COVER PAGE

Title of Proposal: **MathVRE: Collaborative ecosystems for mathematical research and software development**

Date of preparation: **December 2, 2014**

Participant no	Participant organisation name	Country
1 (Coordinator)	Université Paris Sud	FR
2	University of St Andrews	UK
3	Logilab	FR
4	Université Bordeaux	FR
5	University of Kaiserslautern	DE
6	University of Oxford	UK
7	University of Warwick	UK
8	University of Silesia	PL
9	Université de Versailles Saint-Quentin	FR

Contents

♠**EC Commentary**: *Please follow the structure of this template when preparing your proposal. It has been designed to ensure that the important aspects of your planned work are presented in a way that will enable the experts to make an effective assessment against the evaluation criteria. Sections 1, 2 and 3 each correspond to an evaluation criterion for a full proposal.*

Please be aware that proposals will be evaluated as they were submitted, rather than on their potential if certain changes were to be made. This means that only proposals that successfully address all the required aspects will have a chance of being funded. There will be no possibility for significant changes to content, budget and consortium composition during grant preparation.

Page limit: The cover page, and sections 1, 2 and 3, together should not be longer than 70 pages. All tables in these sections must be included within this limit. The minimum font size allowed is 11 points. The page size is A4, and all

margins (top, bottom, left, right) should be at least 15 mm (not including any footers or headers). If you attempt to upload a proposal longer than the specified limit, before the deadline you will receive an automatic warning, and will be advised to shorten and re-upload the proposal. After the deadline, any excess pages will be overprinted with a 'watermark', indicating to evaluators that these pages must be disregarded.
Please do not consider the page limit as a target! It is in your interest to keep your text as concise as possible, since experts rarely view unnecessarily long proposals in a positive light.♠

Outline of Project (for Proposers)

♠**TO DO**: *This is the place for various READMEs not included in the final submission*♠

Mission statement for the grant

Our mission is to promote the next generation of community-developed open source software, databases, and services adapted to the needs of collaborative research in pure mathematics and applications.
Our research will cover a wide variety of aspects, ranging from software development models, user interfaces ♠**TO DO**: *virtual environments?*♠, deployment frameworks and novel collaborative tools, component architecture, design, and standardization of software ♠**TO DO**: *system?*♠ and databases, to links to publication, data archival and reproducibility of experiments, development models and tools, and social aspects.
It will consolidate Europe's leading position in computational mathematics and build on the remarkable success of the ecosystem of projects GAP, Python/Sage, Pari, Singular, LMFDB.
♠**TO DO**: *What do we meand by "new generation"*♠.

1 Excellence

The focus of **MathVRE** is on promoting community-developed open source software, databases, and services adapted to the needs of collaborative research in pure mathematics and applications.
It will consolidate Europe's leading position in computational mathematics and build on the remarkable success of the ecosystem of projects like `GAP`, `Python`, `Sage`, `Pari`, `Singular`, `LMFDB`.
♠**TO DO**: *Something about the special value of collaborative VREs for mathematics – research groups are small and can be isolated.*♠

1.1 Objectives

♠**EC Commentary**: *1-2 pages*♠ ♠**EC Commentary**: Describe the specific objectives for the project, which should be clear, measurable, realistic and achievable within the duration of the project. Objectives should be consistent with the expected exploitation and impact of the project (see section 2).♠

Aim 1: Improve the productivity of researchers by promoting collaborations on Mathematical *software*, *data*, and *knowledge*.

Aim 2: Make it easy for small to large teams of researchers to setup custom collaborative Virtual Research Environments adapted to their needs and workflow, supporting the entire life-cycle of computational work in mathematical research, from initial exploration to publication, teaching and outreach.

Aim 3: Enable reproducible research and promote ♠**TO DO**: *dissemination, easier access, reuse, sharing ...*♠

♠**TO DO**: *make those into concrete objectives*♠
Our research will cover a wide variety of aspects, ranging from software development models, user interfaces ♠**TO DO**: *virtual environments?*♠, deployment frameworks and novel collaborative tools, component architecture, design, and standardization of software ♠**TO DO**: *system?*♠ and databases, to links to publication, data archival and reproducibility of experiments, development models and tools, and social aspects.
The concrete objectives of **MathVRE** are:

Objective 1: To provide interoperable collaborative open source tools for research in Mathematics that can be combined with off-the shelf non-mathematical infrastructure into Virtual Research Environments running on a variety of platforms, including standard e-infrastructures. This fulfills part of Aim 1 and 2.

♠**TO DO**: *Maybe split this into architecture/framework – basically a collection of APIs; and an initial core set of components*♠

♠**TO DO**: *Semantics of APIs*♠

♠**TO DO**: *Urgent talk to EPCC re e-infrastructure standards*♠

♠**TO DO**: *SL Write detailed description*♠

Objective 2: Community building

Bring together the communities (IPython, `Sage`, `Singular`)

♠**TO DO**: *NT expand and write detailed description*♠

Objective 3: Update existing components for seamless deployment and efficient execution on a wide range of computing environments (workstation, HPC, cloud). This fulfills part of Aim 2.

♠**TO DO**: *SL: write detailed description*♠

Objective 4: Explore the social aspects: how do researchers collaborate in Mathematics? What can be the role of Virtual Research Environments? ♠**TO DO**: *UM: write detailed description*♠ ♠**TO DO**: *SL: semantics from mathematical text?*♠

Objective 5: Identify and promote software development best practices that will ensure the long term sustainability of an ecosystem of interoperable open source components, developed by overlapping communities. ♠**TO DO**: *NT: write detailed descriptions*♠

Objective 6: Identify and extend ontologies and standards to allow easy, safe and efficient storage, reuse, interoperation and sharing of rich mathematical data taking account of provenance and citability. ♠**TO DO**: *POD – write detailed description*♠

Objective 7: Demonstrate the effectiveness of Virtual Research Environment built on top of **MathVRE** components for a number of real-world use cases taken from different domains, or crossing previously little connected domains.

Objective 8: Effective Dissemination ♠**TO DO**: *Vivienne – write detailed description*♠

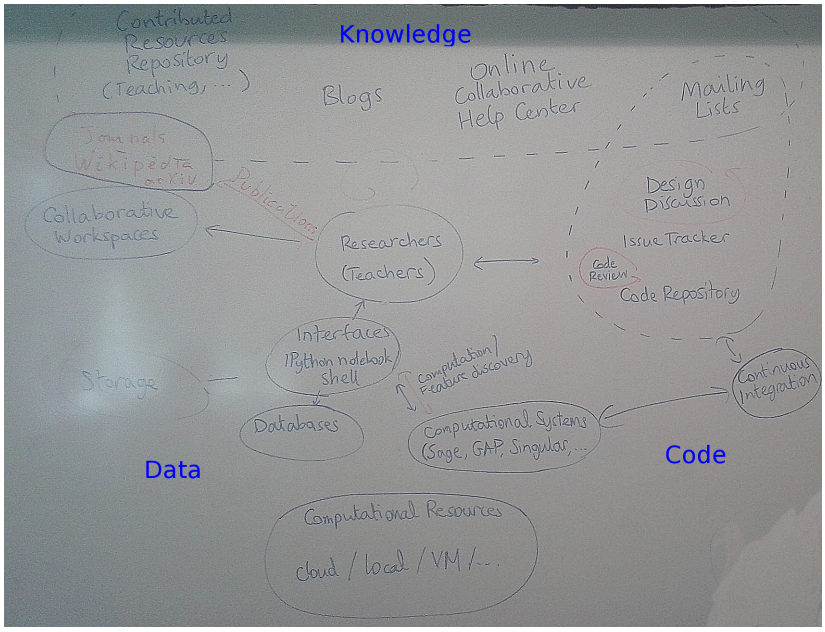
1
♠**TO DO**: *The pieces of material below need to be recombined in a flowing story.*♠

1.1.1 Key ideas

:

- Experimental maths has become a core asset for research in pure mathematics and its applications.
- Over the last decades, mathematicians have gained strong experience in collaborative software development, with pioneering work and continuing leadership of Europe.
- Mathematicians have a strong tradition of sharing knowledge openly (arxiv, Wikipedia, ...).
- Mathematicians have been building and sharing databases for a long while; the needs for such is growing tremendously, and the process needs to be streamlined.

This project gathers European core developers of leading mathematical software (GAP, Pari, Sage, Singular, ...), databases (LMFDB, ...), and critical components (IPython stack), together with researchers in computer and social sciences, with mission to promote a new generation of community-developed open source software ♠**TO DO**: *more precisely what's new is the combination thereof!*♠, databases, and services, adapted to the needs of collaborative research in pure mathematics and its applications.
♠**TO DO**: *Keyword: flexible virtual environment*♠



Our research will cover a wide variety of aspects, ranging from software development models, user interfaces ♠**TO DO**: *virtual environments?*♠ deployment frameworks and novel collaborative tools, component architecture, design, and standardization of software components and databases, to links to publication, data archival and reproducibility of experiments, development models and tools, and social aspects. It will build on the remarkable success of the open source ecosystem and consolidate Europe's leading position in computational mathematics.
Following the call specifications, all software, data, and publications resulting from this proposal will be open.

1.1.2 Why collaborative development of open source software?

From their early days, computers have been used in pure mathematics, either to prove theorems or, like the telescope for astronomers, to explore new theories. Major achievements include the proof of the four color theorem or ♠**TO DO**: *Nice flashy example?*♠. Usage has grown to the point that certain areas of mathematics now completely depend on experimental methods, with major efforts spent on software development. As the sophistication of the required computations increased, supported by the boom of the available computational power, it became vital to share those efforts at the scale of large research communities. European mathematicians have been pioneers and have grown a steady tradition of collaborative open source software development, with systems like GAP, Singular, or Pari/GP playing a major role for decades.

1.1.3 Importance of experimental tools in maths

The field of computer algebra allows us to compute in and with a multitude of mathematical structures. It is interdisciplinary in nature, with links to quite a number of areas in mathematics, with applications in mathematics and other branches of science and engineering, and with constantly new and often surprising developments. Quite a number of these developments, in fact the creation of whole subareas of the field, have been iniated by European researchers who made crucial contributions at all levels. These include the design of fundamental algorithms, the development of major computer algebra systems, applications of the computational methods in various fields, and the creation of widely used databases.

Particular fruitful interactions unfold between computer algebra and algebraic geometry, number theory, and group theory. Algebraic algorithms open up new ways of accessing subareas of these key disciplines of mathematics, and they are fundamental to practical applications of the disciplines. Conversely, challenges arising in algebraic geometry, number theory, and group theory quite often lead to algorithmic breakthroughs which, in turn, open the door for new theoretical and practical applications of computer algebra.

Based on exact computer aided calculations, the experimental method has now been added to the toolbox of the pure mathematician. Experiments lead to new conjectures which may have a deep impact on the future development of mathematics. An outstanding example is the Birch and Swinnerton-Dyer conjecture which is one of the Clay Millenium Problems. Databases relying on computer calculations such as the Small Groups Library or the Modular Atlas in group and representation theory provide indispensible tools for researchers. A constructive way of understanding proofs of deep theorems yields algorithmic tools to deal with highly abstract concepts. These tools make the concepts available to a broader class of researchers, with many potential applications. A prominent example from algebraic geometry is the desingularization theorem of Hironaka, for which Hironaka won the Fields Medal, and its algorithmization by Villamayor.

Spectacular theoretical breakthrougs such as Wiles' proof of Fermat's last theorem are based on interdisciplinary approaches. Current developments on the algorithmic side allow one to conquer crossconnections between different areas of mathematics also computationally and, thus, to arrive at cutting-edge applications which previously were inconceivable.

1.2 Relation to the Work Programme

♠**EC Commentary**: *1-2 pages; Eugenia will help there*♠
♠**EC Commentary**: *Indicate the work programme topic to which your proposal relates, and explain how your proposal addresses the specific challenge and scope of that topic, as set out in the work programme.*♠
♠**EC Commentary**:

```
This is a synthesis of Antonios Barbas' slides describing the Call 3
Topic 9-2015 EINFRA-9: e-Infrastructure for Virtual Research Environment

See file:../Documentation/VirtualEnvironmentsWorkProgramme2014-2015.ppt

** Suggested EU contribution per proposal: 2 to 8 M euros ; Total budget: 42 Meuros
** Dates: 14/01/2015
- H2020-EINFRA-2014-1 15/04/2014
- H2020-EINFRA-2014-2 02/09/2014
- H2020-EINFRA-2015-1 14/01/2015(tbc)
** European contacts: Antonios Barbas
   See file:Documentation/Contacts.docx
** Definition:
- Groups of researchers, typically widely dispersed, who are working
  together
- through ubiquitous, trusted and easy access to services for
  scientific data, computing and networking
- in a collaborative, virtual environment:
  > the e-Infrastructures
** Characteristics:
- Address the needs of specific scientific communities { in support of
  e-Science;
- Have users from both academia and industry;
- Involve bottom-up research and develop user-oriented services;
- Are based on e-infrastructures
** Specific challenge:
- Capacity building in interdisciplinary research
- through community-led development and deployment of service-driven
  digital environments
- for large-scale cross-disciplinary research collaboration and data
  interoperability
** Expected impact:
- More effective collaboration between researchers and increased
  take-up of collaborative research by new disciplines;
- Easier discovery, access and re-use of data, resulting in higher
  productivity of researchers;
- Accelerate innovation via access to integrated digital research
  resources across disciplines;
*** Scope: Proposals are expected to
Notations: [X]: easy to argue; [?]: we have some lead, but that will take some arguing
- [?] Integrate resources across all layers of the e-infrastructure
  (networking, computing, data, software, user interfaces) to foster
  cross-disciplinary data interoperability
- [?] Build on requirements from real use cases, i.e. integrate
  heterogeneous data from multiple sources and re-use tools and
  services from existing infrastructures
- [X] Target any area of Science and Technology, especially
```

```
  interdisciplinary ones, including ICT, mathematics, web science and
  social sciences and humanities
- [X] Use standardised building blocks and workflows, well-documented
  interfaces and interoperable components;
- [?] Define semantics, ontologies and metadata to enable data citation
  and promote data sharing, as to ensure interoperability;
- [X] Target easy-to-use functionalities; and indicate the number of
  researchers they target as potential users;
** Specific conditions for the Call on e-Infrastructures:
- [X?] Proposals should be structured around Networking, Service
  and Joint Research Activities
- [X] The Software to be developed needs to be open source
- [ ] A Data Management Plan to be developed enabling data preservation,
  on-line discoverability, authorisation and re-use of data
- [X] Clear Metrics (KPIs) to be proposed and used;
- [?] Open Access to Publications resulting from the project;
- [X] Usefulness of services to the end user community and
      financial sustainability to be ensured;
** Where should the emphasis be?
- [?] Services
- [X] Thinking innovation
      With both suppliers or users
- [X] Mainstreaming skills development
- [ ] Integration between data and computing
- [X] Business plans for financial sustainability
  ...and partnerships with the private sector
- [ ] Supporting policies
- [X] open data and software
- [X] Sharing basic operations services and building blocks
- [X] Monitoring performance (KPIs)
** Simplified funding model
   - Up to 100% for Research and Innovation
     - Flat 25% rate for indirect costs (overhead?)
```

♠

1.3 Concept and Approach

♠**EC Commentary**: *5-8 pages*♠ ♠**EC Commentary**: *– Describe and explain the overall concept underpinning the project. Describe the main ideas, models or assumptions involved. Identify any trans-disciplinary considerations; – Describe and explain the overall approach and methodology, distinguishing, as appropriate, activities indicated in the relevant section of the work programme, e.g. Networking Activities, Service Activities and Joint Research Activities, as detailed in the Part E of the Specific features for Research Infrastructures of the Horizon 2020 European Research Infrastructures (including e-Infrastructures) Work Programme 2014- 2015;*
– Describe how the Networking Activities will foster a culture of co-operation between the participants and other relevant stakeholders.
– Describe how the Service activities will offer access to state-of-the-art infrastructures, high quality services, and will enable users to conduct excellent research.
– Describe how the Joint Research Activities will contribute to quantitative and qualitative improvements of the services provided by the infrastructures.
– As per Part E of the Work Programme, where relevant, describe how the project will share and use existing basic operations services (e.g. authorisation and accounting systems, service registry, etc.) with other e-infrastructure providers and justify why such services should be (re)developed if they already exist in other e-infrastructures. Describe how the developed services will be discoverable on-line.
– Where relevant, describe how sex and/or gender analysis is taken into account in the project's content.♠

1.3.1 Linked research and innovation activities

♠**EC Commentary**: *Describe any national or international research and innovation activities which will be linked with the project, especially where the outputs from these will feed into the project;*♠
♠**TO DO**: *For each item below, write a paragraph describing the project and one describing how it connects with this proposal*♠

DFG Priority Project SPP 1489 `computeralgebra.de`
♠WD [**WRITE HERE**: Summarize the DFG Priority project description below into a paragraph]♠
The field of computer algebra allows one to compute in and with a multitude of mathematical structures. It is interdisciplinary in nature, with links to quite a number of areas in mathematics, with applications in mathematics and other branches of science, and with constantly new and often surprising developments.
Particular fruitful interactions unfold between computer algebra and algebraic geometry, number theory, and group theory. Algebraic algorithms open up new ways of accessing subareas of these key disciplines of mathematics, and they are fundamental to practical applications of the disciplines. Conversely, challenges arising in algebraic geometry, number theory, and group theory quite often lead to algorithmic breakthroughs which, in turn, open the door for new theoretical and practical applications of computer algebra.
The goal of the DFG Priority Program SPP 1489 is to considerably further the algorithmic and experimental methods in the afore mentioned disciplines, to combine the different methods where needed, and to apply them to central questions in theory and praxis. Moreover, the programme is meant to support the further development of free computer algebra systems which are (co-)based in Germany, and which in the framework of different projects, may require crosslinking on different levels.
Of particular interest are interactions with application areas inside and outside of mathematics such as system- and control theory, coding theory, cryptography, CAD, algebraic combinatorics, and algebraic statistics as well as hybrid methods which combine numerical and symbolic approaches.
♠WD [**WRITE HERE**: One paragraph description of how this relates to this project]♠

IPython/Jupyter grant from the Alfred P. Sloan foundation ♠IPython [**WRITE HERE**: Proofread description of the Sloan grant and link to this project]♠
The IPython project received a $1.15M grant from the Alfred P. Sloan foundation that is supporting IPython development for two years (1/1/2013-12/31/2014), in particular at the University of California, Berkeley and California Polytechnic State University, San Luis Obispo. This grant enabled the project to focus on developing the IPython Notebook as a general tool for scientific and technical computing that is open, collaborative and reproducible. This

goes a long way toward Aim ♠**TO DO**: *... and ...*♠ of **MathVRE**, especially given the current rapid evolution of IPython toward its language agnostic avatar Jupyter.
MathVRE will build on the outcome of the Sloan grant, and further develop the critical IPython/Jupyter component in close collaboration with the IPython/Jupyter team. In particular, we plan to hire some of the European developers that are currently funded by the Sloan grant to work in California and wish to later return to Europe.

Sage-Combinat grant ♠NT [**WRITE HERE**: ...]♠

Logilab: simulagora, cubicweb, ... ♠Logilab [**WRITE HERE**: One paragraph description of simulagora, cubicweb, ...]♠ ♠Logilab [**WRITE HERE**: How does it relate to this project]♠

Sage Math Cloud

FLINT grant?

LMFDB grant The L-functions and Modular Forms Database (LMFDB) project originated at a meeting at The American Institute for Mathematics (AIM) in 2007. L-functions are ubiquitous in number theory, and have applications to mathematical physics and cryptography. The simplest example of an L-functions is the Riemann zeta function. Two of the seven Clay Mathematics Million Dollar Millennium Problems deal with properties of these functions, namely the Riemann Hypothesis and the Birch and Swinnerton-Dyer Conjecture. As well as providing a central repository of data as a resource for researchers, through its website `www.lmfdb.org`, the LMFDB provides a modern handbook, including tables, formulas, links and references, concerning particular specific L-functions and their sources. Between 2008 and 2012 the LMFDB was funded through a US National Science Foundation (NSF) Focussed Research Grant (FRG) of around $1M. Since 2013, the funding of the LMFDB has passed to Europe through a six year £2.2M Programme Grant from the UK Engineering and Physical Sciences Research Council (EPSRC), held at the universities of Warwick and Bristol, with Professor John Cremona (Warwick) as its Principal Investigator (see `http://www2.warwick.ac.uk/fac/sci/maths/people/staff/john_cremona/lmf`). This grant supports six three-year postdoctoral research fellows, mathematical researchers who work on the mathematical aspects of the project full-time, biannual workshops, equipment and a portion of the investigators' own time.
Almost all contributors to the LMFDB project, including those directly supported by the EPSRC grant and the larger world-wide team of 30-50 contributors of data and code, are pure mathematicians. Most of these have good computational skills, but are not professional programmers or software developers. The LMFDB has a great need to broaden the support it can call upon from software developers, to enhance the project in several ways, including the computation of number-theoretic data but more specifically in supporting the database management and website user interface, in order to make the data more accessible and useful to others. The codebase of the LMFDB project is entirely open source and hosted at github (https://github.com/LMFDB/lmfdb), written in python with specialist modules such as flask and pymongo to manage the website and database interface, and Sage for higher-level mathematical computations. The LMFDB project would therefore benefit greatly from collaboration with **MathVRE** as it would connect the project with a pool of experts. Joint workshops between the LMFDB and **MathVRE** will stimulate and develop such collaboration: the LMFDB places great importance on its workshops, which are small gatherings of around 30 invited participants who work throughout one week on certain specific aspects of the project, coming together in plenary sessions to make decisions, plan and collectively approve of proposed developments.

1.4 Ambition

♠**EC Commentary**: *1-2 pages*♠
♠**EC Commentary**: *– Describe the advance your proposal would provide beyond the state-of-the-art, and the extent the proposed work is ambitious. Your answer could refer to the ground-breaking nature of the objectives, concepts involved, issues and problems to be addressed, and approaches and methods to be used.*
– Describe the innovation potential which the proposal represents. Where relevant, refer to products and services already available, e.g. in existing e-Infrastructures.♠

2 Impact

♠**TO DO**: *Orsay's grant services will help here in December*♠

2.1 Expected Impacts

♠**EC Commentary**: *Please be specific, and provide only information that applies to the proposal and its objectives. Wherever possible, use quantified indicators and targets.*
Describe how your project will contribute to:
– the expected impacts set out in the work programme, under the relevant topic (including key performance indicators/metrics for monitoring results and impacts);
– improving innovation capacity and the integration of new knowledge (strengthening the competitiveness and growth of companies by developing innovations meeting the needs of European and global markets; and, where relevant, by delivering such innovations to the markets;
– any other environmental and socially important impacts (if not already covered above).
Describe any barriers/obstacles, and any framework conditions (such as regulation and standards), that may determine whether and to what extent the expected impacts will be achieved. (This should not include any risk factors concerning implementation, as covered in section 3.2.)♠

2.2 Measures to Maximise Impact

2.2.1 Dissemination and Exploitation of Results

♠**EC Commentary**: *– Provide a draft 'plan for the dissemination and exploitation of the project's results'. The plan, which should be proportionate to the scale of the project, should contain measures to be implemented both during and after the project.*
Dissemination and exploitation measures should address the full range of potential users and uses including research, commercial, investment, social, environmental, policy making, setting standards, skills and educational training.
The approach to innovation should be as comprehensive as possible, and must be tailored to the specific technical, market and organisational issues to be addressed
– Explain how the proposed measures will help to achieve the expected impact of the project . Provide a draft business plan for financial sustainability as stated in the Part E of the Specific features for Research Infrastructures of the Horizon 2020 European Research Infrastructures (including e-Infrastructures) Work Programme 2014-2015.
– Where relevant, include information on how the participants will manage the research data generated and/or collected during the project, in particular addressing the following issues: What types of data will the project generate/collect? What standards will be used? How will this data be exploited and/or shared/made accessible for verification and re-use (If data cannot be made available, explain why)? How will this data be curated and preserved?

– Include information about any open source software used or developed by the project.
You will need an appropriate consortium agreement to manage (amongst other things) the ownership and access to key knowledge (IPR, data etc.). Where relevant, these will allow you, collectively and individually, to pursue market opportunities arising from the project's results.
The appropriate structure of the consortium to support exploitation is addressed in section 3.3.

– Outline the strategy for knowledge management and protection. Include measures to provide open access (free on-line access, such as the "green" or "gold" model) to peer-reviewed scientific publications which might result from the project.
Open access publishing (also called 'gold' open access) means that an article is immediately provided in open access mode by the scientific publisher. The associated costs are usually shifted away from readers, and instead (for example) to the university or research institute to which the researcher is affiliated, or to the funding agency supporting the research.
Self-archiving (also called "green" open access) means that the published article or the final peer-reviewed manuscript is archived by the researcher - or a representative - in an online repository before, after or alongside its publication. Access to this article is often - but not necessarily - delayed ("embargo period"), as some scientific publishers may wish to recoup their investment by selling subscriptions and charging pay-per-download/view fees during an exclusivity period.♠

Long term sustainability The success of large specialized software like Pari, Singular or GAP in the last decades has shown the viability of the academic open source development model for such. For a long time, it was bitterly debated whether this model would have any chance to scale to general purpose systems for pure mathematics. The rapid take off of Sage in the last 10 years has proven the viability of the "developed by users for users" model: despite its large community of 300 developers, it's running on a tiny specific budget, with most activities being funded indirectly by research grants that require specific development.
This was made possibly by reusing existing components whenever possible (e.g. hundreds of specialized open source math libraries, or the Python programming language with its developers tools and huge library), and outsourcing software development (e.g. the Cython compiler) to larger communities whenever possible.
♠**TO DO**: *This piece of argument is tricky to setup!!!*♠
Yet, long term critical non mathematical features like portability, modularization, packaging, user interfaces, large data, parallelism, or outreach toward related software, have been lagging behind. Indeed they can hardly be implemented as a side product of research projects, and **need to be assigned to full time developers**. Regular funding is also needed to better structure the computational mathematics community in Europe and support its upcoming major widening through training, development workshops, exchanges, ...
The purpose of this grant is to initiate this process. The principle is that, with the growth of the user base, a tiny

number of institutions or companies will hire a single full-time developer, and this will be sufficient because they critically need it to support their research.
The number of such required full time developers will be made even tinier because most of the efforts now will be focused toward outsourcing more components to reduce the recurrent needs.
For example, this project will save much recurrent efforts to the mathematics community by outsourcing the development of the user interface to IPython. This grant will provide the required temporary boost to make IPython stand to the stringent needs of the community. Later on, thanks to its large user base, both in academia and industry, IPython will continue to thrive without specific funding or major contributions from the mathematics community.
Another big focus of this project will be on the study of open source development models for mathematical software and how they can be made more productive, in particular by better processes and collaboration between components, which will also reduce the number of required full time developers.

2.2.2 Communication activities

♠**EC Commentary**: *Describe the proposed communication measures for promoting the project and its findings during the period of the grant. Where appropriate these measures should include social media and public events with user participation. Measures should be proportionate to the scale of the project, with clear objectives. They should be tailored to the needs of various audiences, including groups beyond the project's own community. Where relevant, include measures for public/societal engagement on issues related to the project.*♠

3 Implementation

♠**TO DO**: *Typical granularity: 5-8 work packages with 3-5 tasks and one deliverable per task; 10 milestones*♠

3.1 Work Plan — Work packages, deliverables and milestones

♠**EC Commentary**: *Please provide the following:*

- *brief presentation of the overall structure of the work plan;*
- *timing of the different work packages and their components (Gantt chart or similar);*
- *detailed work description, i.e.:*
  - *a description of each work package (table 3.1a);*
  - *a list of work packages (table 3.1b);*
  - *a list of major deliverables (table 3.1c);*
- *graphical presentation of the components showing how they inter-relate (Pert chart or similar).*

♠

Overall Structure of the Work Plan

The work plan is broken down into XX workpackages as shown in Figure **??**: WP2 deals with ... In addition, there is one management work package (WP1) and one general dissemination work package (**??**). The Gantt chart on Page 15 illustrates the timeline for the various tasks for these work packages, including inter-task dependencies.

How the Work Packages will Achieve the Project Objectives

♠ALL [**WRITE HERE**: This needs to explain that we're actually going to meet the objectives. Needs to be done after objectives and WPs.]♠
The project objectives (Section 1.1, page 2) and the corresponding work packages that contribute to achieving those objectives are:

Objective	**Purpose**	**WPs**
Objective 1	XX	**WPX**

Work Programme for Objective 1: Objective 1 is covered by WPX, which will ...

Work Plan Timing: GANTT Chart showing Task Dependencies and Information Flows

Work package list

Work package No	Work package title	Lead partic. no.	Lead short name	Person months	Start month	End month
WP1	Project Management	1	**UPS**		1	60
WP2	Community Building and Engagement	1	**UPS**			
WP3	Component Architecture	1	**UPS**			
WP4	User Interfaces	1	**UPS**			
WP5	HPC and massively parallel components	1	**UPS**			
WP6	Next generation Mathematical Databases	2	**USTAN**			
WP7	Development Models for an Academic Free Software Ecosystem					
WP8	Social Aspects	6	**UO**			
WP9	Dissemination, Exploitation and Communication	2	**USTAN**			
Total				**XXX**		

List of Deliverables

Del. no.	Deliverable name	WP no.	Lead	Type	Dissemination level	Delivery date
D8.1	Requirements Analysis	WP?		R	CO	??

List of milestones

Milestone number	**Milestone name**	**Related work package(s)**	**Estimated date**	**Means of verification** (deliverables shown here + success criteria below)
MS1 MS3	Completed initial requirements analysis.	WPX WPX	1	**??**.

Milestone	**Success Criteria**	**Contributes to Objective(s)**
MS1 MS3	Completed requirements analysis (Deliverable **??**). XX	**1, 3.** **XX**

♠**EC Commentary**: *KEY Estimated date Measured in months from the project start date (month 1) Means of verification Show how you will confirm that the milestone has been attained. Refer to indicators if appropriate. For example: a laboratory prototype that is 'up and running'; software released and validated by a user group; field survey complete and data quality validated.* ♠

Work package description (WP1)

Work package number	WP1		**Start date or starting event:**			Month 1	
Work package title	**Project Management**						
Participant number	1	2					
Participant short name	**UPS**	**USTAN**					
Person-months per participant:	48	48					

Objectives: Objectives: The objectives of WP1 are to undertake all project management activities, including:

- monitoring the overall progress of the project and the use of resources;
- ensuring the timely production of deliverables and other project outputs;
- reporting to the European Commission on financial matters;
- preparing for and attending the annual project review meetings; and
- managing the project Advisory Board.

Description of work:

This workpackage will perform all the activities related to monitoring of progress towards the project milestones shown on Page 3.1 and the deliverables listed on Page 3.1, assuring the quality of the deliverables, ensuring the collation and distribution of the required reports, questionnaires and deliverables including the annual reports to the European Commission, arranging project management meetings, tracking the project budget in terms of expenditure and person-months, obtaining financial certificates as required, convening project management meetings, ensuring that important project documents such as the project contract and the consortium agreement are properly maintained and amended as necessary, ensuring that contractual details are complied with, monitoring compliance with the grant agreement, preparing for the annual review meetings, and reviewing research results against the aims and objectives of the project. It also involves managing and supporting the project Advisory Board, including supporting attendance at project meetings, convening Advisory Board meetings, and obtaining feedback on the project direction and results.

Deliverables:

- D1.1 (Month 1): Internal and external mailing lists.
- D1.3 (Month 1): Internal software repository.♠**TO DO**: *Needed?*♠
- D1.4 (Month 12): Project Periodic Report (first year).
- D1.5 (Month 24): Project Periodic Report (second year).
- D1.6 (Month 36): Project Periodic Report (third year).
- D1.7 (Month 48): Project Periodic Report (fourth year).
- D1.8 (Month 48): Project Final Report

Work package description (WP2)

Work package number	WP2		**Start date or starting event:**			Month 1	
Work package title	**Community Building and Engagement**						
Participant number	2						
Participant short name	**USTAN**						
Person-months per participant:	1						

Objectives: The objective of WP2 is to further develop the community at the European scale, foster cross teams collaborations, spread the expertise, and engage the greater community to participate to the definition of the needs, and the implementation and use of the produced solutions.

Description of work:

We will organize regular open workshops (e.g. Sage Days, Pari Days, summer schools, etc.); some of them will be focused on development and coding sprints, and others on training.
This work package will also provide general travel budget to fund short to long term visits between the participants, to collaborate on specific features. A typical such visit would bring together an IPython developer with a GAP developer for a couple of days to implement a first prototype of notebook interface to GAP.
This work package will complement and lean on a parallel COST network whose role is to build and animate the greater community.

Deliverables:

- **??** (Month 12): Report on community needs
- Workshop 1 ...
- Workshop 2 ...
- Workshop 3 ...
- ♠**TO DO**: *make a list*♠

Work package description (WP3)

Work package number	WP3		**Start date or starting event:**			Month 1	
Work package title	**Component Architecture**						
Participant number	1	9	2				
Participant short name	**UPS**	**UVSQ**	**USTAN**				
Person-months per participant:	24	1	1				

Objectives: The objective of this work package is to develop and demonstrate a set of API's enabling components such as database interfaces, computational modules, separate systems such as GAP or Sage to be flexibly combined and run smoothly across a wide range of environments (cloud, local, server, ...).

Description of work:

Task 3.1 ***Portability***
In order to achieve maximum availability and accessibility, mathematical software must be developed and tested for a wide range of computer architectures and operating systems. However most of open source development happens in POSIX environments (usually Linux or OSX), and almost exclusively on x86 platforms. The vast majority of the developers of mathematical software does not have the expertise, nor the access to appropriate hardware and software, to insure appropriate testing and porting of components. The best incarnation of this issue is the involved installation procedure for Sage on Windows, a major adoption barrier and common source of complaints by end-user.
In this task we will address the common needs of the community in terms of portability layers, building and testing infrastructure.

- *Best practices adopted by the larger open source community will be investigated and leveraged, and existing expertise will be shared between the component developers.*
- *Windows being largely dominant in the desktop/laptop market, a specific focus will be placed on the port of* `Sage`*, and therefore all the components included in its distribution (in particular* `Pari`*,* `GAP`*,* `Singular`*,* `Linbox`*) to this platform (D 3.1, D 3.2).*
- *The deployment of a common infrastructure for multi-platform continuous integration (testing, building and distribution) will be addressed.*

Task 3.2 ***Interfaces between systems***
In this task we will investigate patterns to share data, ontologies and semantics across computational systems, possibly connected remotely. We will leverage the well established semantics used in mathematics (categories, type systems, . . .) to give powerful abstractions on computational objects.
Through well defined APIs, we will enable discovery of subsystems, functionality, documentation and computational resources. The user interfaces shall be enabled to automatically choose the best available algorithms and resources to perform a required computation, as well as clearly and intuitively present the available choices to the expert user.
As a concrete example, we would want GAP's categories to be mapped to Sage's categories, so that a handle to a GAP group would automatically appear within Sage like a native Sage group. GAP's documentation for the group shall be accessible through Sage, and the user interface shall be able to automatically map Sage methods to GAP routines whenever this makes sense. When different algorithms are available, some of them coded in Sage, some of them in GAP, the interface shall offer an easily navigable interface for the expert user to choose among them.

Task 3.3 ***Modularization and packaging***

In this task we will investigate best practices for composing, sharing and interfacing computational components and data for connected mathematical systems.

We will start with a comparative study of the practices adopted in various open source projects, both inside and outside of this project. This will include reviewing non-mathematical systems, e.g.: operating systems, platforms, web frameworks, cloud and HPC infrastructures.

We will address the current shortcomings to promote a new generation of mathematical software that is capable of scaling to large code bases, large datasets, and massively distributed infrastructures. This task also needs to consider the results of work package WP8 on social issues regarding distributed development, community management, acknowledging contributions, etc.

As an example, Sage has a long history of integrating and distributing large mathematical libraries/software as a whole, with relatively few attention given to defining and exposing interfaces. Component re-usability is not a main focus for the Sage community, at the same time the non-standard and relatively underused package system discourages writing and maintaining autonomous libraries. These factors have contributed to make the Sage distribution what is usually described as a "monolith" (Sage library code alone, not counting included libraries, makes up for 1.5M lines of code), hard to distribute, to maintain, to port, and to develop with.

On the opposite side, GAP has been distributing community-developed "GAP packages" for a long time, but faces now fragmentation issues, at the code and at the community level. The rudimentary package system adds more technical difficulties to GAP's development model.

Both models reach the limits of their scalability, and a synthesis is very much needed.

Task 3.4 ***Deployment and distribution***

♠***TO DO****:* NT: what did you have in mind?♠

Task 3.5 ***Component architecture for High Performance Computing and Parallelism***

As in all other areas of science, properly supporting massively parallel architecture is a major challenge. Many of the computational components have already gone a long way in this direction, and further work will happen there within WorkPackage T5.1.

In this task we will investigate and implement parallelism-friendly ways of combining components together, so that calling components can benefit from the parallelism features of called components, with self-adaptation to the environment and cooperative sharing of resources. ♠***TO DO****:* deliverables♠

Deliverables:

Deliverable 3.1 *(Month 12) one-click install Sage distribution for Windows with Cygwin 32bits*

Deliverable 3.2 *(Month 24) one-click install Sage distribution for Windows with Cygwin 64bits*

♠**TO DO**: *Make all those deliverables one-liners, and move the material to the task descriptions*♠

- **??** Make sure that Sage and therefore all the components it depends on (including GAP, Linbox, Pari, Singular, ...) have standard packages in the main Linux distributions: Debian/Ubuntu, Redhat, Gentoo, ...

  ♠**TO DO**: *Get feedback from our experts, and make this precise; what can we actually promise to achieve? how much work is this? Do we have personnel for this? There is strong expertise in Logilab with a Debian developer working there; he could advise someone on this. Logilab is interested in this because it's meeting similar issues with some of its clients software like Salomé.*♠

- **??** (Month 12): Creation, deployment, and distribution of preconfigured virtual machines for Pari, Sage, ... as a cloud service, in particular within the StratusLab infrastructure. This includes build bots and test bots for continuous integration over a variety of operating systems.

- **??** Modularization of the Sage distribution

  Separation of the different components of Sage (communication with third-party softwares, build system, Sage native code). This is a prerequisite for easier packaging and integration in standard Linux distributions and lmonade, native integration within the IPython notebook and other interfaces (larchenv, Spyder, ...) and collaboration with sister projects.

- **??** Add support for the SCSCP interface protocol to all relevant components (e.g. Sage, ...). ♠SL/AK [**WRITE HERE**: Brief description of what SCSCP is, reference to previous grant, relevance to the goals of this grant; maybe this should go in the work package description]♠

  SCSCP (Symbolic Computation Software Composability Protocol) is a remote procedure call protocol by which a computer algebra system (CAS) may offer services to a variety of possible clients, including e.g. another CAS running on the same computer system or remotely; another instance of the same CAS (in a parallel computing context); a simplistic SCSCP client (e.g. C/C++/Python/etc. program) with a minimal SCSCP support needed for a particular application; a Web server which passes on the same services as Web services, etc. A distinctive feature of the protocol is that both instructions and data are represented in the OpenMath format (`http://www.openmath.org/`; previously supported by the EU JEM Thematic Network; EU project 24969 "ESPRIT" and other projects); moreover, OpenMath support is not limited by existing official OpenMath content dictionaries - private encodings may be easily embedded into SCSCP messages.

  SCSCP has been developed in the EU FP6 project 26133 "SCIEnce – Symbolic Computation Infrastructure for Europe" (TODO: fix and insert URL) and by now is supported by a number of computer algebra systems, including GAP, Macaulay2, Maple, TRIP and others. To facilitate SCSCP implementations, there are also APIs for Java, C and C++, and a simple Python SCSCP client (for a full overview, see SCIEnce website).

  TODO: more on the relevance to to the goals of this grant. This is useful to exchange information between systems for problems that can not be solved within any single system; for storing and retrieving information (in databases) immediately into the CAS session; for organising distributed computations.

- Some IPython/Jupyter deliverables here. ♠**TO DO**: *review what it can already do in term of choice of computational resource and storage back-end.*♠

- Contribution by Kaiserslautern: libSingular, pySingular?, GAP-Singular, Singular-Sage.

  Moving code from Sage into Singular when relevant

- **??** (Month ...) Configure the components of Sage's distribution (e.g. Atlas, Linbox, GAP, Singular, ...) to be systematically HPC-enabled, and make sure that Sage's calls to such components indeed enable HPC.

- Transparent integration of Ipython capabilities for cluster computing.

- Implementation of a transparent abstraction over mpi.

- Develop or integrate existing solutions for MapReduce operations over big data.

- FLINT development (key component for several systems)?

Raw material:

Component Architecture

Recomputation connection belongs here?

Collaboration with unreliable (or restricted!) networking connections
(peer-to-peer, opportunistic syncing, 3rd world). This is technically
interesting, and gets in support for non-networked working. Not sure
if it belongs here or not.

- Security concerns

Goal: Fostering collaborations/integration between components in an open source ecosystem
==

- How to make systems "cooperate" rather than "predate each other".
- E.g. reduce the version issues

- Foster collaboration with upstream libraries by sharing the
  development and maintenance of the interfaces, typically as
  standalone upstream Python bindings (e.g. py-Singular).

- How to make it easy to develop simultaneously two interdependent
  components (e.g. Sage+Singular)

- Foster communication

- Social aspect:
  Credit, Citations, Recognition, Funding

Documentation system
====================

In which package?

Improvements to Sphinx

Sage heavily customizes the Sphinx documentation system, hacking deep
in it in some cases, with quite some duplication in some cases.
Refactor the whole thing, generalizing and contributing back upstream
as much as possible (e.g. parallel compilation).

Work package description (WP4)

Work package number	WP4		**Start date or starting event:**		Month 1		
Work package title	**User Interfaces**						
Participant number	1	3	2	5	4		
Participant short name	**UPS**	**Logilab**	**USTAN**	**UK**	**UB**		
Person-months per participant:	1	1	1	1	1		

Objectives: The objective of this work package is to provide a modern, robust, and flexible user interface for computation, supporting real-time sharing, integration with collaborative problem-solving, multilingual documents, paper writing and publication, links to databases, etc.

Description of work:

♠**TO DO**: *What is a notebook interface*♠
♠IPython [**WRITE HERE**: improve this draft presentation of IPython]♠
IPython is a leading notebook interface in the world of interactive computations, and use massively by biologists, physicists, ♠**TO DO**: *and outside academia!*♠. Originally tailored for Python, it has been language agnostic, and can communicate through a standardized interface to various computation kernels ♠**TO DO**: *cite a few*♠. It can transparently run kernels locally or remotely (e.g. on the cloud), and has built in support for parallel computing and HPC.

Task 4.6 ***Uniform notebook interface for all interactive components***
In this task, we will implement Jupyter interfaces for the interactive computation components of ***MathVRE****, including GAP, Pari, Sage, and Singular. A first release D 4.3 will focus on basic functionality, and a second release D 4.4 will cover advanced features like 3D graphics or transparent documentation browsing (as live worksheets whenever relevant).*
Sage itself will require a specific treatment as it already has a notebook interface. Its development started about at the same time as the IPython notebook, with similar target features but a different agenda: the Sage notebook had to be available very quickly to solve pressing needs of the Sage community; instead the IPython notebook was to take its time and build robust foundations from the ground up. The two projects have exchanged a lot, and the IPython notebook, which benefits from a much larger user base and thus developer pool, has mostly caught up with the Sage notebook in terms of functionality. It's thus time for the Sage community to outsource this key but non disciplinary component and phase out the Sage notebook in favor of the IPython notebook.
The Sage and IPython convergence D 4.5 will require:

- *Robust migration path and tools for Sage worksheets,*
- *Support for math, 2D, and interactive 3D output.,*
- *Import (and export?) of ReST documents, with full support for Sage's specific roles (math, ...),*
- *Support for remote Sage kernel, typically on the cloud, or running with a different Python version (Sage as a library),*
- *A migration path for interactive widgets implemented with Sage's* `@interact` *functionality.*

Joint meetings and visits between the developers of Jupyter and of the computing components will be a key asset for this task.
♠***TO DO****:* convergence with TeXmacs?♠

Task 4.7 ***Notebook interface usability***
In this task, we will further improve Jupyter as a uniform user interface for interactive computation and database query.
♠*IPython [**WRITE HERE**: Split off the work into an appropriate list of deliverables]*♠

- *Collaborative and shared worksheets,*
- *Multilingual notebooks?*
- *Improved 2D/3D graphics: maybe architecture for integrating VPython, vispy, ...? 3D visualization of algebraic surfaces (in collaboration with Singular)*
- *Native folding support for sections and the like, with mouse and keyboard commands.*
- *Version control*
- *Reproducibility:* **??** *Support for tested notebooks. The writer specifies the expected outputs, e.g. in text format, and can check at any point that the full execution of the notebook yields exactly the expected output, as can be done with e.g. ReST files in Sage: sage −t notebook.rst*
- **??** *(Month ???) Heavyweight (e.g. QT based and not web based) user interface with docking support in the style of* `Spyder` `https://code.google.com/p/spyderlib/`*. Can possibly be implemented by extending the QT IPython console, or by letting* `Spyder` *use the IPython protocol (*♠***TO DO**: If that's not yet the case♠*).*

Task 4.8 ***Dynamic documentation and exploration system***

Introspection has become a critical tool in interactive computation, allowing user to explore on the fly the properties and capabilities of the objects under manipulation. This becomes particularly acute in systems like Sage where large parts of the class hierarchy is built dynamically, and static documentation builders like Sphinx cannot anymore render all the available information.
In this task, we will investigate how to further enhance the user experience. This will include:

- *On the fly generation of Javadoc style documentation, through introspection, allowing e.g. the exploration of the class hierarchy, available methods, etc.*
- ♠*Logilab [**WRITE HERE**: Inclusion of database queries and views]*♠
- **??** *(Month 36) Exploratory support for semantic-aware interactive widgets providing views on objects represented and or in databases*

  Preliminary steps are demonstrated in the `Larch Environment` *project (see demo vidéo on* `http://www.larchenvironment.com/`*) and sage-explorer.*

  Ultimate goal: automatically generated LMFDB-style interfaces.

Whenever possible, those features will be implemented generically for any computation kernel by extending the Jupyter protocol with introspection and documentation queries.

.

Task 4.9 ***Structured documents***
Support for writing interactive structured documents, and in particular papers, books, experimentation log books and reports, presentations, course notes, etc, with the following features:

- *Static printed/PDF/HTML version and interactive version.*
  Achieved by either importing or exporting document files in some standard format (LaTeX, ReST, Markdown, ...).

- *Tests (see above).*
- *Collaborative edition.*
- *Version control.*

♠**TO DO**: *include here everything about this topic in Needs.rst*♠
♠**TO DO**: *Wherever relevant, create tickets with details, and refer to them here.*♠

Deliverables:

Deliverable 4.3 *(Month 12) Basic Jupyter interface for GAP, Pari, Sage, Singular*

Deliverable 4.4 *(Month 12) Full featured Jupyter interface for GAP, Pari, Singular*

Deliverable 4.5 *(Month 12) Sage notebook / IPython notebook convergence*

♠SG/PB [**WRITE HERE**: Evaluation of man.months for the Pythran tasks]♠

Work package description (WP5)

Work package number	WP5		**Start date or starting event:**			Month 1	
Work package title	**HPC and massively parallel components**						
Participant number	1	3	2	5	4	10	
Participant short name	**UPS**	**Logilab**	**USTAN**	**UK**	**UB**	**UJF**	
Person-months per participant:	1	1	1	1	1	1	

Objectives: The objective of this work package is to improve the performance of the computational components of **MathVRE**, in particular on massively parallel architectures. This includes notably:

- Fine grained High Performance Computing on many-cores architectures.
- Coarse grained or embarrassingly parallel computing on grids or on the cloud.
- Compilation of high level interpreted code to optimized parallel native code.
- Develop novel HPC infrastructure in the context of combinatorics.

A key aspect will be to foster further sharing expertise and best practices between computational components.

Description of work:

As in all other areas of science, properly supporting massively parallel architecture is a major challenge. Many of the computational components in **MathVRE** have already gone a long way in this direction. For example, an adaptation of the `GAP` kernel for HPC was developed during the 2009-2013 EPSRC project. The expertise gained there was then transferred to the ongoing `Singular`-HPC project, in particular through the rehiring of one of the developers of `GAP`-HPC.
In this work package, we will build on this momentum to further implement HPC support in the components Tasks 5.10, 5.11, and 5.12.
♠**TO DO**: *transition*♠
Many of the computational components of **MathVRE** use a high level interpreted language for their library. This is notably the case of `Sage`. Performance is achieved by compiling critical sections using the `Cython` `Python`-to-C compiler, to the expense of a lower level implementation. In Tasks 5.14 and 5.15, we will also boost performance by further developing and applying such compilation tools, while keeping a high-level approach.

Task 5.10 ***Pari***
♠*KB [**WRITE HERE**: Task around HPC/parallelism in Pari?]*♠
♠***TO DO**:* deliverable♠

Task 5.11 ***Linbox***
♠*JGD/CP [**WRITE HERE**: Task around HPC/parallelism in Linbox]*♠
♠***TO DO**:* deliverable♠

Task 5.12 ***Singular***
♠*WD [**WRITE HERE**: Task around HPC/parallelism in Singular]*♠
♠***TO DO**:* deliverable♠

Task 5.13 ***HPC infrastructure for combinatorics***
♠*FH [**WRITE HERE**: Task around HPC infrastructure for combinatorics]*♠ ♠***TO DO**:* deliverable♠

♠PB/SG [**WRITE HERE**: Add references to the deliverables from the tasks; see e.g. UserInterfaces.tex]♠

Task 5.14 ***Pythran-Cython convergence***

Pythran is a Python to C++ compiler for a subset of the Python language. It is meant to efficiently compile scientific programs, and takes advantage of multi-cores and SIMD instruction units. Thanks to type inference, it requires little annotations. Its rutime supports a subset of the Numpy package.
Cython is a Python to C compiler that was originally developed for Sage and is now a thriving project of its own. It can handle essentially any Python code, and in particular classes, but relies heavily on annotations for producing optimized code.
Therefore, Pythran and Cython are similar in spirit but have complementary feature sets: Pythran can heavily optimize high level Numpy constructs and Cython has broader Python support. In this task, we will investigate the opportunity and feasibility of a convergence between Cython and Pythran: depending on the code at hand, one strategy or the other would be automatically selected, eventually using Pythran generated called from Cython when relevant. This would result in compiler-runtime cooperation driven by the Cython compiler thanks to part of the Pythran-runtime and the extra typing information provided by Cython. An effort will be made to improve more and more the parallelism in the Pythran runtime.
This work will be achieved through a close collaboration between the Pythran developers hired for ***MathVRE*** *and Cython developers involved in the Sage project. It should quicken Sage execution time at least on numpy centric codes, while not putting an extra burden on the developpers.*

Task 5.15 ***Pythran for Sage and Sage Users***

Currently, Sage doesn't provide facilities to improve user written Python code improvement without the modifications implied by the use of the Cython compiler. As Pythran doesn't need these these rewrite, a notebook interface to compile Pythran compliant code will he added in Sage to improve user kernels.
In a similar perspective, testing and improving the integeration between mpi4py and Pythran could provide an efficient toolchain for HPC while keeping full backward compatibility with pure Python code.
Internally, Sage uses Cython for compiling the critical sections of its libraries. In this task, we will explore opportunities to benefit from Pythran compilation within the Sage library, in particular toward better support for parallelism. A specific challenge is that the Sage library uses quite heavily object-oriented programming.
This task will strongly benefit from Task 5.14, while providing in return a real life large-scale use case for it.
A first step to support object-oriented programming will be to make Pythran type inference more accurate, which will also improve error feedback provided for the user.

Task 5.16 ***Explorative task: Add support for classes in Pythran.***
♠*PB/SG [**WRITE HERE**: Make this into a task?]*♠

Deliverables:

Deliverable 5.6 *(Month XX) Implement Pythran runtime support in Cython toward a unified interface.*

Deliverable 5.7 *(Month XX) Improve Pythran runtime support for parallelism.*

Deliverable 5.8 *(Month XX) Facility to compile Pythran compliant user kernels.*

♠PB/SG [**WRITE HERE**: Make the deliverables more explicit?]♠

Deliverable 5.9 *(Month XX) Test and improve Pythran.*

Deliverable 5.10 *(Month XX) Make Pythran typing better to improve error information.*

Work package description (WP6)

Work package number	WP6		**Start date or starting event:**			Month 1	
Work package title	**Next generation Mathematical Databases**						
Participant number	2						
Participant short name	**USTAN**						
Person-months per participant:	1						

Objectives: The objectives of WP6 is to design interfaces that can be used for a wide range of mathematical data, facilities for accepting contributions while tracking provenance and credit, standard metadata allowing database discovery, versioning allowing stable references, citability, recomputation without copying entire database into VM. Participants: Warwick, Paul Oliver Dehaye, Logilab (SME in Paris)

Description of work:

This workpackage ...

Deliverables:

- Databases
  - Polytopes in Polymake
  - Finite groups (Max)
  - Lattices
- **??** (Month X): Shared persistent memoization library for Python/Sage. Typical use case: A group of collaborators is using intensively a given function (in Sage, or in their private code). They want to memoize the results, as with e.g. Sage's cached_method, but across sessions. They further want to share the underlying growing database between themselves, and maybe eventually publish it.

  Features:
  - Use, further extend, and contribute back to some established (Python?) persistent memoization infrastructure. E.g.
    * `https://pythonhosted.org/joblib/memory.html`
    * `github.com/vivekn/redis-simple-cache`
    * `bitbucket.org/zzzeek/dogpile.cache`
  - Apply not only to user-level functions, but also to lower level functions, e.g. in the Sage library, so that indirect calls to the function also get memoized.
  - Trivial to setup and configure for the end user: in a single line, the user selects an existing function, a backend (with a default value), maybe provide some semantic information, and voilà.
    Typical interface: a decorator to be set on appropriate functions. ♠**TO DO**: *Mock code*♠
  - Trivial to setup and configure for groups of researchers, with a wide range of storage backends (e.g. shared dropbox folder, remote directory, database, git repository, ...).
  - Easy to setup data-bot: e.g. launching a virtual machine that systematically fills up the shared database.
  - Versioning and provenance tracking (user, algorithm, software version, ...), for quality certification, credit, ...
  - Recomputation?
  - Ease of publishing, importing, ...
  - Usual database properties: atomicity, merging (easy since the results are supposed to be immutable: just need to merge the tracking info), alerts in case of divergence.

Work package description (WP7)

Work package number	WP7		**Start date or starting event:**			Month 1	
Work package title	**Development Models for an Academic Free Software Ecosystem**						
Participant number	2						
Participant short name	**USTAN**						
Person-months per participant:	1						

Objectives: The objectives of WP7 are to:

-
-
-
-
-

Description of work:

This workpackage ...

Deliverables:

- **??** (Month X): X.

♠DP/UM [**WRITE HERE**: Workpackage Social Aspects]♠

Work package description (WP8)

Work package number	WP8		**Start date or starting event:**			Month 1	
Work package title	**Social Aspects**						
Participant number	6						
Participant short name	**UO**						
Person-months per participant:	1						

Objectives: The objectives of WP8 are to:

-
-
-
-
- Supporting the Mathematical Process

Description of work:

This workpackage ...

Deliverables:

Deliverable 8.11 *(Month 12) ...*

Work package description (WP9)

Work package number	WP9		**Start date or starting event:**			Month 1	
Work package title	**Dissemination, Exploitation and Communication**						
Participant number	2						
Participant short name	**USTAN**						
Person-months per participant:	1						

Objectives: The objective of this work package is to organize and optimize the communication with the larger community. This includes:

- reviewing emerging technologies;
- disseminating research results to the scientific community;
- ensuring awareness of the results in the user community;
- raising general public awareness of the **MathVRE** project;
- defining individual exploitation plans; and,
- managing existing and new intellectual property.

Description of work:

Dissemination: software, APIs, technologies, research results, ...

Task 9.17 ***Reviewing emerging technologies***
In this task, we will produce periodic reviews of emerging technologies and relevant developments elsewhere, and implications for our plans. This include the review of standard components and service for storage and sharing, computational resources, authentication, package management, etc. This may further include negotiating access or shared development when appropriate. This information will be fed to the other work packages, in particular Work Package WP3 Component Architecture.

Task 9.18 ***Dissemination and Communication activities***

♠***TO DO****:* scale this down as appropriate♠
This task comprises all forms of direct dissemination and public communication activities such as press releases, creation of the project web-site including visitor analysis and monitoring tools (9.13), scientific and technical publications, outreach activities (seminars, keynote talks, media interviews, press releases), pro- motion through social media (e.g. twitter, facebook, linkedin), technical workshop organisation, creation of advertisement materials such as flyers, posters, and electronic feeds as well as their distribution.
At least two press releases will be generated in the course of the project (9.12, **??***), and the project will organise at least one open technical workshop each year.*

Deliverables:

Deliverable 9.12 *(Month 3) Press release announcing start of **MathVRE**.*

Deliverable 9.13 *(Month 3) Project web site*

Deliverable 9.14 *(Month 12) Year 1 report*

Deliverable 9.15 *(Month 24) Year 2 report*

Deliverable 9.16 *(Month 36) Year 3 report*

Deliverable 9.17 *(Month 48) Year 4 report*

Raw material:

- Documentation improvements: overview, cross links, overview of recent improvements
- Thematic tutorials
- Collections of pedagogical documents
  E.g. a complete collection of interactive class notes with computer lab projects for the "Algèbre et Calcul formel" option of the French math aggregation (starting from 2014-2015, only open-source systems will be supported, and Sage is a major player).
- Localization of the Sage user interface and key documents in various European languages.
- Distribution of the documents either in the main distribution of Sage or through the online repository (see collaborative tools).
- Massive online introduction course to Sage, drawing on the sage tutorial/notebooks. Could be "First year Sage course in a box".
- Taking the opportunity of Python courses to propose Sage as a natural extension for mathematics; an example is French's "Classes préparatoires"[1], where Python has been recently selected as the language to learn programming[2].

♠**TO DO**: *Milestones need to be discussed and then described here.*♠

[1] `http://en.wikipedia.org/wiki/Classe_prparatoire_aux_grandes_coles`
[2] See the "Annexe" at `http://www.education.gouv.fr/pid25535/bulletin_officiel.html?cid_bo=71586`

♠**TO DO**: *Check this for any necessary changes.*♠

3.2 Management Structure and Procedures

♠**EC Commentary**: *Will get help from Orsay's grant services*♠
♠**EC Commentary**: *Describe the organisational structure and the decision-making (including a list of milestones (table 3.2a)).*
Explain why the organisational structure and decision-making mechanisms are appropriate to the complexity and scale of the project.
Describe, where relevant, how effective innovation management will be addressed in the management structure and work plan.
Describe any critical risks, relating to project implementation, that the stated project's objectives may not be achieved. Detail any risk mitigation measures. Please provide a table with critical risks identified and mitigating actions (table 3.2b).♠

3.3 Consortium as a Whole

♠**EC Commentary**:

- *Describe the consortium. How will it match the project's objectives? How do the members complement one another (and cover the value chain, where appropriate)? In what way does each of them contribute to the project? How will they be able to work effectively together?*
- *If applicable, describe the industrial/commercial involvement in the project to ensure exploitation of the results and explain why this is consistent with and will help to achieve the specific measures which are proposed for exploitation of the results of the project (see section 2.3).*
- *Other countries: If one or more of the participants requesting EU funding is based in a country that is not automatically eligible for such funding (entities from Member States of the EU, from Associated Countries and from one of the countries in the exhaustive list included in General Annex A of the work programme are automatically eligible for EU funding), explain why the participation of the entity in question is essential to carrying out the project*

♠
♠**TO DO**: *The participants are core developers of the involved components*♠
♠**TO DO**: *Experience in community building and engagement*♠
♠WD [**WRITE HERE**: highlight the existing tight collaborations between the members by select some events from the list on `computeralgebra.de`]♠
♠**TO DO**: *User interfaces: recruitement of IPython developers*♠
♠**TO DO**: *Pythran and HPC: a key asset will be the recruitment of two of the lead developers of the* `Pythran` `Python`-*to-C compiler.*♠
♠JGD/CP [**WRITE HERE**: Linbox: recruitement of previous ANR developper]♠
♠**TO DO**: *Explanation of why we want to include Seattle (sage-math cloud, is a key component; access to IP).*♠

3.4 Resources to be Committed

♠**EC Commentary**: *Will get help from Orsay's grant services*♠
♠**EC Commentary**: *Please provide the following:*

- *a table showing number of person/months required (table 3.4a)*
- *a table showing 'other direct costs' (table 3.4b) for participants where those costs exceed 15% of the personnel costs (according to the budget table in section 3 of the administrative proposal forms)*

♠

Summary of staff effort

♠**EC Commentary**: *Please indicate the number of person/months over the whole duration of the planned work, for each work package, for each participant. Identify the work-package leader for each WP by showing the relevant person-month figure in bold.*♠

♠**TO DO**: *Update this once the list of parthers and the WPs are finalised.*♠

Partic. no.	**Partic. short name**	**Work package** WP1	WP2	WP3	WP4	WP5	WP6	WP7	WP8	WP9	**Total PMs**
1	**UPS**										
2	**USTAN**										
3	**Logilab**										
4	**UB**										
5	**UK**										
6	**UO**										
7	**UW**										
Total PM											

♠**EC Commentary**: *Please complete the table below for each participant if the sum of the costs for' travel', 'equipment', and 'goods and services' exceeds 15budget table in section 3 of the proposal administrative forms).*♠

Other direct cost items

	Cost (€)	**Justification**
Travel		
Equipment		
Other goods and services		
Total		

Management Level Description of Resources and Budget

♠**TO DO**: *This needs to be updated in line with the rest of the project.*♠
The project will employ XX person-months of effort over YY years, comprising ...

♠**EC Commentary**: *This section is not covered by the page limit.*
The information provided here will be used to judge the operational capacity.♠

4 Members of the Consortium

4.1 Participants

♠**EC Commentary**: *Please provide, for each participant, the following (if available):*

- *a description of the legal entity and its main tasks, with an explanation of how its profile matches the tasks in the proposal;*
- *a curriculum vitae or description of the profile of the persons, including their gender, who will be primarily responsible for carrying out the proposed research and/or innovation activities;*
- *a list of up to 5 relevant publications, and/or products, services (including widely-used datasets or software), or other achievements relevant to the call content;*
- *a list of up to 5 relevant previous projects or activities, connected to the subject of this proposal;*
- *a description of any significant infrastructure and/or any major items of technical equipment, relevant to the proposed work;*
- *any other supporting documents specified in the work programme for this call.*

♠

♠SL [**WRITE HERE**: Saint Andrews]♠

Université Paris Saclay

University Paris-Sud is among the 40 top universities worldwide in the 2013 Shanghai ranking, and is one of the two best French research universities. With about 27000 students, 1800 permanent teaching staff and 1300 permanent research scientists from national research organisations (CNRS, Inserm, INRA), it is the largest campus in France. Since 2006, scientists from the University were awarded two Fields medals, one Nobel Prize and a number of other international (European Inventor Award 2013, Wolf Prize 2010, Holweck Prize 2009, Japan prize 2007) and national prizes. The Université Paris-Sud has a complete array of competences, ranging from the purest of exact sciences to clinical practices in medicine, covering life and health sciences, legal sciences and economics. Research at the Université Paris-Sud, an essential part of academic understanding, is complemented by research activities with a high valorisation potential. Research contracts and partnership with companies make the Université Paris-Sud a key actor and a major player in French research. The University is located close to the Plateau de Saclay, the largest cluster of public and private R&D institutions in France (with ca. 16000 research staff), and is one of the core members of the University Paris Saclay – a world class university and a world-renowned research and innovation hub.
In the context of this project, the Université Paris Saclay is the home of one of the largest group of Sage developers worldwide. It's a member of the Open Source Thematic Group of the Systematic Paris Region Systems and ICT Cluster. The University also hosts a major research group working on proof assistants (Coq), which naturally opens the door for reaching toward this neighbor community.

Curriculum vitae of the investigators

Florent Hivert Professor at the Laboratoire de Recherche en Informatique, Florent Hivert is a senior researcher in Algebraic Combinatorics with 29 papers in international journals and 15 communications in international conferences. With 100 tickets (co)authored and as many refereed, Hivert is himself a core Sage developer, with contributions including key components of the Sage infrastructure (documentation, automated test, combinatorics infrastructure, paralellism, ...), specialized research libraries.

First name Last name Maître de Conférences at the Laboratoire de Recherche en Informatique, Viviane Pons is a young researcher in Algebraic Combinatorics with 3 papers in international journals and 3 communications in international conferences.
She discovered Sage during her first sage-days in 2010 and has since been an active user and contributor with 10 (co)authored tickets. She is very involved in the promotion of Sage, participating in Sage-Days and proposing Sage introduction tutorials or Sage presentations in various conferences.

Nicolas M. Thiéry Professor at the Laboratoire de Recherche en Informatique, Nicolas M. Thiéry is a senior researcher in Algebraic Combinatorics with 15 papers published in international journals. Among other things, he is a member of the permanent committee of FPSAC, the main international conference of the domain, and has collaborators in Canada, India, and in the US where he spent three years (Colorado School of Mines, UC Davis); he also coorganized fourteen international workshops, in particular Sage Days, and the semester long program on "Automorphic Forms, Combinatorial Representation Theory and Multiple Dirichlet Series" hosted in Providence (RI, USA) by the Institute for Computational and Experimental Research in Mathematics.
Algebraic combinatorics is a field at the frontier between mathematics and computer science, with heavy needs for computer exploration. Pioneer in community-developed open source software for research in this field, Thiéry founded in 2000 the Sage-Combinat software project; with 50 researchers in Europe and abroad, this project has grown under his leadership to be one of the largest organized community of Sage developers, gaining a leading position in its field, and making a major impact on one hundred publications[3]. Along the way, he coauthored part of the proposal for NSF Sage-Combinat grant OCI-1147247.
With 150 tickets (co)authored and as many refereed, Thiéry is himself a core Sage developer, with contributions including key components of the Sage infrastructure (e.g. categories), specialized research libraries (e.g. root systems), thematic tutorials, and two chapters of the book "Calcul Mathématique avec Sage".

[3] `http://sagemath.org/library-publications-combinat.html`, `http://sagemath.org/library-publications-mupad.html`

Publications, achievements

♠**TO DO**: *Il faut être plus formel dans la description des projets antérieurs : Acronyme, titre, agence de financement, durée. Pareil pour les publi - auteurs, titre exact, année etc.*♠

1. Lead of the Sage-Combinat software project.
2. Coauthoring of the open source book "Calcul Mathématique avec Sage", the first of its kind comprehensive introduction to computational mathematics in Sage for education.
3. XXX tickets contributed to Sage.

Previous projects or activities

1. Home of six one week-long Sage Days workshops.
2. Co-Organizer of ♠**TO DO**: *XXX*♠ Sage Days.
3. Founder and regular organizer of a bimonthly Sage User Group meeting in the greater Paris area.
4. Expertise exchanges with Logilab
5. ♠**TO DO**: *XXX*♠

Significant infrastructure

The Université Paris Sud hosts the lead developers of the open source cloud infrastructure `Stratuslab` and its reference infrastructure (♠**TO DO**: *XXX cores*♠). The participants are regular users of this infrastructure, and in close contact with the developers.

4.1.1 University of Kaiserslautern

Principal investigator Prof. Dr. Wolfram Decker

Personal Data

Gender:	male
Nationality:	German
Address:	Department of Mathematics Technical University of Kaiserslautern P.O. Box 3049 Germany
Phone:	+49 631 205 2253
Email:	decker@mathematik.uni-kl.de
URL:	www.mathematik.uni-kl.de/agag/mitglieder/professoren/ prof-dr-wolfram-decker/
Status:	Professor

Scientific Qualification

PhD 1984	University of Kaiserslautern
Habilitation 1989	University of Kaiserslautern

Academic Career

1984 – 1990	Assistant Professor, University of Kaiserslautern
1990 – 2009	C3 Professor, University of the Saarland
2009 – present	W3 Professor, Technical University (TU) Kaiserslautern

Synergistic Activities

PhD students

Hirotachi Abo, Holger Crñi, Hiep Dang, Hanieh Keneshlou, Dereje Kifle, Michael Messollen, Ngoc Anh Pham, Sorin Popescu, Andreas Steenpaß, Isabel Stenger, I. Made Sulandra, Shrawan Tiwari

University Service

1994–1996	Dean, Dep. of Math., University of the Saarland
2014–present	Dean, Dep. of Math., TU Kaiserslautern

Scientific Service

1996–1999	Coordinator of EuroProj (a European algebraic geometry network)
2000–2004	Chair of the programme management committee of EAGER (a European algebraic geometry network)
2010–present	Coordinator of the DFG Priority Programme SPP1489 'Algorithmic and experimental methods in algebra, geometry, and number theory'

Conferences (co)organized

1997– 2004	About 30 conferences, summer schools, and workshops in the framework of EuroProj and EAGER.
1992–present	More than 20 coding sprints, conferences, summer schools, and workshops outside EuroProj and EAGER, including 3 conferences at Dagstuhl and 1 at Banff.
2000	Chair of the Minisymposium on computer algebra, third ECM.

Selected Grants

1986-1987	NATO-Grant of the DAAD (visit UC Berkeley)
1987–1994	In: DFG Priority Programme 'Complex manifolds'
1993	Grant: Japanese Society for the Promotion of Science (visit Kyoto)
1992–1997	In: DFG Priority Programme 'Algorithmic number theory and algebra'
2002–2006	In: DFG Priority Programme 'Global methods in complex geometry'
2010–present	In: DFG Priority Programme SPP1489 (two grants)
1997–2004	Seven grants for EU Highlevel Scientific Conferences

Selected Publications

1. On the uniqueness of the Horrocks-Mumford-bundle (with F.-O. Schreyer). *Math. Ann.* **273** (1986), 415–443. (MR0824431)
2. Stable rank 2 vector bundles with Chern-classes $c_1 = -1$, $c_2 = 4$. *Math. Ann.* **275** (1986), 481—500. (MR0858291)
3. Construction of Surfaces in $\mathbb{P}_4$ (with L. Ein, F.-O. Schreyer). *J. Algebraic Geometry* **2** (1993), 185–237. (MR0858291)
4. Computational algebraic geometry today. In: *Applications of algebraic geometry to coding theory, physics and computation (Eilat, 2001)*, 65–119, NATO Sci. Ser. II Math. Phys. Chem., 36, Kluwer Acad. Publ., Dordrecht, 2001. (MR1866896)
5. Computing in Algebraic Geometry. A Quick Start using SINGULAR (with C. Lossen). *Algorithms and Computation in Mathematics, 16*. Springer, Berlin, 2006. xvi+327 pp. (MR2220403)
6. Parallel algorithms for normalization (with J.Böhm, S. Laplagne, G. Pfister, A. Steenpaß, S. Steidel). *J. Symbolic Comput.* **51** (2013), 99–114. (MR3005784)
7. A first course in computational algebraic geometry (with G. Pfister). *African Institute of Mathematics (AIMS) Library Series*. Cambridge University Press, Cambridge, 2013. viii+118 pp. (MR3052757)
8. Local analysis of Grauert-Remmert-type normalization algorithms (with J.Böhm, M. Schulze). *Internat. J. Algebra Comput.* **24** (2014), 69–94. (MR3189667)

Selected Mathematical Software

1997–present	Coauthor of SINGULAR libraries for adjoint ideals, absolute factorization, integral bases, invariant theory, parametrization of rational curves, primary decomposition, normalization, and sheaf cohomology
2009–present	Head of the SINGULAR developers group

Publications, products, achievements

1. ♠WD [**WRITE HERE**: ...]♠

Previous projects or activities

1. ♠WD [**WRITE HERE**: ...]♠

Significant infrastructure ♠WD [**WRITE HERE**: ...]♠

University of Oxford

♠DP/UM [**WRITE HERE**: Description of the university of Oxford]♠

Curriculum vitae

♠UM [**WRITE HERE**: CV Ursula Martin]♠ ♠DP [**WRITE HERE**: CV Dima Pasechnik]♠

Publications, products, achievements

1. ♠DP/UM [**WRITE HERE**: Publications Oxford]♠

Previous projects or activities

1. ♠DP/UM [**WRITE HERE**: Projects and activities in Oxford]♠

Significant infrastructure

♠DP/UM [**WRITE HERE**: Significant infrastructure in Oxford]♠

Université de Bordeaux

♠VD [**WRITE HERE**: Bordeaux]♠

Curriculum vitae

Publications, products, achievements

1. ♠VD [**WRITE HERE**: ...]♠

Previous projects or activities

1. ♠VD [**WRITE HERE**: ...]♠

Significant infrastructure

♠VD [**WRITE HERE**: ...]♠

Université de Grenoble

♠JGD [**WRITE HERE**: Description of UJF]♠

Curriculum vitae

Jean-Guillaume Dumas Professor at the Laboratoire Jean Kuntzmann, Jean-Guillaume Dumas is a senior researcher in Computer Algebra with 40 papers published in international journals or refereed international conferences. Among other things, he is vice-president of ACM Special interest group on symbolic and algebraic manipulations (SIGSAM), department chair within his Laboratoire (6 research teams, 130 members) and has collaborators in USA, Canada, Ireland, Germany and Luxembourg; he has also co-organized fifteen international conferences.
Computer Algebra is a field at the frontier between mathematics and computer science, with heavy needs for computer exploration. Jean-Guillaume Dumas is the main developer of the LinBox and Givaro C++ libraries (libgivaro1, libgivaro-dev, libgivaro-doc, liblinbox0, liblinbox-dev in Debian) used, e.g., by Sage respectively as its exact linear algebra and its finite fields.
Along the way, he coauthored part of the proposal for NSF-INRIA grant QOLAPS on Quantfier elimination, Optimization, Linear Algebra and Polynomial Systems and he is the director of the French ANR program on High-Performance Algebraic Computations.
♠CP [**WRITE HERE**: CV Clement Pernet]♠ ♠PB [**WRITE HERE**: CV Pierrick Brunet]♠

Publications, products, achievements

1. ♠JGD/CP [**WRITE HERE**: ...]♠
2. Coauthoring of the open source book “Calcul Mathématique avec Sage”, the first of its kind comprehensive introduction to computational mathematics in Sage for education.

Previous projects or activities

1. Direction of the ANR program on High-Performance Algebraic Computations 2012-2015.

Significant infrastructure

♠JGD/CP [**WRITE HERE**: Significant infrastructure in Grenoble (or remove section)]♠

Logilab

Logilab (`http://www.logilab.fr/`) is a french SME focused on using the web and free software to help scientists. It has been in business since 2000 and counts over 20 engineers and PhDs proficient in software engineering, knowledge representation, design and management of IT infrastructure, etc.
Logilab invests 15% of its turnover in research and development and has been part of several R&D projects at the national and european levels, always to provide technical expertise and support to the other partners.
In the context of this project, Logilab will innovate to support the partners with tools and infrastructure, including open databases to flexibly store mathematical objects, user interfaces to visualize complex mathematical properties, fluid workflow tools to ease large-scale collaboration, etc.
Logilab's PIC number is 948455525.

Curriculum vitae

Publications, products, achievements

1. ♠XXX [**WRITE HERE**: ...]♠
2. CubicWeb (mention of prize at DataConnexion#3)
3. publi Brainomics

Previous projects or activities

1. ♠XXX [**WRITE HERE**: ...]♠
2. ASWAD (eu)
3. KIDDANET (eu)
4. PYPY (eu)
5. OpenHPC (fr/FUI)
6. BRAINOMICS (fr/ANR)

Significant infrastructure

♠XXX [**WRITE HERE**: ...]♠

University of Warwick

The Mathematics Institute at the University of Warwick was ranked 23rd worldwide in the 2013 QS world university subject rankings. In 2014 one of its members, Regius Professor Martin Hairer, was awarded a Fields Medal. The Number Theory research group at Warwick was started only in 2006, but has rapidly risen to international status and one of the largest and most vibrant groups in Europe, comprising 25 members (professors, lecturers, postdoctoral researchers and early stage researchers). Mathematics and Statistics at Warwick currently hold £35.8M in research grants from EPSRC (the next highest in the UK being Cambridge at £22.8M and Oxford at £24.2M). [ERC? at least one, check for more.]

In 2013 Professors John Cremona and Samir Siksek, together with co-investigators at Bristol, were awarded a six-year major grant of £2.2M from the UK Engineering and Physical Sciences Research Council (EPSRC) to support the L-functions and Modular Forms Database (LMFDB) project. This grant funds three postdoctoral researchers at Warwick, computer equipment to host its database and website, and regular LMFDB workshops. Several members of the Number Theory group at Warwick are developers for Sage, notably John Cremona who has contributed thousands of lines of code to Sage since 2006 both through his `eclib` C++ library and through original `Python` code which forms part of the `Sage` library.

Each year Warwick hosts a year-long Warwick EPSRC Symposium focussing on one area of mathematical research. The 2012-13 Number Theory Symposium included six research workshops and a summer school "Number Theory for Cryptography" and raised the international profile of the number theory group substantially.

Curriculum vitae

Professor John Cremona. DPhil (Oxford, 1981) under Birch. Previous posts: Michigan, Dartmouth (US), Exeter, and Nottingham (as chair and Head of Pure Mathematics). Cremona has around 50 publications, including a book and papers in Compositio and Crelle. He has held grants from EPSRC and other UK sources worth £2.5M as well as €2.5m from the EU for Marie-Curie Research Training Networks in 2000-2004 and 2006-2010. He was a Scientist in Charge of one of twelve teams in both of these networks, and leader of the research project "Effective Cohomology Computations" in the second. He has been on the Scientific Committee of 30 international conferences (including several Sage Days), and given many invited lecture series. He co-organised semester-long research programmes at IHP Paris (2004) and MSRI (2011, joint with CI Conrey). He has been an editor for five journals. He has supervised 16 PhD students, a dozen Masters students, two EU-funded postdoctoral fellows and currently has three EPSRC-funded postdoctoral research assistants. Cremona has given over 30 invited conference addresses and seminars in 9 countries in the last 10 years.

Cremona's research includes areas of particular relevance to the current project. His methods for systematically enumerating elliptic curves, which are the subject of a book and numerous papers, have been used to compile a definitive database of elliptic curves which is very widely cited, and now forms part of the LMFDB. Cremona's experience in managing such computations and the management, publication and electronic dissemination of the resulting large datasets set a standard which large-scale number-theoretical database projects such as the LMFDB now seek to match. Cremona's experience and reputation in this field has been vital for the management and success of the LMFDB project.

Cremona has been the leading computational number theorist in the UK since his PhD thesis in 1981, following in the tradition of Birch and Swinnerton-Dyer. He has written thousands of lines of code in his C++ library eclib (one of the standard packages included in Sage since its inception) which includes his widely-use program `mwrank` for computing ranks of elliptic curves. As well as writing thousands of lines of new python code for Sage, he has also contributed to the active number-theoretical packages Pari/GP and Magma.

Publications, products, achievements

1. ♠XXX **[WRITE HERE**: ...]♠

Previous projects or activities

1. ♠XXX **[WRITE HERE**: ...]♠

Significant infrastructure

Computing infrastructure available to the group is excellent, with seven dedicated machines (over 300 cores) as well as access through Warwick's Centre for Scientific Computing which hosts a 6000-core linux cluster and a 3500-core cluster of workstations.

♠Logilab [**WRITE HERE**: Description of Logilab]♠ ♠MK [**WRITE HERE**: Description of University of Silesia]♠ ♠SL/WS [**WRITE HERE**: Description of Seattle]♠ ♠LD [**WRITE HERE**: Description of Versailles]♠ ♠Logilab [**WRITE HERE**: PIC for Logilab is 948455525]♠ ♠MK [**WRITE HERE**: PIC for University of Silesia]♠ ♠LD [**WRITE HERE**: PIC for Versailles]♠
[**?**]

4.2 Third Parties Involved in the Project (including use of third party resources)

♠**EC Commentary**: *Please complete, for each participant, the table (see page 27 of "VRETemplate.PDF"), or simply state "No third parties involved", if applicable.*♠
No third parties involved.
♠**TO DO**: *Or Seattle?*♠

5 Ethics and Security

♠**EC Commentary**: *This section is not covered by the page limit.*♠

5.1 Ethics

♠**EC Commentary**: *If you have entered any ethics issues in the ethical issue table in the administrative proposal forms, you must:*
• submit an ethics self-assessment, which:
– describes how the proposal meets the national legal and ethical requirements of the country or countries where the tasks raising ethical issues are to be carried out;
– explains in detail how you intend to address the issues in the ethical issues table, in particular as regards: research objectives (e.g. study of vulnerable populations, dual use, etc.), research methodology (e.g. clinical trials, involvement of children and related consent procedures, protection of any data collected, etc.), the potential impact of the research (e.g. dual use issues, environmental damage, stigmatisation of particular social groups, political or financial retaliation, benefit-sharing, malevolent use , etc.)
• provide the documents that you need under national law(if you already have them), e.g.:
– an ethics committee opinion;
– the document notifying activities raising ethical issues or authorising such activities
If these documents are not in English, you must also submit an English summary of them (containing, if available, the conclusions of the committee or authority concerned).
If you plan to request these documents specifically for the project you are proposing, your request must contain an explicit reference to the project title♠

5.2 Security

Please indicate if your proposal will involve:

- activities or results raising security issues: NO
- 'EU-classified information' as background or results: NO

References