the implementation of Programmes and running of the Steering Committee.
- To maintain the established web portal system.
- To develop the portal system to incorporate new and changing aspects of the Network.
- Ensure smooth operation of the Network programmes, and dissemination of research.

Description of work

Task One: Operational Management (MGT)

a) Management of the Network Project Office in Southampton. (*Soton-ECS*)

b) Overall responsibility for the operation and upgrading of the PASCAL web portal. (*Soton-ECS*)

c) Responsible for operational management of the site activity review process. (*Soton-ECS, UCL*)

d) Responsible for monitoring of expenditure on budgets. (*Soton-ECS*)

e) Support for Workshop or Challenge organisation. (*Soton-ECS, UCL*)

f) Preparing and updating consortium agreements. (*Soton-ECS*)

g) Collating deliverables. (*Soton-ECS, UCL*)

Task Two: Knowledge Management (RTD)

a) Re-alignment of existing organisational portal with new structures: Programme management, Site Management, Mailing Lists, Network View, My PASCAL, WiKi, Document Store & FAQ (*Soton-ECS, UCL*)

b) Enhancement and maintenance of dissemination activities: Publications Database (*Soton-ECS*), Video Lectures (*JSI*),

c) Enhanced Workshop and Challenge Hosting (*Soton-ECS, UCL*)

d) Extension of GForge system to Dataset repository to support challenges and additional collaborative services to support joint authorship of papers and teaching materials. (*Soton-ECS*)

Deliverables (brief description) and month of delivery

D2, Structures Consolidation, M9

D5, 1st Year Summary, M13

D8, 2nd Year Summary, M25

D11, 3rd Year Summary, M37

D14, 4th Year Summary, M49

D17, 5th Year Summary, M60

Work package number	WP2	**Start date or starting event:**		Month 1		
Work package title	Scientific Integration & Co-ordination					
Activity type	RTD					
Participant number	2	1	6	4	3	
Participant short name	*UCL*	*SOTON-ECS*	*JSI*	*CNRS*	*UEDIN*	
Person-months per participant[2]	27	3	2	2	1	

Objectives

- To oversee the scientific progress of the project
- To identify through appropriate calls thematic programmes for years 3 – 5 of the project
- To monitor the overall balance of network actions between theoretical and applied activities and among different application domains.
- To monitor progress towards overall network goals such as integration, outreach, excellence and equal opportunities including gender balance.
- To take (or encourage other programme tasks to take) any actions needed to improve progress towards these goals.

Description of work

Task One: Scientific Management

a) Convening and chairing steering committee meetings (*UCL*)

b) Collating inputs from the Industrial Pool, Balance and Integration Programme Task, etc., concerning the opportunities and directions that the Network should adopt (*UCL*)

c) Monitoring the operation of all aspects of the Network to ensure smooth running and identify new opportunities (*UCL, Soton-ECS*)

Task Two: Balance & Integration

a) To provide input where appropriate to Site Activity, pump-priming and harvest programme tasks to ensure maximum responsiveness to the global objectives (*UCL, Soton-ECS*)

b) To develop mentoring/outreach programmes and resources designed to encourage equal participation. (*JSI*)

c) Monitoring the project to ensure resources are used in a way that will promote the overall goals of the network: integrating theoretical and applied research, building a vibrant and sustainable European research community, promotion of the results in education, wider dissemination and practical exploitation, promotion of diversity and equal opportunities. (*CNRS*)

d) Direct intervention when necessary to perform corrective actions. (*JSI*)

Task Three: Site Management & Activity

a) To develop procedures and guidelines for membership enlargement. (*UCL, Soton-ECS, CNRS*)

b) There is an annual review of each site's level of network-related activity, in order to verify that funding was well spent and set funding levels for the next year. The review process includes setting appropriate activity classes and weightings, data collection and annual reports, site scoring and setting the corresponding funding level, and running an appeals process. (*CNRS, UEDIN*)

Task Four: Theme Management

a) To issue calls for proposals of new Thematic Programmes on the most pressing themes for the overall scientific development and impact of the network at the end of years 2, 3 and 4 (forming the basis for the associated milestones, MS7-12). (*UCL*)

b) To subject proposals to peer review and then decide by Steering Committee as to which should be adopted for the next period of the network's activities. (*UCL*)

Deliverables (brief description) and month of delivery
D1, Leveraging complex prior knowledge for learning, M7
D3, Procedures definition, M9
D4, Multi-component learning, M13
D5, 1st Year Summary, M13
D6, Partial or delayed feedback, M19
D7, Representation/Architecture, M25
D8, 2nd Year Summary, M25
D9, Thematic Programme 5, M31
D10, Thematic Programme 6, M37
D11, 3rd Year Summary, M37
D12, Thematic Programme 7, M43
D13, Thematic Programme 8, M49
D14, 4th Year Summary, M49
D15, Thematic Programme 9, M55
D16, Thematic Programme 10, M60
D17, 5th Year Summary, M60

Work package number	WP3	**Start date or starting event:**		Month 1			
Work package title	Joint Research						
Activity type	RTD						
Participant number	4	5	17	9	16	13	14	1-17
Participant short name	*CNRS*	*XEROX SAS*	*TKK*	*UNIMAN*	*BIU*	*FRAUNHOFER*	*CWI*	*NETWORK*
Person-months per participant[2]	8	12	5	4	2	2	1	340

Objectives

- To inject targeted resources that could not be delivered at a local level to further the leading edge of European research in machine learning.
- To cultivate the transfer of research results to industry and users in general.
- To concentrate the efforts of the community on potentially ground-breaking areas.
- To incubate new areas of research around applications, theoretical questions or exploratory challenges.
- To provide an extensive assessment/benchmarking of algorithms.
- Develop competition challenges to support the goals of the Harvest programme.
- To promote interaction between researchers and research fields by supporting conferences and workshops.
- To promote dissemination of research results

Description of work

Task One: Infrastructure/Pump-priming

a) Develop procedures and guidelines for funding applications, calls, review process and reporting, based on the model developed in PASCAL (*UniMan, Fraunhofer*)

b) Manage the review process and report the results to the steering committee (*UniMan*)

c) Monitoring the resulting contribution to the Network (*Fraunhofer, UniMan*)

d) Undertaking of pump-priming activity. (*NETWORK*)

Task Two: Harvest

This programme will fund the preparation and the execution of joint projects between academic partners and industrial members or scientists from research communities outside the scope of PASCAL 2. Harvest projects will be of short duration (e.g. three months) and will be carried out by teams of around 6 members who will gather to work in the same physical location.

a) To design and deploy a web-based infrastructure for publishing project proposals, managing the bidding, review and publishing process (with WP1). (*XEROX SAS*)

b) To define the rules, reviewers, selection criteria and the calendar for proposal publication and bidding. It is anticipated that these criteria will favour projects requiring inter-disciplinary and inter-site collaborations. (*XEROX SAS, CNRS*)

c) To prepare emplate agreements identifying alternative ways in which IPR will be dealt with in each project; these templates will serve as a basis for defining the actual rules on a project-by-project basis. (*XEROX SAS, CNRS*)

d) To monitor the projects with a lightweight process. The process, via documentation, will seek approval from the Steering Committee on minimal software requirements that Harvest projects should satisfy. (*XEROX SAS*)

e) To organise Harvest workshops to showcase the results from the Harvest projects (*XEROX SAS, CNRS*)

f) Undertaking of harvest activity. (*NETWORK*)

Task Three: Challenges

a) To extend the procedures and guidelines for funding calls, and reporting, based on the model developed in PASCAL (*CNRS, BIU*)

b) Manage the review process and report the results to the steering committee (*CNRS,BIU*)

c) Monitor the balance of challenge types and contribution to the Network (*BIU,CNRS*)

d) Undertaking of challenge activity. (*NETWORK*)

Task Four: Conference & Workshop Organisation

a) Develop procedures and guidelines for funding applications, calls, and reporting, based on the model developed in PASCAL (*TKK, CWI*)

b) Manage the review process and report the results to the steering committee (*TKK*)

c) Monitoring the quality and topic distribution of workshops (*CWI,TKK*)

d) Undertaking of conference and workshop organisation activity. (*NETWORK*)

Deliverables (brief description) and month of delivery

D5, 1st Year Summary, M13

D8, 2nd Year Summary, M25

D11, 3rd Year Summary, M37

D14, 4th Year Summary, M49

D17, 5th Year Summary, M60

Work package number	WP4	**Start date or starting event:**		Month 1		
Work package title	Mobility					
Activity type	RTD					
Participant number	7	2	10	14	1-17	
Participant short name	*UNIMI*	*UCL*	*UH*	*MPG*	*NETWORK*	
Person-months per participant[2]	4	10	3	2	340	

Objectives

- To strengthen integration within the Network.
- To assist in the organisation and funding of visits between sites of the Network.
- Support attendance of members at relevant international conferences or workshops.
- Support attendance of student members at international schools or similar tutorial events in relevant research areas.
- To bring external experts into the Network to complement existing expertise.

Description of work

Task One: Internal Visiting

a) Develop procedures and guidelines for funding applications, calls, review process and reporting, based on the model developed in PASCAL (*UCL*)

b) Manage the review process and report the results to the steering committee (*UCL*)

c) Monitoring the resulting contribution to the Network (*UCL*)

d) Undertaking of visiting activity. (*NETWORK*)

Task Two: Conference & Workshop Attendance

a) Develop procedures and guidelines for funding applications, calls, review process and reporting, based on the model developed in PASCAL (*UH,MPG*)

b) Manage the review process and report the results to the steering committee (*UH*)

c) Monitoring the resulting contribution to the Network (*MPG,UH*)

d) Undertaking of conference attendance activity. (*NETWORK*)

Task Three: External Visitor

a) Develop procedures and guidelines for funding applications, calls, review process and reporting, based on the model developed in PASCAL (*UNIMI,UCL*)

b) Manage the review process and report the results to the steering committee (*UNIMI*)

c) Monitoring the resulting contribution to the Network (*UCL,UNIMI*)

Deliverables (brief description) and month of delivery

D5, 1st Year Summary, M13

D8, 2nd Year Summary, M25

D11, 3rd Year Summary, M37

D14, 4th Year Summary, M49

D17, 5th Year Summary, M60

Work package number	WP5	**Start date or starting event:**		Month 1		
Work package title	Outreach					
Activity type	RTD					
Participant number	8	4	6	5	1	2
Participant short name	*UoB*	*CNRS*	*JSI*	*XEROX SAS*	*SOTON-ECS*	*UCL*
Person-months per participant[2]	5	7	6	2	1	1

Objectives

- To promote, discuss and propose courses, curricula, summer schools and bootcamps related to machine learning.
- To contribute through courses of different levels to pedagogical archives, in formats ranging from print (slides, papers, e-prints) to downloadable videos
- To bring to the Network the challenges and problems identified in related projects in robotics, cognitive systems and neuroscience that are of particular interest and have potential to be enhanced with machine learning techniques.
- To provide a dissemination route for techniques and algorithms developed within PASCAL to ensure that where appropriate researchers in related projects are aware of the developments and their potential.
- To ensure the effective communication of results to different audiences: potential industrial users of the technology (technology transfer); the general public (public outreach programme); the wider scientific community.
- To forge a bidirectional interface between academia and industry to increase the application of machine learning techniques.

Description of work

Task One: Curriculum Development

a) To develop the criteria and process for obtaining a "PASCAL certified PhD" (*CNRS*)

b) To propose courses that prepare students for research work and industrial applications in machine learning. The PASCAL bootcamp is one such initiative, but others, with in mind the possibility of an Erasmus Mundus master project, will be encouraged. (*UCL, CNRS*)

c) A third task consists in organising the information about teaching machine learning through the PASCAL website. An easy access to the different available documents is needed: Books and manuscripts, web pages, packages, etc. (*CNRS*)

Task Two: Liaison

a) Keep abreast of developments within other projects in order to ensure timely links are developed and opportunities for collaboration are not missed. (*CNRS, JSI*)

b) Disseminate to other PASCAL2 programmes relevant results, ideas and data from sister projects. (*CNRS*)

Task Three: Brokerage of Expertise & Public Outreach

a) Create and maintain a list of industrial and scientific partners. (*UoB, UCL*)

b) Develop case studies, demonstrations, proofs of concept, as part of every research undertaking through the standard call model. (*UoB*)

c) Create a dedicated web-site to facilitate the communication between firms and the experts in the network, where specific requests, suggestions, offers can be posted, and people with the relevant expertise can be put in touch with potential users in the EU. (*UoB*)

d) Organise events to bring together the community of technology users with those researchers who have developed tools, proofs of concept and demonstrations of feasibility. These events will also be used for public outreach purposes. (*UoB*, *UCL*)

Task Four: Industrial Club, IPR & Exploitation

a) To implement proactive coordination of the Harvest Programme in a way that industrial partners will be individually contacted with concrete proposals for collaboration using PASCAL2 funds. (*JSI, XEROX SAS*)

b) Special attention will be given to the standardisation of the results of the Harvest programme – this will ensure

(in combination with other software development activities within the network) an increased level of connectivity between various research project results. (*JSI, XEROX SAS*)

c) To help collaboration between industrial and academic partners by promoting internships and other forms of exchange of students or senior staff in both directions. (*JSI*)

d) Evaluation of the "Industrial Club" programme will be mainly through the number of successful demonstrators and software prototypes that arise as the result of joint projects. Another form of evaluation will be the number of follow-up collaborations in the form of PASCAL2 independent contracts or further joint EU projects arising from such collaborations. (*JSI*)

Deliverables (brief description) and month of delivery

D5, 1st Year Summary, M13

D8, 2nd Year Summary, M25

D11, 3rd Year Summary, M37

D14, 4th Year Summary, M49

D17, 5th Year Summary, M60

B.1.3.6 Efforts for the full duration of the project

Workpackage	WP1	WP2	WP3	WP4	WP5	TOTAL[2] per Beneficiary
SOTON-ECS	56	3	10	10	1	**80**
UCL	14	27	150	150	1	**342**[4]
UEDIN	0	1	10	13	0	**24**
CNRS	0	2	48	40	7	**97**
XEROX SAS	0	0	22	10	2	**34**
JSI	5	2	10	10	6	**33**
UNIMI	0	0	10	14	0	**24**
UOB	0	0	10	10	5	**25**
UNIMAN	0	0	14	10	0	**24**
UH	0	0	10	13	0	**23**
IDIAP	0	0	10	10	0	**20**
CWI	0	0	11	10	0	**21**
FRAUNHOFER	0	0	12	10	0	**22**
MPG	0	0	10	12	0	**22**
TKK	0	0	15	10	0	**25**
BIU	0	0	12	10	0	**22**
UPMC	0	0	10	10	0	**20**
TOTAL	**75**	**35**	**374**	**352**	**22**	**858**

[4] The large amount of effort associated with UCL is explained in section B.2.1.4.

Activity Type	SOTON-ECS	UCL	UEDIN	CNRS	XEROX SAS	JSI	UNIMI	UOB	UNIMAN	UH	IDIAP	CWI	FRAUNHOFER	MPG	TKK	BIU	UPMC	TOTAL[2] ACTIVITIES
RTD/Innovation activities																		
Operational & Knowledge Management	8	14				5												**27**
Scientific Integration & Co-ordination	3	27	1	2		2												**35**
Joint Research	10	150	10	48	22	10	10	10	14	10	10	11	12	10	15	12	10	**374**
Mobility	10	150	13	40	10	10	14	10	10	13	10	10	10	12	10	10	10	**352**
Outreach	1	1		7	2	6		5										**22**
Total 'research'	**32**	**342**	**24**	**97**	**34**	**33**	**24**	**25**	**24**	**23**	**20**	**21**	**22**	**22**	**25**	**22**	**20**	**810**
Demonstration activities																		
Total 'demonstration'	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Consortium management activities																		
Operational & Knowledge Management	48																	**48**
Scientific Integration & Co-ordination																		**0**
Joint Research																		**0**
Mobility																		**0**
Outreach																		**0**
Total 'management'	**48**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**48**
Other activities																		
Total 'other'	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
TOTAL BENEFICIARIES	**80**	**342**[4]	**24**	**97**	**34**	**33**	**24**	**25**	**24**	**23**	**20**	**21**	**22**	**22**	**25**	**22**	**20**	**858**

B.1.3.7 List of milestones and planning of reviews

List and schedule of milestones

Milestone no.	Milestone name	WPs no's.	Lead beneficiary responsible for the milestone	Delivery date from Annex I	*Comments* *Success criteria / expected achievements and Means of verification*	
MS2	Structures	WP1	SOTON-ECS	Month 9	Efficient web implementation	Verified by D2
MS3	Procedures	WP2	UCL	Month 9	Equitable and efficient procedures for Network funding	Verified by D3
MS1	Themes A	WP2	UNIMAN	Month 6	Workshop. Challenge, Pump-Priming, Harvest, Publication, Lecture, … Activity	Verified by D1
MS4	Themes B	WP2	UOB	Month 12		Verified by D4
MS5	Themes C	WP2	UNIMI	Month 18		Verified by D6
MS6	Themes D	WP2	UCL	Month 24		Verified by D7
MS7	Themes E	WP2	UCL	Month 30		Verified by D9
MS8	Themes F	WP2	UCL	Month 36		Verified by D10
MS9	Themes G	WP2	UCL	Month 42		Verified by D12
MS10	Themes H	WP2	UCL	Month 48		Verified by D13
MS11	Themes I	WP2	UCL	Month 54		Verified by D15
MS12	Themes J	WP2	UCL	Month 60		Verified by D16

Tentative schedule of project reviews

Review no.	Tentative timing, i.e. after month X = end of a reporting period
1	After project month: 12
2	After project month: 24
3	After project month: 36
4	After project month: 48
5	After project month: 60

B2. Implementation

B.2.1 Management structure and procedures

B.2.1.1 Management Principles

Management will be driven by a number of over-riding principles:

(i) *Equity*. To approach the ideal of the network as a harmonious community of scientists working efficiently towards a common vision, it is important that the network conduct itself in an equitable and ethical manner.

(ii) *Efficiency*. To promote efficient day-to-day management of the whole network. That is, efficient running of the individual programme tasks and their co-ordination across the network.

(iii) *Transparency.* To provide a clear dissemination of the procedures used and the decisions reached by the Network.

(iv) *Scientific Excellence.* To promote, champion and disseminate research of the highest scientific standards.

(v) *Risk Minimisation*. Under the umbrella of project management, principles of quality and risk management will be used to meet deliverables at cost, on time and with high quality. Risk management involves reduction or elimination of events which upset these aims as well as preparation of contingencies.

B.2.1.2 Management Structure

The above management principles will be achieved through five principle components:

(i) *Strategic Advisory Council*. The Council will comprise senior outside scientific figures (Sue Becker, Mike Mozer, Helge Ritter, Satinder Singh and Sebastian Thrun), and may also include representatives of national and international funding agencies and of our industrial club. There are two major duties:
   (a) To provide input to the debate concerning the overall strategic direction and vision
   (b) To monitor the equity and ethical stance of the network including an appeal procedure in the event of disputes.

(ii) *Steering Committee*. This is the core committee with overall responsibility for the network. It will comprise a Chair, the Scientific Coordinator, the Operational Coordinator, the managers of each of the programmes (deputised by their co-managers), the gender issues representative, the Financial Officer and the Executive Officer (see below). It will have the following overall responsibilities:
   (a) The main scientific programme
   (b) Internal bidding system (see below)
   (c) Overseeing the management of individual programmes and the co-ordination between programmes
   (d) Liaison with other related projects and networks
   (e) Overseeing the evolution of the Network

(iii) *Executive Committee*. This will be chaired by the Executive Officer (based at University College, London), include the Scientific and Operational Coordinators, the Financial Officer (based at the University of Southampton) and the Web Manager (University of Southampton). It will have day-to-day responsibility for all administrative matters including:
   (a) Budgets and current expenditure on the different budgets
   (b) Quality and risk management systems
   (c) Project management and project management systems
   (d) Report management and reporting standards

(iv) *Individual Programme Committees*. Each programme will have a Programme Chair and may create its own committee to suit administrative convenience. Thus, for example, the Chair may appoint other scientists, local administrators, etc. to form the committee.

(v) *Assembly*. This comprises one representative from each beneficiary and has formal responsibility for accession of any amendments to the Technical Annex (this document).

B.2.1.3 Management Process

The individual programme managers will be members of and report to the Steering Committee, who will set the strategic direction for the Network taking into account advice from the Strategic Advisory Council and input from

researchers at contractor and associate sites. The Executive Committee will report regularly to the Steering Committee and may deal individually with the programme managers on matters of budget and detailed administration.

The Steering Committee will meet biannually. Business arising from the different programmes will be considered, including the ratification of results of any calls for proposals having a significant financial or scientific impact on the network. The submissions will have already been refereed and prioritised by the relevant Programme Manager.

The Scientific Management Programme will bring strategic considerations to the attention of the Steering Committee for discussion taking into account input from the Balance and Integration Programme among others, and including whenever possible views from the Strategic Advisory Council. Proposals for future Thematic Programmes will be considered and any decisions on these or other strategic decisions will be passed to the relevant Programme Managers for them to take appropriate actions.

The Executive Committee functions around the secretariat that will be established at the Scientific Coordinator's site. The Committee responds to the day-to-day requests for information, collating relevant information for the programme managers, preparing documentation for the Steering Committee and Strategic Advisory Council meetings. Requests for financial reports and knowledge management issues will be passed to the Financial Officer and Web Manager respectively. The Financial Officer will be responsible for keeping track of all expenditure in order to monitor the budgets of the individual programmes preparing reports for Steering Committee meetings to ensure that as up to date a picture of the expenditure is available to the Programme Managers.

When urgent matters arise between meetings these will when possible be dealt with by the Executive Committee. If there is a need to consult the Steering Committee this will be made through email exchanges. There will be clear document trails established for all the programmes through the use of web based procedures that ensure all information is automatically logged in relevant databases. These will include calls for submissions, the applications for funding, referee reports, minutes of meetings considering different proposals, proposed rankings and final decisions either taken within the programme or by the Steering Committee. Reports of the results of the funding will also be filed including a record of funds actually used.

The Strategic Advisory Council will meet occasionally to provide an opportunity to reflect on progress within the project. Members of the council will be given any deliverables submitted during the previous period together with any further documentation that they choose to request from individual programmes. Meetings which may be distributed either through video conferencing or emailing will usually be held shortly before a Steering Committee meeting to ensure that any issues highlighted by the Strategic Advisory Council can be rapidly acted on by the Steering Committee.

B.2.1.4 Budgeting

Following the successful implementation of the strategy in PASCAL, PASCAL2 will adopt a dual band funding scheme. Note that the allocation of funding for non-beneficiary members will follow the same procedure as for beneficiaries, so that in this section 'sites' refers to all member sites. The only difference for non-beneficiary members is that any allocated funds will be held by UCL, who will distribute them against that site's expenditure as described in B.2.4. This explains UCL's relatively large budget allocation and project effort.

Approximately 33% of the funds will be allocated through direct funding to the sites through the Site Activity Programme. The remaining 67% will be allocated to the other programme managers for distribution through requests submitted to their programmes. We describe these two components separately:

(i) *Site Activity Programme* The Site Activity Programme aims to provide a base level of funding to support the individual partners involvement in the network. This ensures that requests for additional funding can be restricted to larger scale activities, such as workshop or challenge organisation, etc. The funding allocation to each site is set on a yearly basis in response to their level of contribution to the network during the previous year so that the funding also provides a positive feedback mechanism to encourage participation in the network. The method of measuring the partners' activity is outlined above in the description of the workings of the Site Activity Programme. For the first year the level of funding will be set using the evaluation from the final year of PASCAL, with adjustments for any changes in the constitution of sites or new sites.

(ii) *Remaining Programme Funding* The remaining 67% of the funds will be allocated to budgets for each of the programmes with a reserve fund held centrally to assist in any special needs arising during the year. A few clarifications are in order:

(a) The percentage allocated to the Infrastructure/Pump-priming Programme is high because this represents a significant investment in the development of key components and testing of novel directions for the network.

(b) The percentage allocated to the Conference and Workshop Organisation Programme is high as this provides a key vehicle for fostering new research directions and collaborations across different themes and between different groups.

(c) The percentage allocated to the Challenge Programme is high as these provide an important way establishing active two-way communication between algorithms and analysis research groups and those working in relevant application areas.

The budget allocations to the different programmes will be decided by the Steering Committee at the beginning of the year though allocations for the first year in percentages are as follows:

Table 5 Indicative Budget breakdown by Programme Task

WP	Programme	Budget (%)
WP1	Knowledge Management	6.0
	Operational Management	9.5
WP2	Scientific Management	4.5
	Balance & Integration	0.5
	Site Activity	0.5
	Theme Management	0.5
WP3	Infrastructure/pump-priming	12.0
	Harvest	5.0
	Challenges	6.0
	Conference & Workshop Organisation	11.0
WP4	Internal Visiting	4.0
	Conference & Workshop Attendance	2.0
	External Visitor	1.0
WP5	Curriculum Development	0.5
	Liaison Programme	0.5
	Brokerage of Expertise & Public Outreach	1.0
	Industrial Club, IPR & Exploitation	0.5
	Site Activity Funding	33.0
	Reserve	2.0
	Total	100.0

For site activity funding of and additional funding requests from non-beneficiary members, approval will be sought from the responsible European Commission department / officer. The absence of a response from the Commission within 45 days of receipt of such a request constitutes approval. The approval request specifies the member involved, the Network activity / programme, the goal(s) of the specific action, as well as the budget requested or to be granted. Approval requests for site activity funding will normally be bundled and submitted en-bloc at regular intervals (e.g., subsequent to Steering Committee decisions). Requests through other individual PASCAL2 programmes will be submitted to the Project Officer through the web portal after his being directly alerted once the relevant Programme Manager has approved the request. In addition to email confirmations of individual approvals, a summary of all approved requests will be submitted at regular intervals for recording purposes.

B.2.1.5 Internal bidding

The Programme Manager will be responsible for instigating and fielding calls for funds under their programme, balancing the needs of individual requests against the overall budget allocation for that programme in their funding recommendations. In the event of the needs outstripping the budget, he or she may request additional funds from the central reserve fund. Such funding will be agreed through an email exchange between all members of the steering committee if urgent or discussed at the next six monthly meeting.

Any individual member of the network (beneficiary and non-beneficiary site members) may bid for additional funding via the appropriate programme, either in response to a call, or if a permanent call is open pro-actively. Bids will be decided upon according to a three-stage process: (i) refereeing by the manager either alone or with one or more additional members depending on the scale of funding requested, (ii) recommendation by the programme manager, and in the event of larger amounts of funding, (iii) final decision by the full Steering Committee which will be required to take into account the strategic directions of the network in making its decisions. In all cases a complete document trail will be maintained through the website for monitoring of procedures. Codes of practice will be drawn up for the decision process covering ethical standards, disclosure and appeal procedures.

B.2.2 Beneficiaries

B.2.2.1 University of Southampton

University of Southampton is an internationally-recognised research-led university with strong interdisciplinarity. It has a worldwide reputation for innovative research which was confirmed in the recent RAE, where every department in the School of Engineering, and Computer Science gained 5*. The Information: Signals, Images and Systems Research Group is situated within the School of Electronics and Computer Science and is currently composed of approximately 70 researchers. ISIS activities are centred in fundamental theory and algorithm development associated with adaptive data modelling, machine learning, control theory, computer vision and signal processing. The research group is recognised as one of the leading machine learning groups in the world with expertise in theory (e.g. frequentist and Bayesian approaches to learning), algorithms (e.g. kernel methods, SVMs, subspace methods) and their application to many diverse domains. This research is developed through a combination of theoretical modelling together with verification and validation in the real-world problem domains of vision and image processing, signal processing, document analysis, guidance and control, biomedical fields, etc. Research is also focused on the development of methodologies for integrating such techniques in overall systems engineering.

Steve Gunn is a Professor in the Information: Signals, Images and Systems Research Group, and has published over 80 research papers in the areas of image processing and machine learning. His current research interests are in the area of sparse representations, feature selection and subspace methods for identification of salient parts of the data space for prediction. He has recently published a book entitled *Feature Extraction: Foundations and Applications*.

Adam Prügel-Bennett is a Reader in the Information: Signals, Images and Systems Research Group. His main research interests are in mathematical modelling applied to machine learning, including models of the brain, Bayesian learning, computer vision, bioinformatics and optimisation techniques. He is a member of PASCAL and has been a guest speaker at a PASCAL workshop.

Craig Saunders main research interests currently lie in the application of Machine Learning methods to structured data. He has developed various kernel methods for structured data and has shown that they can be successfully applied in the domains of text, biological sequences and molecular data. Recently he has been involved in the development of algorithms which encompass structured output data, which have been shown to achieve good performance on hierarchical data (such as classifying documents into a category hierarchy, and enzyme function classification). This work is currently being extended to the fields of enzyme function prediction and Statistical Machine Translation. His work is currently funded by the EU as part of the SMART project (FP6--033917) and industrially through GlaxoSmithKline. He has been funded and worked on many previous EU and UK EPSRC projects, and is currently and active participant in the PASCAL network.

B.2.2.2 University College London

University College has recently created a Centre for Computational Statistics and Machine Learning (CSML) linking the Departments of Computer Science and Statistical Sciences with the Gatsby Computational Neuroscience Unit. This combination and breadth of expertise makes the group a valuable contributor to the PASCAL2 network. Furthermore the Scientific Coordinator of PASCAL, John Shawe-Taylor, has been appointed to the Directorship of this Centre, equipping UCL to undertake a more prominent role in the management of PASCAL2.
The research represented in the Centre ranges from pure machine learning covering both frequentist and Bayesian approaches and a strong experience with kernel methods, through traditional statistics to one of the leading centres for the application of learning methods to the analysis of computational neuroscience. This link is particularly relevant for the new focus of PASCAL2 on Cognitive Science. Furthermore, the Centre includes additional researchers from related areas such as for example Nick Chater a recognised expert in Cognitive Science from the Department of Psychology at UCL. The Centre for CSML at UCL is an example of the restructuring of research in the area of machine learning that has resulted from the PASCAL Network's activities. Its mix of expertise is particularly well-suited for its position as one of the PASCAL2 main contractors.

John Shawe-Taylor is the original Coordinator (since his move to UCL the Scientific Coordinator) of PASCAL. He has further coordinated two earlier Networks (Working Groups), known as NeuroCOLT and NeuroCOLT2. These projects laid much of the groundwork for the PASCAL network. He also coordinated the STREP project Kernel Methods for Images and Text (KerMIT) involving two commercial and four university partners. In his research he has pioneered the development and application of principled machine learning methods with his contributions to the theoretical underpinnings of techniques such as kernel and boosting methods.

Yee Whye Teh joined the Gatsby Computational Neuroscience Unit as a lecturer in 2007. Prior to this appointment he obtained his PhD from the University of Toronto and worked as a postdoc at the University of California at Berkeley and as Lee Kuan Yew Postdoctoral Fellow at National University of Singapore.

Massimiliano Pontil is an EPSRC Advanced Research Fellow and Reader in the Department of Computer Science at University College London (UCL). He joined UCL as a lecturer in 2003. He received the equivalent of an MSc and a PhD in Physics from the University of Genova in 1994 and 1999, respectively.

Peter Dayan is Director of the Gatsby Computational Neuroscience Unit. His research interests lie at the intersection between methods and techniques of machine learning, and theoretical neuroscience.

Nick Chater is Professor of Cognitive and Decision Sciences at the Department of Psychology at UCL, and specializes on computational models of learning and decision making. He has especially focussed computational methods for learning syntactic categories and structures in language acquisition; on deriving learnability results for language acquisition, using Kolmogorov complexity theory; and building Bayesian probabilistic models of many aspects of human reasoning.

Maneesh Sahani is a Lecturer in Computational Neuroscience at the Gatsby Computational Neuroscience Unit, and a member of the Centre for CSML. He received his PhD in 1999 from the California Institute of Technology (Caltech) under the supervision of Profs. John Hopfield and Richard Andersen. After postdoctoral work at the Gatsby Unit and at the University of California at San Francisco, he joined the UCL faculty in 2004.

B.2.2.3 University of Edinburgh

The researchers are members of the Institute for Adaptive and Neural Computation (IANC) and the Institute for Perception, Action and Behaviour (IPAB) withing the School of Informatics, University of Edinburgh.

Informatics is one of seven schools in the College of Science and Engineering, at The University of Edinburgh. It was the only department in the UK awarded the top 5*A rating in Computer Science in the 2001 Research Assessment Exercise. With 87.1 research active staff submitted for assessment, we are also the UK's biggest research group in this area. We received a top excellent rating in the most recent (1994) SHEFC Teaching Quality Assessment exercise.

Prof Chris Williams obtained his PhD from the University of Toronto in 1994, under the supervision of Prof Geoffrey Hinton. Subsequently he was first a research fellow and then Lecturer at Aston University. He moved to Edinburgh in 1998 and was subsequently promoted to Reader (2000) and Professor of Machine Learning (2005). He is also Director of IANC. Prof Williams is interested in a wide range of theoretical and practical issues in machine learning, statistical pattern recognition, probabilistic graphical models and computer vision. He has particular interests in visual object recognition, time series analysis (e.g. intensive-care condition monitoring), Bayesian methods, and sensor networks.

Dr Sethu Vijayakumar obtained his PhD in 1998 from the Tokyo Institute of Technology and then held positions at RIKEN Brain Science Institute (1998-2000) and the University of Southern California (2001-2003). He joined the School of Informatics in 2003 and is currently a Reader, as well as Director of IPAB. In 2007 he was awarded a prestigious 5-year Royal Academy of Engineering Senior Research Fellowship. His research interests include statistical machine learning, visual attention & oculomotor control, and real-time learning for robot control. These methods address problems of adaptive, autonomous control in robotic cognitive systems.

Dr Amos Storkey obtained his PhD in 2000 from Imperial College London. He then worked as a research fellow in the School of Informatics, University of Edinburgh from 2001 to 2005, and was appointed a Lecturer in the same school in 2005. His areas of research interest include machine learning and probabilistic modelling, image understanding, time series analysis, and methods in medical imaging.

The Edinburgh team will also include post-docs and PhD students who are working on relevant topics in IANC and IPAB.

B.2.2.4 Centre National de Recherche Scientifique

Founded in 1939 and one of the largest fundamental research organizations in Europe, CNRS is present in all fields of knowledge, ranging from Mathematics and Information Sciences to Humanities. Most of its 11,500 researchers work in joint labs, partnered with universities in all regions. There are a number of laboratories that will participate via the CNRS partner and the description is extended to reflect this:

LJK-CNRS – Laboratoire Jean Kuntzmann (UMR 5224), Grenoble

Laboratoire Jean Kuntzmann (LJK) is an Applied Mathematics and Computer Science laboratory created in January 2007 in Grenoble, France. LJK is a Joint Research Unit/Unité Mixte de Recherche (UMR) set up by the following entities: the Université Joseph Fourier (- Grenoble 1), the CNRS, the Université Pierre-Mendez France (- Grenoble 2), and the Institut National Polytechnique de Grenoble. It combines the forces of applied mathematicians and statisticians from the former laboratories LMC and LabSAD with graphics and computer vision experts from the former laboratory GRAVIR. Its expertise centres on the computational and statistical sciences and their uses in analysing natural phenomena, with applications ranging from environmental modelling through life sciences, nanosciences, visualisation and signal processing to mathematical finance.

Four LJK research teams will participate in PASCAL2: AI (Apprentissage et Interfaces), whose focus is on learning based approaches to treating large-scale and sensed data; **LEAR** (Learning for Recognition), whose focus is computer vision and particularly visual recognition; **MISTIS**, whose focus is statistical modelling of images and signals, with a particular emphasis on Markov field and mixture models; **SMS** (Statistics and Stochastic Modelling), whose interests cover several areas of statistical modelling and machine learning including functional estimation, statistical analysis of stochastic processes, stochastic algorithms, models and inference for reliability and maintenance, and various applications of these.

The participants from the LJK site include:

Anestis Antoniadis, who is Professor of Statistics in the Université Joseph Fourier, director of the Statistics department of the Laboratoire Jean Kuntzmann, and fellow of the American Statistical Association and of the Institute of Mathematical Statistics. His research interests include: machine learning and ICA models; inverse problems; functional data analysis; nonparametric and semi-parametric regression (single index models, dimension reduction, wavelet denoising and regularization); stochastic sensitivity analysis; and various applications (analysis of microarray and proteomic data, curve classification, time series prediction).

Florence Forbes, who is a research scientist in the Institut National de Recherche en Informatique et Automatique (INRIA) and leader of the statistics team MISTIS. Her research activities include Bayesian image analysis, Markov models and hidden structure models, data fusion and integration, and inference methods using variational approximation techniques.

Anatoli Juditsky, who is a professor of statistics at the Université Joseph Fourier and leader of the SMS team. He received his PhD from ICS Moscow in 1989 and he was an INRIA research fellow from 1990-1999. His research interests include nonparametric statistics, system identification and stochastic optimization.

Cordelia Schmid, who is a research director in INRIA and leader of the computer vision team LEAR. Her main research interests are visual object recognition, image description and machine learning.

Bill Triggs, who is a research scientist in the Centre National de Recherche Scientifique (CNRS) and deputy director of the Laboratoire Jean Kuntzmann. His research interests include learning based methods for visual object recognition, scene understanding and human-machine interfaces. He will serve as scientific coordinator for the LJK site in PASCAL2.

UJM-CNRS – Laboratoire Hubert Curien (UMR 5516), Saint Etienne

The computer science team at the Laboratoire Hubert Curien (Saint Etienne) is a Joint Research Unit/UMR set up by the Université Jean Monnet and the CNRS.

It comprises 12 Professors (Professeurs) and Associate Professors (Maîtres de Conférences), and currently 6 non permanent members (doctoral students and post doctorates). The team is in charge of the computer science studies at the University of Saint-Etienne; these correspond to the BSc (Licence d'Informatique) and Masters (Master de Web Intelligence).

Its main research activities are in the fields of Machine Learning and Data Mining. Researchers from the Laboratoire Hubert Curien have successfully been publishing in the main conferences (ICML, COLT, ECML, ALT, ICGI...) and journals (Machine Learning Journal, Journal of Machine Learning Research, PAMI, PR journal…) of their area. Results from the team have been theoretical, but also applied, with specific interest to image analysis, web mining and speech technologies.

Members of the Laboratoire Hubert Curien team have been more particularly successful in the subfield of Grammatical Inference, also known as Grammar Induction. The team is arguably one of the most active in the world on this topic. They have presented papers at the relevant conferences of Grammatical Inference, and have been involved with all the scientific operations (conferences, tutorials, workshops, special issues…) regarding this theme over the past 5 years.

During PASCAL (2003-2006), the team has actively participated to joint activities with several nodes from the Network: Université Paris VI (common research projects, common workshops), XEROX (student exchange), Universitat Politecnica de Catalunya (long term stays, visits), Helsinki University of Technology (long term stays, visits), University of Warwick (long term stays, visits), Southampton (student exchange), Royal Holloway (student exchange, visits, common papers), University of London (students, common papers, visits), University of Alicante (long term stays, visits, common papers).

They have also associated to PASCAL, through various ways, other teams: for example LIM at Marseilles, France and Demokritos in Athens, Greece. The participants from the Laboratoire Hubert Curien site include:

Marc Sebban, whose primary interests are in Boosting and statistical machine learning, and who is involved with learning edit distance weights and in a more general way with techniques dealing with trees and distributions over strings and trees.

Colin de la Higuera who is interested in grammatical inference (with many aspects of this field), combinatorics over strings and trees and query (or active) learning in imperfect settings.

LM-CNRS, Laboratoire de Mathématiques (UMR 8628), Université Paris Sud

Founded in the 50s, the Laboratoire de Mathématiques at the Université Paris-Sud is a world-renowned Maths Labs (Field medals in 2006, 2002, 1994). It is a Joint Research Unit/UMR setup by the CNRS and the Université Paris Sud (Paris 11). It comprises about 40 Professors and 80 Associate Professors.

The Pascal participants from the LM site include:

Pascal Massart, Professor of Mathematics, head of the Probabilities and Statistics group. He received the Committee of Presidents of Statistical Societies (COPSS) award for outstanding research in 1998. His research interests include empirical processes, concentration inequalities, nonparametric statistics and model selection.

Elisabeth Gassiat, Professor of Mathematics. Her research interests include sequence modelling, lossless coding strategies and Gaussian processes.

Gilles Celeux, Senior scientist at INRIA. His research interests include model selection, mixture models, unsupervised clustering.

LRI-CNRS, Laboratoire de Recherche en Informatique (UMR 8623), Université Paris Sud

Founded in the 70s, the Laboratoire de Recherche en Informatique at the Université Paris-Sud comprises about 40 Full Professors and 80 Associate Professors.. It is a Joint Research Unit/UMR setup by the Université Paris Sud (Paris 11) and the CNRS.

The Pascal participants from the LRI site include:

Michele Sebag, Senior scientist at CNRS. Her research interests include relational learning, evolutionary optimization, autonomic computing. She was site manager of Université Paris Sud in Pascal-1 and member of the Steering Committee. In Pascal-2 she will be co-manager of the Challenge and Harvest programmes.

Marc Schoenauer, Senior scientist at INRIA. His research interests include evolutionary optimization, optimal design, robotics, inverse problems. He is editor in chief of the MIT Review Evolutionary Computation.

Olivier Teytaud, Scientist at INRIA. His research interests include dynamic programming, reinforcement learning, optimization, resampling methods. He is one of the authors of the award winner MoGo program, which was acknowledged a major breakthrough in the Computer-Go world.

B.2.2.5 XEROX SAS

The Xerox Research Centre Europe (XRCE) in Grenoble, France will be the Xerox entity collaborating in the project. XRCE will lead the Harvest Programme.

Dr Nicola Cancedda manages the "Cross-Language Technologies" Area at XRCE. He joined XRCE in 1999 after receiving a PhD in Computer Engineering from the University of Rome "La Sapienza". Before joining XRCE he was international fellow at the Artificial Intelligence Center of SRI International in Menlo Park, California, consultant for Finsiel SpA in Rome, Italy, and visiting student at the IRST in Trento, Italy. His main research interest is the design of Machine Learning methods for natural language processing and textual information access. He was XRCE site manager for the IST Network of Excellence PASCAL, as well as a member of its Steering Committee. Nicola Cancedda is currently coordinator of the FP6 STReP "Statistical Multilingual Analysis for Retrieval and Translation" (SMART). In PASCAL2 he will be in charge of the Harvest programme.

Dr Marc Dymetman is a principal scientist at XRCE, where he has been working since 1994. He holds a Doctorat d'Etat in Computer Science from the University of Grenoble. He is the author of numerous publications in computational linguistics, especially in the areas of symbolic and probabilistic machine translation, mathematical models of parsing and generation, multilingual document authoring. He has served on the programme committees of conferences in computational linguistics, machine translation, logical approaches to natural language. He is currently primarily involved in projects involving the use of machine learning techniques for computational linguistics tasks. In PASCAL2 He will be co-managing the thematic programme on "Leveraging Complex Prior Knowledge for Learning".

B.2.2.6 Jožef Stefan Institute

Jožef Stefan Institute (JSI) is the central research institution for natural sciences in Slovenia. It consists of over 800 researchers within 25 departments working in the areas of computer science, physics, and chemistry and biology. The Department of Knowledge Technologies is one of the largest European research groups working in the areas of machine learning and data mining. It has approx. 40 researchers covering different aspects of data analysis with special emphasis on textual data, social networks/graphs, complex data visualization, cross modal analysis, temporal (stream) data and in particular on scalability of approaches and deployability of research results in real world environments. In the recent years the research shifted towards semantic technologies, where the main goal is to combine modern statistical data analytic techniques with more traditional logic based knowledge representations and reasoning techniques. The department developed several software tools, among others: *Text-Garden* suite of text mining tools, *OntoGen* system for ontology learning, *Document-Atlas* for complex visualization. The Centre for Knowledge Transfer, also participating on ACTIVE proposal, consists from 10 researchers and technical staff working in the areas of research results dissemination and e-learning. In particular, the centre is well known by the portal http://videolectures.net with multimedia materials of numerous scientific events, on-line training materials, and collection of tutorials on different scientific fields. The centre is covering training and dissemination activities for many EU FP6 projects. Both of the above described JSI departments participate on several European projects from the three areas: Semantic & Knowledge Technologies (SEKT-IP, NEON-IP, ALVIS-STREP, TAO-STREP, SWING-STREP), KM for Networked Organizations (ECOLEAD-IP, E4-STREP), and Cognitive Systems (PASCAL-NoE, SMART-STREP).

Marko Grobelnik is researcher and manager of a research group of 15 people at the department of Knowledge Technologies working primarily in the areas of text-mining, semantic web and social network analysis. He is coauthor of several book and numerous scientific papers. Marko is technical director of FP6 IST World project on analysis of European research, he is a member of management board of several FP6 projects (SEKT-IP, NEON-IP, PASCAL-NoE), he participates in W3C standardizing committees. He co-organized over 10 international workshops and tutorials on text mining and link analysis at main international conferences, such as IJCAI, ACM-KDD, IEEE-ICDM, ICML.

Mitja Jermol is head of Centre for knowledge Transfer at JSI working in the area of e-learning, dissemination and promotion of research results. His main research area is Knowledge Management enriched with modern analytic techniques in the context of improved business processes for (networked) organizations. Among others, he is JSI's representative in the leading FP6 IP ECOLEAD in the area of networked organizations and virtual enterprises. Before joining JSI, Mitja was heading the group for distance education and e-learning in one of the largest national publishing houses of educational material.

Dunja Mladenič is researcher, project manager and financial deputy head of the Department of Knowledge Technologies. She is active researcher in the area of machine learning and text mining. She co-edited several books on data mining, she is author of several papers on machine learning, data mining and text mining, she is program chair of European Machine Learning Conference 2007. In FP5 she was coordinating R&D project SolEuNet on "*Data-Mining and Decision Support for Business Competitiveness: A European Virtual Enterprise*" (3.5M euro, 12 partners). She is also on management board of several EU FP6 projects including SWING, SMART, IMAGINATION, KDubiq. Among others, Dunja was long-term visitor at School of Computer Science, Carnegie Mellon University USA in 1997 and in 2001.

B.2.2.7 Università degli Studi di Milano

The department of the University of Milan responsible for the activities of the network is the Information Sciences Department (DSI). DSI is the largest computer science department in the Milan area and offers programs in computer science for undergraduate and graduate students. There are 30 faculty members whose research interests include: machine learning, databases and security, distributed algorithms and mobile computing, performance evaluation, computational logic, analysis of algorithms. Members will be involved in management of the thematic programme "Learning with partial or delayed feedback" and of the "External Visitor Programme".

Nicolò Cesa-bianchi received his Ph.D. degree in computer science from the University of Milan in 1993. Since 2004, he is professor of computer science at the University of Milan. His main research interests are in the areas of game-theoretic learning, statistical learning theory, and machine learning for pattern analysis. He is author of the monograph "Prediction, Learning, and Games" (Cambridge University Press, 2006). He was program chair of the thirteenth annual conference on Computational Learning Theory (COLT 2000) and of the 13th International Conference on Algorithmic Learning Theory (ALT 2002). He was site manager in the FP5 project KerMIT, site manager and steering committee member in the FP6
Network of Excellence PASCAL, and site manager in the FP6 STREP SMART. Since 2005, he is action editor for the Machine Learning journal. Since 2006, he is president of the Association for Computational Learning.

Giorgio Valentini graduated in Biology and Computer Science at the University of Genova, and he received a PhD in Computer Science from the same university in April 2003. He is currently assistant professor of computer science at the University of Milan. In 2001 he was visiting fellow at the CS Dept. of the Oregon State University (USA), working on ensemble methods based on bias-variance analysis (under the supervision of Tom Dietterich). He is author or co-author of 16 publications in international bioinformatics and machine learning journals and of more than 40 publications in peer-reviewed proceedings of national and international conferences. He is member of several scientific program committee of bioinformatics conferences, chair of CIBB 2007, Fourth International Conference on Computational Intelligence Methods for Bioinformatics and Biostatistics, Portofino, Italy, and chair of the SUEMA 2007 workshop on Supervised and Unsupervised Ensemble Methods and Their Applications (in conjunction with IbPRIA2007), Girona, Spain. He is member of ISCB (International Society of Computational Biology), INNS (International Neural Networks Society), BITS (Italian Bioinformatics Society).

B.2.2.8 University of Bristol

Nello Cristianini is affiliated to both the Engineering Mathematics and the Computer Science Department. Tijl De Bie is affiliated with the Engineering Mathematics Department. Both departments are part of Bristol University's closely integrated Faculty of Engineering. The research of the Artificial Intelligence Group was rated 5 in the last RAE. This proposal connects with several broad themes of research in Bristol, from the Pattern Analysis to the Exabyte initiative. Importantly, this project will interface with two existing grants held by Nello Cristianini, funded by the NIH and the European Union. The first one involves integration of heterogeneous genomic data, the other the analysis of multilingual text.

Nello Cristianini is a Professor of Artificial Intelligence at the University of Bristol since March 2006, and a holder of the Royal Society Wolfson Merit Award. He has wide research interests in the area of computational pattern analysis and its application to problems ranging from genomics, to computational linguistics, web analysis and artificial intelligence systems. In the past two years he has been involved in extensive preliminary work on automatic media monitoring, coordinating a group of students and programmers, and developing a prototype infrastructure. The system, that is currently working, will form the basis for the design of this project. Before the appointment to Bristol he has held faculty positions at the University of California, Davis, and visiting positions at the University of California, Berkeley, and in many other institutions. Before that he was a research assistant at Royal Holloway, University of London. He has also covered industrial positions. He has a PhD from the University of Bristol, a MSc from Royal Holloway, University of London, and a Degree in Physics from University of Trieste. Since 2001 has been Action Editor of the Journal of Machine Learning Research (JMLR), and since 2005 also Associate Editor of the Journal of Artificial Intelligence Research (JAIR). He is co-author of the books 'An Introduction to Support Vector Machines' and 'Kernel Methods for Pattern Analysis' with John Shawe-Taylor (both published by Cambridge University Press). A third book, "Introduction to Computational Genomics" is in Press. He is currently involved in a NIH funded project for the development of algorithms to combine heterogeneous data sources in genomics and in a long term effort to apply modern learning algorithms to the analysis of web data. This project has been partially supported by various sources, including Fair Isaac Co., and the University of California. Another project, (SMART) sponsored by the EU, involves the development of methods for Statistical Machine Translation. Expertise from this project will be directly leveraged in the media-content pattern analysis.

Tijl De Bie has been appointed Lecturer in Artificial Intelligence at the University of Bristol in January 2007. Prior to that he has worked at the K.U.Leuven, Belgium, and Southampton University. During his PhD research at the K.U.Leuven, he spent a total of one year as a visiting research scholar at U.C.Berkeley and U.C.Davis. His research interests range from statistical analysis of machine learning algorithms, the use of optimization to design such algorithms, to their efficient implementation and application to genomic problems. He has published in each of these areas, and his tools are currently used in genomic laboratories. He is currently working on a joint book on Pattern Analysis. He has been invited to lecture about pattern analysis in international conferences, and has published in international journals and conferences. His collaboration with Nello Cristianini has taken place during various research visits in California and UK, and has been fruitful in terms of applications, as well as publications.

B.2.2.9 University of Manchester

The University of Manchester is the U.K.'s largest single site university with an international reputation for research led teaching. The principal contribution to PASCAL from Manchester will be from the School of Computer Science which was judged to be one of the six top ranked (5*) UK Computer Science departments in the most recent research assessment exercise. Manchester has a long standing interest in machine learning through **Dr Jon Shapiro** (expertise: dynamics of learning, statistical mechanics analysis of learning) and **Dr Magnus Rattray** (expertise: Bayesian methods, phylogenies, computational biology). Over the last two years Manchester has invested heavily in this area through three recent appointments **Dr Neil Lawrence** (expertise: Bayesian methods, Gaussian processes, computational biology), **Dr Joshua Knowles** (expertise: multi-objective optimisation, computational biology) and **Dr Gavin Brown** (expertise: ensemble learning theory, semi-supervised learning). **Dr Elzbieta Pekalska** is an EPSRC research fellow in the group (expertise: representation of dissimilarities). As well as these researchers (who each have machine learning and optimisation as their core interests) the School has research groups in advanced interfaces (led by Prof. Roger Hubbold) with interests in haptics, virtual reality and visualisation. A particular connection from this group with the areas of interest within PASCAL is through **Dr. Aphrodite Galata** who works in Computer Vision and behaviour modelling with a particular emphasis on interpretation of dynamic scenes. The Division Imaging Science and Biological Engineering group (led by Prof. Chris Taylor) is joint hosted with the Medical School and has a strong focus on image understanding, with particular inspiration from the field of Computer Vision. These groups both provide applications and expertise that is of interest across all the core themes of the PASCAL network.

As well as interest in vision, haptics and image understanding outlined above, Manchester is a key site in the exploitation of machine learning techniques within Computational and Systems Biology. Lawrence and Rattray are key contributors to the existing PASCAL Thematic Programme in Learning in Computational and Systems Biology. This programme has already engaged in a highly successful workshop in Parameter Estimation in Systems Biology, the themes of which are expected to feature prominently in TP1: Leveraging complex prior knowledge for Learning. Manchester has a large EPSRC funded Systems Biology centre housed within the Manchester Interdisciplinary Biocentre. Through this centre and the Faculty of Life Sciences, Manchester bridges the gap between computation and biology to provide access to expertise in systems modelling and raw data for challenging inference problems in Biological Systems Analysis.

Manchester will contribute to PASCAL2 in several key ways, as well as providing a bridge to computational and systems biology, vision, haptics and image understanding as applications of the technologies PASCAL2 develops Manchester will undertake core research in machine learning technologies with a particular emphasis on the Bayesian approach. We expect across the board contributions from the Manchester site, but they are likely to be particularly active in TP1 (co-managed by Lawrence), particular in the context of biological applications of machine learning.

B.2.2.10 University of Helsinki

University of Helsinki (UH) is a research intensive university founded in 1640. It is Finland's most multidisciplinary higher education institution, and it was in 2002 elected to be one of the research-intensive, highly ranked European universities forming the League of European Research Universities (LERU) consortium. In recent evaluations, the Department of Computer Science at the University of Helsinki was awarded the highest possible score for its quality both in research and teaching. The research at the Department combines theory with interaction with the application fields in a well-balanced whole. Pattern analysis, statistical modelling and computational learning belong to the focal areas of research at the department, and the department hosts several internationally well-known research groups working on various topics related to these areas. The past application areas of this research include bioinformatics, mobile computing, user modelling, context-aware computing and information retrieval.

Petri Myllymäki leads the Complex Systems Computation research group focusing on Bayesian and information-theoretic modelling. He is also the head of the Intelligent Systems specialization area at the Department. In PASCAL2 he is the manager of Work Package 11 (Conference and Workshop Attendance Programme).

Juho Rousu's research interests lie in algorithms and machine learning, especially kernel methods for structured domains, with applications to bioinformatics and NLP.

Jyrki Kivinen is the head of the Algorithms specialization area at the Department. His main research interest is computational learning theory, in particular on-line learning.

Esko Ukkonen is Research Director of the Basic Research Unit of HIIT and the head of a Center-of-Excellence in Algorithmic Data Analysis of the Academy of Finland. His research area is combinatorial pattern matching and its applications in systems biology.

Patrik Floréen works on ubiquitous computing and sensor networks with methods from combinatorial optimization and probabilistic modelling.

Jorma Rissanen is best known for his seminal work on arithmetic coding and the Minimum Description Length (MDL) Principle.

Hannu Toivonen works on data mining, especially on pattern and link discovery in structured and heterogeneous data, with applications in bioinformatics, genetics, and in ubiquitous computing.

B.2.2.11 Fondation de l'Institut Dalle Molle d'Intelligence Artificielle Perceptive

The IDIAP Research Institute is an independent, not-for-profit, research institute located in Martigny (Switzerland). IDIAP is involved in numerous national and international (EU and US) projects, as well as in multiple collaborative projects with industry. With a research staff of more than 75 scientists, the primary missions of IDIAP are research, education, and technology transfer in the areas machine learning, speech and audio processing, computer vision, information retrieval, biometric authentication, multimodal interaction, and multiple multimodal research activities across these disciplines. At the national level, IDIAP is also the "Leading House" of the National Centre of Competence in Research (NCCR) on "Interactive Multimodal Information Management" (IM2, www.im2.ch). At the EU level, IDIAP is involved in numerous projects and Networks of Excellence, and is currently managing 2 Integrated Projects (AMIDA www.amidaproject.org and DIRAC www.diracproject.org), and one STREP (MAIA http://www.maia-project.org). In the US, IDIAP is also involved in numerous collaborations with ICSI Berkeley, partner of a large DARPA project (GALE) and coordinator of a DTO project (Roadmap, as part of the VACE program).

José del R. Millán is a senior researcher at the IDIAP Research Institute in Martigny, Switzerland, where he explores the use of brain signals for multimodal interaction and, in particular, the development of non-invasive brain-controlled robots and neuroprostheses. In this multidisciplinary research effort, Dr. Millán is bringing together his pioneering work on the two fields of brain-computer interfaces (BCI) and adaptive intelligent robotics. He is also an adjunct professor at the Swiss Federal Institute of Technology in Lausanne (EPFL). He received his Ph.D. in computer science from the Univ. Politècnica de Catalunya (Barcelona, Spain) in 1992. Prior to joining IDIAP, he has been a research scientist at the Joint Research Centre of the European Commission in Ispra, Italy, a visiting professor at the EPFL, and a visiting scholar at Stanford Univ. His research on BCI was nominated finalist of the European Descartes Prize 2001 and he has been named "Research Leader 2004" by the journal Scientific American for his work on brain-controlled robots. The journal Science has reviewed his work as one of the world's key researchers in the field of BCI. Dr. Millán is the coordinator of a number of European projects in the field of BCI and also is a frequent keynote speaker at international events.

François Fleuret got the PhD in Probability from University of Paris VI in 2000. After one year at the department of Computer Science of University of Chicago, he was hired as a full researcher at the INRIA in France. In 2004 he joined the CVLab at EPFL where he spent three years as a senior researcher, before coming to IDIAP in 2007 as a senior researcher in machine learning. His main research interests are at the interface, centred on the development of algorithmically efficient machine learning techniques.

Alejandro Jaimes is Scientific Manager and Senior Researcher at IDIAP Research Institute where he is responsible for managing the research efforts of 12 partners within the EU-funded AMIDA integrated project, as well as leading the Human-Machine Interaction unit of the Swiss NCCR IM2. His technical research focuses on creating new technical approaches for computer-understanding of multimedia content and for human interaction with computers in creative environments, mainly developing computer vision techniques that use machine learning, that involve humans directly, and that are rooted in principles, theories, or techniques from cognitive psychology, the arts, and information sciences, among others. He is a founding member of the IEEE Computer Society Taskforce on Human-Centered Computing, as well as founder, co-organizer, and co-curator of the ACM Multimedia Interactive Arts program (2004-2007).

B.2.2.12 Stichting Centrum voor Wiskunde en Informatica

CWI is the National Research Institute for Mathematics and Computer Science in the Netherlands. This application concerns the CWI research group information-theoretic learning (INS 4.2), which is headed by **Dr. Peter Grünwald**. Grünwald is an expert on information-theoretic approaches to machine learning and statistics. He has recently completed the first comprehensive book on the Minimum Description Length (MDL) principle, the standard information-theoretic approach to statistics. Complexity and representation of prior knowledge are central notions in this approach, which is ideally suited for dealing with structured domain (indeed, Grünwald has applied the approach to language and grammar learning). This work lies in the intersection of computer science, theoretical statistics and information theory, and Grünwald has published in top journals in all three fields. Grünwald also holds an interest in applications of statistics and learning in psychology and cognitive systems, as witnessed by several publications in leading psychological journals (Journal of Mathematical Psychology, Psychological Science) and the annual cognitive science conference, and the fact that in 2005, he co-edited a book with two psychologists. In the Sixth framework PASCAL Network of Excellence, Grünwald was member of the steering committee and manager of the highly succesful conference and workshop organization programme. He has intensely collaborated with three other PASCAL sites: University of Helsinki, University College London and EURANDOM Eindhoven. In 2004, he has been awarded the prestigious VIDI-grant by the Netherlands Organization for Scientific Research for the research project titled learning when all models are wrong.

The other senior group member, **Professor Paul Vitányi** is a leading authority on Kolmogorov complexity and co-author of the standard textbook in the field. Reports of his recent research into compression-based clustering and similarity analysis (related to the aforementioned MDL Principle) can be found in, among others, the New Scientist, Pour la Science (the French scientific American) and Izvestija. This work was partly funded by a PASCAL pump-priming grant. Like Grünwald, Vitányi has a keen interest in psychology and cognitive science, witness publications in the Journal of Mathematical Psychology and Trends in Cognitive Science. His recent co-authors include the cognitive psychologist Nick Chater of University College London. During his career, Vitányi has given keynote lectures and numerous conferences, and has received numerous awards, including the Kolmogorov medal and the Medal of the University of Helsinki.

B.2.2.13 Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.

The Intelligent Data Analysis (IDA) research group is part of the German Fraunhofer-Gesellschaft. The Fraunhofer-Gesellschaft undertakes applied research of direct utility to private and public enterprise. The IDA group is comprised of around 35 members (1 full Professor, 2 tenured senior researchers, 10 PostDocs, and 12 researchers). IDA is concerned with learning systems for data analysis. In this context, IDA is developing tools for high-dimensional multivariate statistics based on methods that stem from classical statistics, neural networks, signal processing and machine learning communities. The major objective of IDA is to pursue research all the way from theory to application. To this end, IDA is contributing to the theoretical development of new data analysis techniques, to their implementation according to software-engineering standards, and, finally, to their industrial and academic application. In terms of mathematical models, IDA is mainly working in the areas of classification and regression, time series analysis, and dimensionality reduction techniques. Here we strive for improved learning methods particularly for the cases of multivariate small/very large sample statistics. The goal is to develop more robust, stable and efficient learning techniques that also allow us to gain insights and understand data of a complex origin. In the scientific community the IDA group is known as one of the most active research centers in the field of machine learning, especially in the topics of support vector machines (SVMs), kernel-based learning methods, boosting, brain-computer-interfacing, and independent component analysis (ICA).

Klaus-Robert Mueller is the head of the IDA group at Fraunhofer FIRST. He holds a MSc degree in physics and a PhD degree in computer science from Uni Karlsruhe. He was appointed as a professor of computer science at Uni Potsdam in 1999. Since May 2006, he is a full professor for machine learning at TU Berlin. His research focus is on Support Vector Machines and other kernel-based learning methods, Boosting and ensemble algorithms, neural network models, and source separation.

Anton Schwaighofer is a senior researcher at Fraunhofer FIRST. His research interests are on Bayesian learning techniques, in particular robust and large-scale methods, and industrial applications of ML techniques. He has been site manager for Fraunhofer FIRST in Pascal, member of the Pascal steering committee and active in the organization of several Pascal workshops and challenges.

B.2.2.14 Max-Planck-Gesellschaft zur Förderung der Wissenschaften

The Empirical Inference Department at the Max Planck Institute for Biological Cybernetics, Tuebingen, Germany is directed by Prof. Bernhard Schoelkopf. Empirical Inference, the title of the department, describes the task of drawing conclusions from the observation of empirical data. The department was started around statistical learning theory and support vector machines (SVM). We have recently placed particular emphasis on semi-supervised learning and Bayesian methods. Our largest application areas are machine vision, biological vision, robotics, bioinformatics, brain-computer interface and computer graphics.

In PASCAL, our department members have been active in workshop organization and related activities. We co-organized one thematic programme and 2 challenges, and in total 13 members have been involved in the organisation of workshops and summer schools. Our members visited other sites 27 times and hosted 44 visits. Our department is deeply engaged in the former PASCAL network and has been one of the central sites.

Prof. Bernhard Schölkopf was born in Stuttgart on 20 February, 1968. He received an M.Sc. in mathematics and the Lionel Cooper Memorial Prize from the University of London in 1992, followed in 1994 by the Diplom in physics from the Eberhard-Karls-Universität, Tübingen, with a thesis written at the Max-Planck-Institute for Biological Cybernetics. Three years later, he obtained a doctorate in computer science from the Technical University Berlin. His thesis on Support Vector Learning won the annual dissertation prize of the German Association for Computer Science (GI). In 1998, he won the prize for the best scientific project at the German National Research Center for Computer Science (GMD). He has researched at AT&T Bell Labs, at GMD FIRST, Berlin, at the Australian National University, Canberra, and at Microsoft Research Cambridge (UK). He has lead the New York City Office of Biowulf Technologies, and taught at the Humboldt University and the Technical University Berlin. In July 2001, he was elected scientific member of the Max Planck Society and director at the MPI for Biological Cybernetics; in October 2002, he was appointed Honorary professor for Machine Learning at the Technical University Berlin.

Koji Tsuda will be the representative of the department for PASCAL 2. He is Project Leader for Bioinformatics. In the formal PASCAL framework, he is now organizing a workshop *Mining and Learning with* Graphs in Italy in August, 2007. He will be the co-manager of the Conference and Workshop Attendance Programme.

B.2.2.15 Helsinki University of Technology

The Laboratory of Computer and Information Science (CIS) at Helsinki University of Technology (TKK) is famous for its pioneering research on neural networks, Self-Organizing Maps (SOMs), Independent Component Analysis (ICA) and data mining. The laboratory currently consists of 5 professors and about 80 staff altogether. The laboratory has been selected for 1994–1999, 2000–2005 and 2006–2011 as one of the Finnish national Centers of Excellence in Research. Most of its research activities are currently coordinated by two research units, Adaptive Informatics Research Centre (AIRC) and Algorithmic Data-Analysis (ALGODAN). To briefly list the main numerical achievements of the last two years, the laboratory produced 9 D.Sc. (Eng.) degrees, 4 Lic.Tech. degrees, and 42 M.Sc. (Eng.) degrees. In PASCAL2 TKK will manage the Conference & Workshop Organization programme and participate actively in the thematic programmes, most notably incorporation of prior knowledge (bioinformatics, language) in modeling and multi-component learning. Several of the research groups of the Adaptive Informatics Research Centre are already working on these problems, with applications in text, images and bioinformatics, including time series data.

Prof. Samuel Kaski is full professor of Computer Science in TKK/CIS and vice director of its Adaptive Informatics Research Centre. His research focus is in machine learning, two of the main application areas being information retrieval and bioinformatics. He leads a research group on Statistical machine learning and bioinformatics.

Prof. Erkki Oja is full professor of Computer Science, head of the laboratory of Computer Science and director of the Adaptive Informatics Research Centre. He is famous of his pioneering work on Independent Component Analysis and pattern recognition, and currently leads the algorithms and methods group of the research centre.

Prof. Heikki Mannila is full professor of Computer Science and currently Academy Professor of the Academy of Finland. He is a distinguished pioneer of algorithmic data analysis and in particular data mining. He leads the TKK branch of ALGODAN Centre of Excellence.

Dr. Jorma Laaksonen is the leader of the content-based information retrieval research group at AIRC and has led the development of the PicSOM content-based image retrieval system. The focus of his current research is especially retrieval of multimedia data.

Dr. Mikko Kurimo leads the Speech Recognition and Multimodal Interfaces research group of AIRC. The current research focus of his speech group is in language independent and unsupervised models for continuous speech with morpheme-based language models for very large vocabulary.

Dr. Timo Honkela leads the Cognitive Systems research group of AIRC. Computational Cognitive Systems group conducts research on artificial systems that combine perception, action, reasoning, learning and communication; this area of research draws upon biological, cognitive and social system approaches to understanding cognition.

Dr. Kai Puolamäki is a lecturer in CIS. He has been active in a key project on proactive information retrieval.

B.2.2.16 Bar-Ilan University

The Department of Computer Science at Bar Ilan University is a major center of research in several fields of theoretical and applied Artificial Intelligence. There has been a long tradition of research on text and language processing, as well as the study of empirical approaches. The School of Engineering also covers research areas related to the PASCAL scope, as detailed below.

Notably, our activities within the first round of PASCAL in the area of Textual Entailment, which was introduced by Dagan and Glcikman at the early PASCAL Workshop on Learning Methods for Text Understanding and Mining in January 2004, made an enormous impact on the research community. Textual entailment is now considered a "hot" and promising applied direction for addressing the highly challenging semantic level of human language, largely based on learning approaches. Clearly, this impact was driven by our PASCAL activities, in particular the challenges on Recognizing Textual Entailment in which about 40 research groups participated. Based on this experience, Bar Ilan will be involved in managing the Challenges program at PASCAL-2 and in directing the first Thematic Program on Leveraging Complex Prior Knowledge for Learning, which is partly focused on learning models for language.

Dr. Ido Dagan, Department of Computer Science: Natural Language Processing (NLP), Information Retrieval (IR) and related empirical methods. Ido Dagan is the PASCAL site manager for Bar Ilan. He is well known for his research in empirical and learning-based approaches to NLP. He is currently focusing on developing robust models of semantic processing that are amenable to automated and semi-automated learning, within the Textual Entailment framework.

Prof. Moshe Koppel, Department of Computer Science: Moshe Koppel's research focuses on text-based applications of machine learning. He is currently studying authorship attribution problems with huge numbers of candidates, collaborative ranking (e.g. Netflix) problems (with Jacob Goldberger), and unsupervised concept discovery from large corpora (with Ari Rappoport). He is on the program committees of SIGIR, ACL, EMNLP, AAAI and other conferences and has helped organize workshops on authorship and writing style at IJCAI, SIGIR and AAAI Fall Symposium. His research in these areas has been reported on in Nature, New Scientist, the New York Times, BBC, NPR and other outlets.

Jacob Goldberger, School of Engineering: Jacob Goldberger’s research interests are in machine learning and information theory with applications in computer vision, medical imaging, speech recognition, neuroscience and statistical signal processing. Jacob did his Master degree at the Tel-Aviv University in the mathematics department and earned his doctoral degree in 1997 from the Tel-Aviv University, Faculty of Engineering. He did a postdoc with Shimon Ullman at the Weizman institute and another postdoc with Geoff Hinton and Sam Roweis at the at the University of Toronto, machine learning group. He has also worked at several research labs including SRI and INRIA. He has recently worked on supervised dimensionality reduction, Bayesian methods, clustering, error-correcting codes and semi-supervised learning.

B.2.2.17 Universite Pierre et Marie Curie-Paris 6

Pierre & Marie Curie University (UPMC) is one of the largest universities teaching science and medicine in France, and in Europe, with 4000 researchers and teaching academics/researchers, 180 laboratories, and some 30,000 students including 8000 in postgraduate studies. Laboratoire d'Informatique de Paris 6 (LIP6) is the computer science lab at UPMC. LIP6 is about 350 researchers including PhDs and academics.

Research at LIP6 covers several domains in Computer Science (Micro-electronics, distributed systems, networks, software agents, machine learning, decision and optimization, etc). Researchers involved in Pascal2 are all from the Machine Learning team although activities relevant to Pascal2 are also developed in other teams. This group is concerned with statistical machine learning and focuses its activities on a few main application areas such as textual information access and information retrieval, user modelling, user interaction, biology. The team has been involved in a number of European and national research projects. It coordinates a French network of excellence on machine learning.

Patrick Gallinari is Professor at University Paris 6, Computer science lab. At UPMC, vice director of the Lab. and head of the Artificial Intelligence department. His research interests are in statistical machine learning, textual information retrieval, user modelling, gesture and handwriting recognition. LIP6 was the French node of Neuronet (Network of Excellence on Neural Networks), and P.G. was in charge of the research committee of this network for 4 years. He is also coordinator of a French excellence network on machine learning.

Massih Amini is assistant professor at University of Paris 6, Computer science lab. He received a PhD in Computer Science from UPMC in 2001. His research interests are statistical learning algorithms and textual information retrieval. He is author or co-author of about 20 publications including conference proceedings and a book chapter.

Thierry Artieres is assistant Professor at University of Paris 6, Computer science lab. He received a PhD in Computer Science from Paris Sud University in 1995. His research interests focus on statistical learning for signals and sequences, applied to speech and speaker recognition, on-line handwriting recognition and pen-based interfaces, user modelling through logs analysis. He is author or co-author of about 20 publications in international conferences and journals.

Ludovic Denoyer has played an active role in PASCAL activities particularly in the organisation of two challenges, the XML Document Mining Challenge, and the Web Spam Challenge. The latter of these two demands linking with research in graph theory and Ludovic will be organising an associated workshop on Graph Labelling at the ECML/PKDD conference in September 2007.

B.2.3 Consortium as a whole

The list of beneficiaries of PASCAL2 corresponds closely to the core sites of the FP6 NoE PASCAL, with some adjustments reflecting the restructuring that has occurred during the intervening period. An example is the inclusion of University College, London following its establishment of a Centre for Computational Statistics and Machine Learning partly inspired by the activities of the PASCAL Network.

In addition to the PASCAL2 beneficiaries who have been assigned the work described in this document, the Network is open for non-beneficiary membership to other parties engaged in research on "learning and adaptation in artificial systems" (see ICT Work Programme 2007-2008, Challenge 2). This type of membership will be established through the signing of a PASCAL2 *Membership Agreement* between the coordinator and the particular institution concerned. The Membership Agreement will flow down the principal rights and obligations from the PASCAL2 Grant Agreement onto the non-beneficiary member. It will allow the non-beneficiary members access to the restricted areas of the PASCAL2 website; they will also be eligible to submit to funding calls published on the PASCAL2 website, under any of the PASCAL2 activities.

Non-beneficiary membership status will be granted by default to all parties identified in the PASCAL2 proposal, section 2.2.18, as "Additional Academic Partners", representing a selection of the relevant experts in the different fields that concern the activities of the Network:

- Kings College London (UK)
- Politècnica de Catalunya (ES)
- NICTA (National Centre for Information and Communication Technologies) (Australia)
- University of Glasgow (UK)
- University of Leoben (AT)
- Fondazione Bruno Kessler (IT)
- Pompeu Fabra University (ES)
- Graz University of Technology (AT)
- Royal Holloway (UK)
- University of Sheffield (UK)
- University of Insubria (Varese) (IT)
- Kungliga Tekniska Hoegskolan (SE)
- Technical University of Denmark (DK)
- Tel Aviv University (IL)
- University of Antwerp (BE)
- Technion – Israel Institute of Technology (IL)
- Lublijana University – Institute of Mathematics, Physics and Mechanics (SI)
- Imperial College (UK)
- Universitaet Stuttgart, Institut fuer Maschinelle Sprachverarbeitung (DE)
- Alicante University (ES)
- Leiden University (NL)
- Oxford University (UK)
- London School of Economics (UK)
- The Hebrew University of Jerusalem (IL)

In addition, PASCAL2 will be open to applications from relevant research institutions interested in joining the Network. These will be referred to referees to assess the appropriateness and quality of the applicants and will then be considered by the Steering Committee at its next bi-annual meeting. We will also encourage Members of the Network to nominate new members that they feel would make an important contribution to the work of the network. Such applicants will be considered using the same procedure described above.

B.2.3.1 Sub-contracting

Not applicable.

B.2.3.2 Other countries

Not applicable.

B.2.3.3 Additional beneficiaries

Where a successful submission (to a call published on the PASCAL2 website, see above) by a non-beneficiary member results in payments for salaries, that member shall be eligible to apply to the Steering Committee to request a change in status to beneficiary. All such requests will be reviewed by the Steering Committee at their next bi-annual meeting. The request will then be forwarded to the European Commission for approval, in accordance with the procedure as set out in article II.37 of the Grant Agreement. Following approval by the Commission, the member will then be required to sign the form of accession to the Grant Agreement to be confirmed as a beneficiary.

B.2.3.4 Third Parties

Under the French "Joint Research Unit" regime the Centre National de la Recherche Scientifique (Partner 4) applies the special clause 10 of the EC Grant Agreement. These Third Party agreements are with:

- Université Joseph Fourier (UMR/JRU n°5224 - Laboratoire Jean Kuntzmann - LJK)
- Université Pierre-Mendès France (UMR/JRU n°5224 - Laboratoire Jean Kuntzmann - LJK)
- Institut National Polytechnique de Grenoble (UMR/JRU n°5224 - Laboratoire Jean Kuntzmann - LJK)
- Université Jean Monnet (UMR/JRU n°5516 - Laboratoire Hubert Curien)
- Université Paris Sud (UMR/JRU n°8628 - Laboratoire de Mathématiques and UMR/JRU n°8623 - Laboratoire de Recherche en Informatique)

Descriptions of the competences of these Third Parties and the work to be realized by them are presented in the section B2.2.

B.2.4 Resources to be committed

All costs are given on the form A3. However, due to the nature of the funding procedure, it is not possible at this time to determine the amount going to each site. An indicative figure has been used on the form A3 to illustrate a likely scenario. Full details of the budget are given in section 2.1.4.

All payments for the activity of researchers working at non-beneficiary member sites (as explained under B.2.3) will normally be made directly to the individuals concerned by UCL reimbursing, subject to the provision of original receipts or in exceptional cases appropriate invoices authorised by the steering committee, 100% of real and actual costs incurred with no additional overheads paid to the institution concerned. Reimbursement will be made for participation expenses of the researchers for travel, subsistence, consumables, workshop organisation and where appropriate equipment. In all cases reimbursement will only be made provided that a corresponding active unspent budget has been allocated for the given expenditure to the non-beneficiary site concerned and its site manager has countersigned the request. Note that no salaries can be reimbursed under this scheme (see B.2.3.3). UCL will be responsible for processing these claims in an efficient manner (normally within one month) and will enter the amounts as part of its cost claims to the EC. This explains UCL's relatively large budget allocation.

Other beneficiaries may also be involved in refunding some of the activities of researchers not associated with their institution.

B3. Potential impact

B.3.1 Strategic impact

Pattern analysis and machine learning are one of the central tools for structuring raw information in the knowledge economy. It is clear that most information will need to be filtered and restructured before it becomes usable, particularly incoming sensor data. Techniques that can quickly analyse complex patterns and adapt to new or difficult-to-interpret data will be indispensable for maintaining a competitive edge in this society. Leading-edge technology will be crucial for the competitiveness of both established companies and start-ups in this area. At the same time the potential social benefits for the knowledge-based society are clear, particularly in enabling wider access to information through user interfaces that are accessible to a more diverse class of users. Furthermore the research proposed in this project will significantly increase the potential for useful applications of machine learning in the analysis of complex scientific data, where advanced machine learning techniques are being increasingly used in place of a traditional research cycle of hypothesise/experiment.

Most importantly, what the proposed research promises over and above a simple extension of data analysis is the ability to bring the complexity of the inferences that can be made to a new 'cognitive' level. By this we mean that rather than just answer specific predefined questions, systems will be able reliably to identify and make use of unforeseen patterns that can guide the formation of new representations able to enhance their operation. This ability will have a fundamental and far-reaching impact, whether the system be an interface providing value added intelligence, a robot sensor system exploring a new environment, or a sensor network monitoring and controlling a complex application.

B.3.2 Spreading excellence, exploiting results, disseminating knowledge

As mentioned in section 1, the distributed nature of the "PASCAL Institute" holds distinct advantages in its ability to reach into an extended range of different national, scientific and commercial communities, while maintaining its own inner coherence and dynamic. This is evident for example in the diverse range of smaller-scale projects being funded through different national programmes – there will be a mutual interchange between the research of PASCAL2 and these different research projects that help to spread the excellence of PASCAL2 research, move to exploit its results and disseminate its knowledge

.

The infrastructure/pump-priming programme will support mini research projects as well as the development of enabling technologies in the form of software packages that will assist researchers from other disciplines to make rapid use of the results of the research. These will also provide exemplars that can motivate the development of spin-off projects and commercial applications. The pump-priming programme is designed to bring new ideas to a level of maturity where funding for focussed projects can be sought. PASCAL2 will continue this tradition in the area of cognitive systems related research and applications.

PASCAL2 includes a set of Additional Industrial Partners that will form an Industrial Club. They will provide input concerning the direction of the research through input to the Scientific Management Programme and hence the Steering Committee for its meeting when the Thematic Programmes are selected. This will also be a forum for more general input concerning the operation of the network and the research directions that could have significant impact in developing applications. The Industrial Club will also have a crucial input to the Harvest Programme both in suggesting topics, hosting collaborations, and potentially exploiting the results of these focused projects. This programme will fund the execution of joint projects between academic partners and industrial members. Unlike pump-priming projects, Harvest projects will be of short duration (e.g. three months) and will be carried out by physically co-located teams of around 6 members with the possible remote involvement of RA and student supervisors. All proposals will necessarily include some software development, and require the participation of at least one third of non-industrial and one third industrial staff. Moreover, all proposals must include some training delivered from the participant academics.

The project team will be expected to do all preliminary work (background reading, identification of main directions) in advance, so as to be quickly up to speed once the project has started. During execution, project advancement will be monitored by some external "project officer" (e.g. a member of the Harvest programme committee) by means of interactions with the coordinator only. The PASCAL Forge will be used to provide a centralised repository, a showcase and a document trail of the project execution.

At the end, all projects will demo their results in an annual dedicated "Harvest Festival" workshop. Resulting software will be reviewed to ensure that it complies with readability and documentation requirements enabling industrial teams to pick up from where the project team ended. It is understood that the software will be proof-of-concept, not a market-ready package. All participating industrial partners will have (possibly royalty-bearing) full commercial rights on the

sources. The latter could also be made available to the general public as an open source project, but with a strictly non-commercial license.

B4. Ethical issues

Not applicable.

B5. Consideration of gender aspects

There will be a Gender aspects representative responsible for monitoring the gender balance and gender relevance of the research championed by the Network. They will be responsible for participating in and initiating relevant initiatives to encourage women to become involved at all levels of science and especially in cognitive science and machine learning research. They will have a position on the Steering Committee and deliver a report at least once a year on developments in this area.