Small or Medium-Scale Focused Research Project (STREP)

ICT Call 1

FP7-???-200?-?

DreamKit: Digital research environment for a* mathematics, as a kit

Acronym: DreamKit

Small or Medium-Scale Focused Research Project (STREP)
Date of Preparation: December 18, 2014
Revision: 000243fof 2014-10-28

Work program topics addressed by DreamKit: Challenge 4: ICT for EU Proposals, **Objective ICT-2012.4.4**: Technology-enhanced Documents, **target outcome b1**) More time for Research, not Proposal writing.

Coordinator: Michael Kohlhase

e-mail: m.kohlhase@jacobs-university.de

tel/fax: +49 421 200 3140/493140

#	Participant organisation name	Short name	Country
1	Universit Paris Sud	UPS	
2	Logilab	Logilab	
3	Universit de Versailles Saint-Quentin	UVSQ	
4	Universit Joseph Fourier	UJF	
5	Universit Bordeaux	UB	
6	University of Oxford	UO	
7	Universit of Sheffield	USHEF	
8	Universit of Southampton	USO	
9	University of St Andrews	USTAN	
10	University of Warwick	UW	
11	Jacobs University Bremen	JU	
12	University of Kaiserslautern	UK	
13	University of Silesia	US	
14	Universität Zürich	UZH	
15	Simula Research Laboratory	Simula	
16	University of Washington at Seattle	UWS	

Abstract

[1] EdN:1

[1] EdNote: *write something*

Contents

B.0.1 Excellence . . . 2
B.0.1.1 Objectives . . . 3
B.0.1.2 Relation to the Work Programme . . . 7
B.0.1.3 Concept and Approach . . . 9
B.0.1.4 Ambition . . . 16
B.0.2 Impact . . . 17
B.0.2.1 Expected Impacts . . . 17
B.0.2.2 Measures to Maximise Impact . . . 18
B.0.3 Implementation . . . 20
B.0.3.1 Work Plan — Work packages, deliverables and milestones . . . 20
B.0.3.2 Work Package Descriptions . . . 23
B.0.3.3 Management Structure and Procedures . . . 25
B.0.3.4 Consortium as a Whole . . . 26
B.0.3.5 Resources to be Committed . . . 27
B.0.4 Members of the Consortium . . . 28
B.0.4.1 Participants . . . 28
B.0.4.2 Third Parties Involved in the Project (including use of third party resources) . . . 50
B.0.5 Ethics and Security . . . 51
B.0.5.1 Ethics . . . 51
B.0.5.2 Security . . . 51

B.0.1 Excellence

MathVRE focuses on delivering easy to setup Virtual Research Environments, customizable to meet the diverse needs of collaborative research in pure mathematics and applications, and built out of a sustainable ecosystem of community-developed open software, databases, and services.

The remarkable emergence and success in the last decade(s) of a large ecosystem of community-developed computational software like e.g. `GAP`, `LinBox`, `PARI/GP`, `Sage`, or `Singular` and of interactive scientific computing environments like `IPython` showcase the viability of community development models for such components.

The recent successes of the Virtual Research and Teaching Environment `SageMathCloud` (hosting more than 10k users and 100k projects after just one year) and of the online number theory database LMFDB showcase the strong need for **integrated solutions** enabling large-scale collaboration on **software**, **knowledge**, and **data**.

Yet setting up infrastructures like those mentioned above currently requires massive ad-hoc efforts. Challenges include portability, compatibility, performance, usability, reproducibility, not to mention the many social aspects involved in communities and ecosystems thereof.

♠**TO DO**: *Better explain the specific needs in maths, and why we focus on VRE building blocks rather than VRE themselves, in the DIY spirit.*♠

A specific challenge in mathematics comes from the vast yet tightly connected array of concepts involved. Different groups of researchers may have radically different needs, workflows, and resources, calling for a highly modular and customizable VRE infrastructure.

MathVRE will attack all those challenges upfront, while consolidating Europe's leading position in this field.

B.0.1.1 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).♠

The specific aims of **MathVRE** are:

Aim 1: Improve the productivity of researchers in pure mathematics and applications by promoting collaborations based on mathematical **software**, **data**, and **knowledge**.

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

Aim 3: Identify and promote best practices in computational mathematical research, including making results easily reproducible, making software sustainable, reusable and easily accessible and sharing data in a semantically sound way. ♠**TO DO**: *insert the* ecosystem *keyword*♠

Aim 4: Maximize sustainability as well as impact in mathematics, neighbor fields, and scientific computing.

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 develop and standardise an architecture allowing a range of mathematical and data and software components to be combined with off-the-shelf non-mathematical infrastructure to produce specialised VREs for different research communities. The architecture will take the form of standards documents and APIs equipped, where appropriate, with formal or informal mathematical semantics to ensure interactions are mathematically sound. This primarily addresses aim 2, thereby contributing to aims 1 and 3. ♠**TO DO**: *This is a bit long*♠

Objective 2: To develop open source core components for such VREs where existing software is not available. These components should support VREs running on a variety of platforms, including standard e-infrastructures. This primarily addresses Aim 2, thereby contributing to Aim 1 and 3.

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

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

Objective 3: Community building

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

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

Objective 4: Update a range of existing open source mathematical software systems for seamless deployment and efficient execution within the VRE architecture of objective 1. This fulfills part of Aim 2.

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

Objective 5: Foster a sustainable ecosystem of interoperable open source components developed by overlapping communities, in particular by identifying and promoting software development best practices, and outsourcing development to larger communities whenever suitable. This fulfills part of Aim 3 and 4.

Objective 6: Explore the social aspects: how do researchers collaborate in Mathematics? What can be the role of Virtual Research Environments? This addresses part of Aim 3 and 2. ♠**TO DO**: *UM: write detailed description*♠ ♠**TO DO**: *SL: semantics from mathematical text?*♠

Objective 7: 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. This fulfills parts of Aim 2 and 3. ♠**TO DO**: *POD – write detailed description*♠

Objective 8: 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. This addresses part of Aim 2.

Objective 9: Effective Dissemination ♠**TO DO**: *VP – write detailed description*♠

Detailed Descriptions of Objectives

Objective 1: Virtual Research Environment Kit

Computational techniques have become a core asset for research in pure mathematics and its applications in the last decades. Mathematics communities have come together to develop powerful computational tools, such as GAP, `PARI/GP`, SAGE or Singular, and valuable on-line services such as the Encyclopedia of Integer Sequences and the ATLAS of Group Representations. ♠**TO DO**: *cite*♠ In building these systems, mathematicians have gained strong experience in collaborative software development, with pioneering work and continuing leadership of Europe.

A number of approaches to linking these resources have been developed, such as the SCSCP protocol from the Framework 6 SCIEnce project, and the incorporation of a variety of free software tools in the SAGE system, but the overall model is still that of a single mathematician running programmes or interacting with a "notebook" page. The software provides little or no support for other aspects of mathematical research: collaboration, archival, reproducibility or linkage between programmes, data and publication. Databases are updated mainly by mathematicians directly, retaining no record of the source of new entries, and providing no way of refering to the actual version of the data used in a particular computation.

The first objective of this project is to design an architecture which will allow existing mathematical software systems (suitably updated), off-the shelf non-mathematical tools and a small number of new components to be flexibly combined to produce a VRE that will effectively support collaborative mathematical research throughout it's entire life-cycle. This will include software APIs and standards, frameworks for assuring the semantic consistency of similar mathematical objects in different systems. It will be informed by the outputs of objective 6, ensuring that the VREs fit the ways that mathematicians actually work.

Objective 2:♠**TO DO: *title, CORE***♠ ♠**TO DO**: *SL Write detailed description*♠

Most of the direct mathematical capabilities of our software will come from existing open source mathematical systems. For instance we will use the power of the GAP Library for computational group theory or `PARI/GP` for number theory. Generic services such as storage, version control, authentication and resource accounting will come from off-the-shelf components building on standard infrastructures.

Nevertheless some new tools will be needed ♠**TO DO**: *what? or at least examples*♠
♠**TO DO**: *Keyword: flexible/versatile 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.

Objective 3: ♠**TO DO: *TITLE: COMMUNITY***♠ ♠**TO DO**: *NT detailed description*♠

Objective 4: ♠TO DO: *TITLE: UPDATES*♠ ♠**TO DO**: *SL detailed description*♠

Objective 5: A sustainable ecosystem of software components

The success of large specialized software like `PARI/GP`, `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 takeoff of Sage in the last decade has proven the viability of the "developed by users for users" model: despite its large international community of about 150 active 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 possible 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), by spinning off software development (e.g. the `Cython` compiler) to larger communities whenever possible, and by carefully designing the development workflow.

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 a few 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, ...

One purpose of this grant is to initiate this process and invest on the long run to reduce the recurrent needs.

The principle is that, with the growth of the user base, a tiny number of institutions or companies will hire a full-time developer to support critical needs of their in-house research or development. Opportunities for such hiring are for example actively investigated at the Laboratoire de Recherche en Informatique. It should be noted that, at the scale of a large university, the cost of licenses for analogous commercial software can reach the same order of magnitude as that of a developer.

To reduce the number of required full time developers **MathVRE** will invest toward, factoring out joint needs, and outsourcing or spinning off more components to larger communities. For example, **MathVRE** will save much recurrent efforts to the mathematics community by providing a temporary boost to outsource the development of the user interface of each computational component to `IPython` and 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.

MathVRE will also foster the productivity within the ecosystem by investigating better collaboration processes between components, and identifying, sharing, and promoting software development best practices.

Objective 6: ♠TO DO: *TITLE: SOCIAL*♠ ♠**TO DO**: *UM detailed description*♠

Objective 7: Next generation mathematical databases ♠**TO DO**: *POD detailed description*♠

Objective 8: ♠TO DO: *TITLE: DEMO*♠ ♠**TO DO**: *detailed description*♠

Objective 9: ♠TO DO: *Title: disseminate*♠ ♠**TO DO**: *VP – write detailed description*♠

The success of any research software or service is strongly related to its ability to attract and convince a great amount of users. Our different communities (Sage, Gap, `PARI/GP`, Singular...) have developed a solid experience and network. As an example, Sage has gathered thousands of users in less than 10 years. This was achieved thanks to a very strong community building philosophy, especially through the organization of Sage-Days all over the world. The first Sage-days happened in 2006, today we count 63 of them, including 10 for 2014 only, and also Sage Education days, Sage Bug days, Sage Doc days and more. Most of us, **MathVRE** project members, have been involved in these events either as organizers or participants and it appears as the most efficient way to promote our software. More precisely, our objective is to create

constant dialogue with the different communities: frequent workshops, conferences, user groups, mailing lists. By building on existing tools, we intend to involve the communities in the development process itself in the spirit of open-source software.

We also intend to reach a larger crowd of researchers by cutting down non-research technical issues to access existing tools: building better documentation and tutorials, developing easy-to-install distributions, easy web and cloud access, better user interfaces, better interactions between different software. Doing so, our objective is to help the communities to grow themselves and interact together using our work.

B.0.1.2 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?)
```

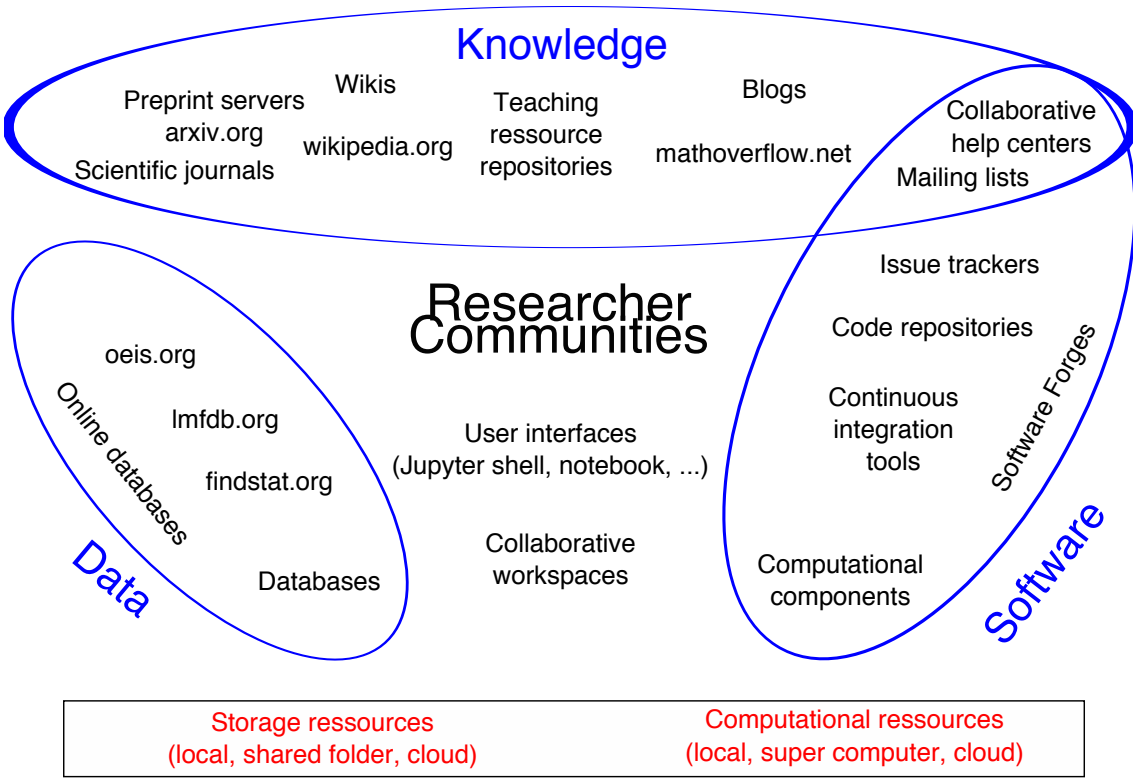

Figure 1: Virtual Research Environments for research in pure mathematics and applications.

B.0.1.3 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.♠

♠**TO DO**: *NT: the purpose of Figure 1 is to give a quick sense of what Virtual Research Environments can be in our context, and a "big picture" for the project. A graphic artist friend of mine is going to help me improve it. I have collected here some material for her.*

What we would like the "big picture" in Figure 1 to highlight:

This is a human centered project: *At the core: researchers and communities thereof.*

The three types of information: *Software, Knowledge, Data (currently in blue)*
How they interact:

- *Knowledge help structure data and software (e.g. through ontologies)*
- *Software produce data*
- *Data is used by researchers to build knowledge*

Physical resources: *(currently in red)*

Virtual Research Environments

- *Researchers in Math have a long tradition of collaborating on Software, Knowledge, and, up to some point, Data*
- *For this they use a variety of collaborative tools which form a loosely knit Virtual Research Environment.*
- ***Aim 2**: make it easy for subcommunities of researchers to setup custom collaborative work spaces / Virtual Research Environments tailored to their needs, by combining:*
  - *Computational resources*
  - *Storage resources*
  - *Computational software components*
  - *Databases*
  - *User interfaces*
  - *Wikis-Knowledge bases (true for findstat, LMFDB): quicker cycle for consolidation of information spread over papers/brains*

  Such VRE shall help them:
  - *collaboratively develop software (e.g. specialized libraries), data and knowledge (e.g. articles) for their research projects.*
  - *contribute back this information to the larger community whenever relevant.*

Processes:
*It would be interesting to depict the following processes. They are indeed about collaboration and sharing (and quality control), that is what **Aim 1** is to promote.*

Software development

- bug reports *and* enhancement requests *emerge from the community, typically through collaborative help centers, and are posted on issue trackers.*
- Design discussions *occur on mailing lists and issue trackers.*
- *Researchers* submit code *to the code repositories.*
- Quality control*: the code is reviewed and tested by continuous integration tools.*
- *Finally the code* integrated *within computational components, and used by the community.*

Researchers (as well as other users: teachers, engineers, ...) interact at each step of the process.

Scientific publication

- *researchers submit articles to journals and post them on preprint servers;*
- *the articles get reviewed by other researchers;*
- *finally they are distributed back to the community*

Improvements to implement:

- *the findstat link does not work for me, kerning looks extremely weird – POD*

- *lmfdb, oeis, and findstat have a strong knowledge component as well, with knowls and wikis, references, ...*
- *arxiv is not far from a database of knowledge*

A collection of links that might give some idea of the look and feel of our universe:

Examples of (computational) components:

- *IPython:* `http://ipython.org/`
- *GAP:* `http://www.gap-system.org/`
- *Singular:* `http://www.singular.uni-kl.de/`
- *Sage:* `http://sagemath.org/`
- *PARI/GP:* `http://pari.math.u-bordeaux.fr/`
- *Linbox:* `http://www.linalg.org/`

Examples of online collaborative tools

- *Issue tracker:* `http://trac.sagemath.org/timeline/`
- *Code repository:* `https://github.com/`
- *Collaborative help center:* `http://ask.sagemath.org/`
- *Collaborative math site:* `http://mathoverflow.net/`

Examples of online databases

- *Online databases:* `http://oeis.org/?language=french`
- *LMFDB:* `http://www.lmfdb.org/EllipticCurve/Q/14.a3`
- *Findstat:* `http://www.findstat.org/`

Example of graphical material

- `http://boxen.math.washington.edu/home/nthiery/main2014.pdf`

♠

B.0.1.3.1 Importance of experimental tools in pure mathematics and applications

From their early days, computers have been used in pure mathematics, either to prove theorems (e.g. the four color theorem) or, like the telescope for astronomers, to explore new theories. By now the experimental method, based on exact computer aided calculations, has now been added to the standard toolbox of the pure mathematician, and its usage has grown to the point that certain areas of mathematics now completely depend on it.

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 breakthroughs such as the recent complete resolution of Serre's conjectures, directly inspired by 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.

The field of computational mathematics 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 initiated by European researchers who made crucial contributions at all levels. These include the design of fundamental algorithms, the development of major computer algebra systems (♠**TO DO**: *this is a bit redundant with below*♠), applications of the computational methods in various fields, and the creation of widely used databases.

Particularly fruitful interactions unfold between computer algebra and algebraic geometry, number theory, combinatorics 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, combinatorics and group theory quite often lead to algorithmic breakthroughs which, in turn, open the door for new theoretical and practical applications of computer algebra.

B.0.1.3.2 A long track of collaboration on software, data, knowledge

Supporting the experimental method requires spending major efforts 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 specialized systems like `GAP`, `Singular`, or `PARI/GP` playing a major role for decades.

The next scale was reached in the last decade with the advent of the general purpose mathematical system `Sage` which proved the viability and sustainability of the "developed by users for users" development model at the international level.

♠**TO DO**: *This is somewhat redundant with the language in Objective 5; see where this belongs best to.*♠

♠**TO DO**: *Develop*♠Similarly, 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.

♠**TO DO**: *Develop*♠Mathematicians have a strong tradition of sharing knowledge openly (arxiv, Wikipedia, ...).

B.0.1.3.3 Early VRE's

♠**TO DO**: *Motivate the relevance of VRE's, in particular by the success of* `SageMathCloud` *or* `Simulagora`*. Mention as well* `LMFDB`*.*♠

♠**TO DO**: *Highlight some other deployed VRE's that would benefit to the sorts of improvements you suggest. You could include Wakari.io and also the tmpnb thing in Nature magazine: http://www.nature.com/news/ipython-interactive-demo-7.21492*♠

B.0.1.3.4 Key concept: bringing communities together toward a VRE kit

♠**TO DO**: *Focus on VRE kit and building blocks*♠

♠**TO DO**: *Why this focus? variability of needs, sustainability, ...*♠

♠**TO DO**: *Bringing communities together*♠

B.0.1.3.5 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`

The SPP1489 "Algorithmic and Experimental Methods in Algebra, Geometry, and Number Theory" is a nationwide Priority Project of the German Research Council DFG which commenced in July 2010 and will end in June 2016. The focus of the programme is on the interactions between computer algebra and algebraic geometry, number theory, and group theory. It combines expertise at all levels of research in computer algebra, be it the design of algorithms, the implementation of algorithms, the application of algorithms, or the creation of mathematical databases. The goal of SPP1489 is to considerably further the algorithmic and experimental methods in the afore mentioned disciplines, to combine the different methods across boundaries between the disciplines, and to apply them to central questions in theory and praxis. A fundamental concern of the programme is the further development of open source computer algebra systems with origins in Germany, which in the framework of different projects will be crosslinked 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 `http://ipython.org/sloan-grant.html`

♠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**: Description of Sage-Combinat grant]♠

HPAC grant from the A.N.R. The french national research agency ANR has funded a 4 years project on High Performance Algebraic Computing (HPAC) focused on the development of parallel exact linear algebra. The consortium gathers research groups from LIP6 (Paris 6), LIRMM (Montpellier), LIP (Lyon) and LIG and LJK (Grenoble). The main goals of the project is to first develop high performance exact linear algebra kernels with dedicated parallel runtime, propose a domain specific language for the parallelization of exact linear algebra libraries and their composition, invent new algorithmic solutions for large scale parallelizations. The output of the project is then twofolds: new computational challenges arising in algebraic cryptanalysis will be addressed, and the open-source libraries maintained by each group will not only integrate these advances, but will expose them in a close integration to high level computer algebra softwares. In this process, Sage will start benefitting from the new shared-memory parallel code of LinBox for the linear algebra over a finite field. The scope of this project is mostly focused on shared memory parallelism (except for some challenge computations). Addressing distributed and heterogeneous infrastructures is the next step after this project, that is be addressed in work-package 5 of the this proposal.

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

SageMathCloud https://cloud.sagemath.com/

♠NT/SL/WS [**WRITE HERE**: Proofread section about SMC]♠

SageMathCloud provides a collaborative online environment for students, teachers and researchers to interact with Sage and with each other. It has Sage and IPython worksheets, powerful LaTeX editing features and a full Linux computer, all accessible from a standard web browser. Its main design feature is to enable and promote collaboration between groups of users. It is for example a natural place to host a course, allowing teachers to collaborate with their students using modern tools like Sage and LaTeX, with facilities for real-time communication through chat, video, and shared editing of documents, programs and worksheets; course material can be provided as worksheets, assignments can be distributed, collected, and returned as well. Launched in 2013, SageMathCloud presently hosts over 100,000 projects and 10,000 weekly active users. This fast adoption by a wide variety of users demonstrates the relevance and the long term impact this kind of collaborative environments can have.

Technically speaking, SageMathCloud is a specific open-source cloud-based Virtual Research and Teaching Environment for mathematics developed since 2013 under the lead of William Stein, with funding from the NSF, and Google's Education Grant program. It's currently deployed at the University of Washington at Seattle, with a business plan in the work for commercial support for massive on line courses, subsidizing a free service for all other academic usage and some further Sage development.

In comparison **MathVRE** focuses on open source building blocks and architecture to easily setup and deploy custom Virtual Research Environments. On the one hand, SageMathCloud will serve as prototype for **MathVRE**, paving the way and showcasing important features from the users perspective. On the other hand, basically each and every task undertaken in **MathVRE** will benefit back SageMathCloud.

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, that were conjectured following computational exploration. 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. As a leading example of the use of databases in mathematical research, the LMFDB will provide **MathVRE** with a real large-scale prototype around which to develop new ideas about the design and implementation of such databases and their associated software. The feasibility of such collaboration was successfully tried at a workshop at the ICMS in Edinburgh in January 2013 on "Online databases: from L-functions to combinatorics", sponsored by the NSF, AIM and the ICMS.

Findstat?

Kwarc group

B.0.1.4 Ambition

♠SL [**WRITE HERE**: Ambition section]♠

♠**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.♠

B.0.1.4.1 Specificities of mathematics

Specific situation of maths w.r.t. Data:

- More often than not data is the result of a computation (and not e.g. an experiment). The role of databases is thus primarily to store results for later reuse (persistent caching), and enable searches. Because of this, many issues (semantic, ontologies, reproducibility, ...) are to be treated upstream at the level of software rather than data.
- extreme reification in mathematics makes classical ontologies techniques/RDF impractical
- interlinking very high
- several alternate and defining description of same objects

B.0.1.4.2 Challenges

Collaboration at the level of large communities.
Promoting collaboration over competition between communities.

B.0.2 Impact

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

B.0.2.1 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.)♠

B.0.2.2 Measures to Maximise Impact

B.0.2.2.1 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 By design (Objective 1), the VRE's promoted by **MathVRE** will consist of a thin layer on top of an ecosystem of components. Hence, the long term sustainability of those VRE is guaranteed by the sustainability of the ecosystem of components, that is by Objective 5.

B.0.2.2.2 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.*♠

B.0.3 Implementation

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

B.0.3.1 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 20 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 B.0.1.1, page 3) 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

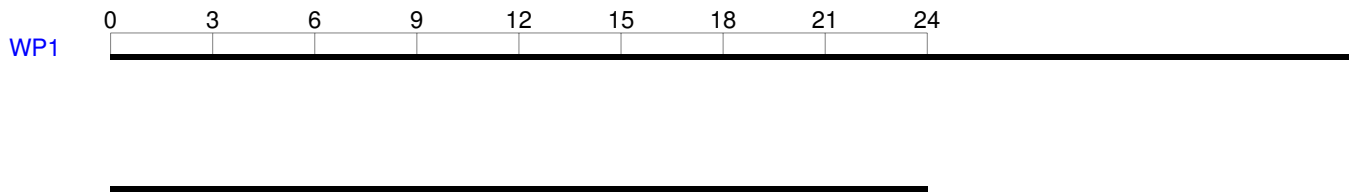

Figure 2: Overview Work Package Activities (lower bar shows the overall effort per month)

WP	Title	type	page	start	end	UPS	Logilab	UVSQ	UJF	UB	UO	USHEF	USO	USTAN	UW	JU	UK	US	UZH	Simula	UWS
WP1	Management	MGT	23	0	48	48								48							
totals						**48**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**48**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
intended totals						**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**	**48**

Efforts in PM; WP lead efforts light gray italicised

Figure 3: Work Packages

#	Deliverable name	WP	Nature	Level	Due
M1.1	Create tickets for all relevant tasks / deliverables	WP1	??	??	1
M1.2	Internal and external mailing lists	WP1	??	??	1
M1.3	Internal software repository	WP1	??	??	1
M1.4	Project Periodic Report (first year)	WP1	??	??	12
M1.5	Project Periodic Report (second year)	WP1	??	??	24
M1.6	Project Periodic Report (third year)	WP1	??	??	36
M1.7	Project Periodic Report (fourth year)	WP1	??	??	48
M1.8	Project Final Report	WP1	??	??	48

The work in the DreamKit project is structured by seven milestones, which coincide with the project meetings in summer and fall. Since the meetings are the main face-to-face interaction points in the project, it is suitable to schedule the milestones for these events, where they can be discussed in detail. We envision that this setup will give the project the vital coherence in spite of the broad mix of disciplinary backgrounds of the participants.[2] EdN:2

#	Name	WPs[1]/Deliverables involved	Mo	Means of Verif.
M1	**Initial Infrastructure**		1	Inspection
Set up the organizational infrastructure, in particular: Web Presence, project TRAC,...				
M2	**Consensus**		24	Inspection
Reach Consensus on the way the project goes				
M3	**Exploitation**		36	Inspection
The exploitation plan should be clear so that we can start on this in the last year.				
M4	**Final Results**		48	Inspection
all is done				

♠**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.* ♠

B.0.3.2 Work Package Descriptions

Work Package 1: Project Management

Objectives

The objectives of this work package 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

This workpackage will perform all the activities related to monitoring of progress towards the project milestones shown on Page **??** and the deliverables listed on Page **??**, 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, in-

[2] EdNote: *maybe automate the milestones*

[1] The work package number is the first number in the deliverable number.

cluding supporting attendance at project meetings, convening Advisory Board meetings, and obtaining feedback on the project direction and results.

♠**TO DO**: *MK: I would combine the first three into one "basic project intrastructure"*♠ **Deliverables**:

M1.1: (Month 1; nature: ??, dissem.: ??) *Create tickets for all relevant tasks / deliverables*
M1.2: (Month 1; nature: ??, dissem.: ??) *Internal and external mailing lists*
M1.3: (Month 1; nature: ??, dissem.: ??) *Internal software repository*
M1.4: (Month 12; nature: ??, dissem.: ??) *Project Periodic Report (first year)*
M1.5: (Month 24; nature: ??, dissem.: ??) *Project Periodic Report (second year)*
M1.6: (Month 36; nature: ??, dissem.: ??) *Project Periodic Report (third year)*
M1.7: (Month 48; nature: ??, dissem.: ??) *Project Periodic Report (fourth year)*
M1.8: (Month 48; nature: ??, dissem.: ??) *Project Final Report*

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

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

B.0.3.3 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).♠

B.0.3.4 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*♠

Writing interfaces between computer algebra systems from different areas and collaboritive software development are important themes within the DFG Priority Project SPP1489. As in the SAGE community, networking measures include the regular exchange of developers and the regular organization of software workshops (coding sprints) which bring whole teams together for solution finding and intense code writing. Particular tight collaborations exist between the GAP and the SINGULAR communities, with major GAP-SINGULAR developers meetings taking alternately place at St. Andrews, Kaiserslautern, and Aachen. See http://www.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).*♠

B.0.3.5 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.*♠

	RTD		Management		Totals		
	(direct costs)				(incl. indirect costs)		
Partner	Personnel	Add. Direct	Personnel	Add. Direct	RTD	MGT	EC Contrib.
UPS							
Logilab							
UVSQ							
UJF							
UB							
UO							
USHEF							
USO							
USTAN							
UW							
JU							
UK							
US							
UZH							
Simula							
UWS							
totals							

Table 3: Overview: Resources to be committed (all in €)

♠**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.♠

B.0.4 Members of the Consortium

B.0.4.1 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; this includes a description of the profile of the to-be-recruited personnel*
- *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.*

♠

University of St Andrews

The Centre for Interdisciplinary Research in Computational Algebra (CIRCA) fosters research at the interface of Mathematics and Computer Science including abstract and algorithmic algebra and combinatorics, formal languages and automata, mathematical software and constraint programming. Our success is founded on the close integration of theoretical and algorithmic research and the development and use of state-of-the-art software.

CIRCA's output includes first class results in Pure Mathematics and in Computer Science, recognised by our highly-cited publications in top international venues in both disciplines and our widely used research software. Beyond the individual international connections of the investigators and research staff, CIRCA as a centre has national importance and international standing. CIRCA has been selected to host major conferences such as CP 2010, BCC 2009, PP 2007, and the "Groups St Andrews" series in 2005 and 2013.

Curriculum vitae

Steve Linton is a Professor of Computer Science at St Andrews. He has worked in computational algebra since 1986 and has helped coordinate the development of GAP since its move from Aachen in 1997. He personally wrote key features of GAP, such as workspaces and exception handling, and has overseen the development and releases of the whole system. He directed CIRCA from 2000–2013. He is an editor of AAECC[2]. He has been PI of four major EPSRC grants and coordinated the EU project SCIEnce. He is the general chair of ISSAC 2015, the main conference in computer algebra.

Alexander Konovalov is a Senior Research Fellow in CIRCA and has worked on GAP for more than 10 years. After holding the fellowship at the Vrije Universiteit Brussel in 2006, researching computational group ring theory, he moved to St Andrews in 2007 to join EU project SCIEnce. He leads many aspects of the GAP project, including release preparation, regression testing and liaison with package authors. He has authored 38 papers and 8 GAP packages, and co-organised a number of events, most recently the LMS/EPSRC Short Instructional Course in Computational Group Theory, the HPC-GAP workshop (2013), and the Summer School on Experimental Methodology in Computational Science Research (2014). He is an editor of Journal of Software for Algebra and Geometry and a Fellow of the Software Sustainability Institute.

Publications, products, achievements

1. ♠SL/AK [**WRITE HERE**: Up to 5: SCIEnce-JSC, ...]♠

Previous projects or activities

1. ♠SL/AK [**WRITE HERE**: Up to 5: SCIEnce, HPC-GAP, Critical Mass, Dominion, ...]♠

Significant infrastructure

♠SL/AK [**WRITE HERE**: ...]♠

[2] Applicable Algebra and Error Correcting Codes

Universit Paris Saclay

♠**TO DO**: *INRIA or INRA below? – sorry to jump in, Paul*♠

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 Universit 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 Universit 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 Universit 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 Universit 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.

Viviane Pons Matre de Confrences at the Laboratoire de Recherche en Informatique, Viviane Pons is a young researcher in Algebraic Combinatorics. She defended her thesis in 2013 and has 3 papers in international journals and 3 communications in international conferences. Before committing herself to research, she spent two years working in industry as a Java and web developer.

She discovered Sage during her first sage-days in 2010 and has since been an active user and contributor with 10 (co)authored tickets improving the support of combinatorial objects into Sage. She is very involved in the promotion of Sage, participating in Sage-Days and proposing Sage introduction tutorials or Sage presentations in various conferences. She is also one of the main developers of the project FindStat dedicated to databases in combinatorics.

Nicolas M. Thiry Professor at the Laboratoire de Recherche en Informatique, Nicolas M. Thiry 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, Thiry 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, Thiry 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 Mathmatique avec Sage”.

Publications, achievements

♠**TO DO**: *Il faut łtre plus formel dans la description des projets antrieurs : Acronyme, titre, agence de financement, dure. Pareil pour les publi - auteurs, titre exact, anne etc.*♠

1. Lead of the `Sage-Combinat` software project.
2. Coauthoring of the open source book “Calcul Mathmatique 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 Universit 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.

♠**TO DO**: *Comments by Olivier Chapuis*♠ Paris Sud also hosts the WILDER platform, an experimental wall-sized high-resolution interactive touch-screen for conducting research on collaborative human-computer interaction and the visualization of large datasets.

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

B.0.4.1.1 University of Kaiserslautern

Principal investigator Prof. Dr. Wolfram Decker Wolfram Decker is a professor of mathematics at TU Kaiserslautern. He formerly was a research fellow at Berkeley with a NATO grant, a visiting researcher at Kyoto with a JSPS grant, and a professor at Saarbrücken, Germany. Decker has more than thirty publications including two books on computational algebraic geometry and papers in Compositio, Crelle, and Mathematische Annalen. He has held several grants in four different priority programmes of the German Research Council DFG and is now coordinator of the priority programme SPP 1489 "Algorithmic and Experimental Methods in Algebra, Geometry, and Number Theory". He was also coordinator of the European algebraic geometry network EuroProj (1996–1999) and Chair of the programme management committee of the European algebraic geometry network EAGER (2000–2004). He held seven grants for EU Highlevel Scientific Conferences and (co-)organized about 50 conferences, summer schools, workshops, and coding sprints. He was Chair of the Minisymposium on Computer Algebra during the third ECM. Decker has supervised 13 PhD students. He has been a frequent lecturer at the African Institute of Mathematics (AIMS) at Cape Town, and he has run 8 schools on computational algebraic geometry in different countries.

Decker's research interests lie in areas of algebraic geometry and computer algebra. In addition to writing theoretical papers, he is a leader in mathematical software development and has written thousands of lines of code himself. He has made contributions to the systems MACAULAY2 and, much more substantially, SINGULAR. Since 2009 he is the head of the SINGULAR development team. Current tasks of the team include crosslinking SINGULAR to other systems, most notably to GAP, and parallelizing SINGULAR. These tasks are fundamental to the **MathVRE** project.

Publications, products, achievements

1. Head of SINGULAR development team.
2. Coordinator of the DFG Priority Project SPP1489 *Algorithmic and Experimental Methods in Algebra, Geometry, and Number Theory'*.
3. Coauthor of the book *Computing in Algebraic Geometry - A Quick Start using Singular*, Springer.
4. Coauthor of the book *A First Course in Computational Algebraic Geometry*, Cambridge University Press.

Previous projects or activities

1. Member of the DFG Priority Project *Algorithmic Number Theory and Algebra*.

Significant infrastructure Excellent computing infrastructure (high end servers), access to different types of compute clusters through the IT-Center of the TU Kaiserslautern.

University of Oxford

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

Curriculum vitae

Ursula Martin Professor Ursula Martin has recently joined the University of Oxford, where she holds a Professorship, in conjunction with a Senior EPSRC Fellowship, on a joint arrangement between the Department of Computer Science and the Mathematical Institute. Her current research concerns social and computational techniques for creating mathematics, building on a significant track record at the interface of mathematics and computing. Prior to this she worked at Queen Mary University of London, where as Vice Principal for Science and Engineering she led strategic change, and was active in knowledge transfer activities and developing young staff.

Her work is characterized by strongly interdisciplinary collaboration in new problem domains at the interface of mathematics and computer science, identifying novel interactions between theory and practice, with real-world problems inspiring scientific advance. Major achievements include results linking randomness and symmetry, new unifying explanations of the power of computational logic, and new practical techniques for using computational logic and algebra in industry.

The work to be undertaken in the Work Package 5 (Social Aspects) fits very well into her current project, which concerns crowdsourced mathematics: the overarching goal is to understand and extend the human and computer creation of mathematics. It is mostly funded by her 2014 EPSRC Advanced Fellowship (EPSRC awards only one or two of these annually in Computer Science) is a partnership of industry, government and international academia. ♠UM [**WRITE HERE**: CV Ursula Martin]♠

First name Last name to be added. ♠DP [**WRITE HERE**: CV Edith Elkind]♠

Dmitrii Pasechnik is a Senior Research Fellow at the Department of Computer Science of the University of Oxford, where he also holds a Lectureship at Pembroke College. Before moving to Oxford in 2013, he taught mathematics for 8 years in Nanyang Technological University (Singapore). While there, he was successful in receiving individual grant funding totalling over 500K, graduated 2 PhD students, supervised post-doctoral researchers, and co-organized a 2-months research program at Singapore Institute for Mathematical Sciences on a range of topics in computational mathematics, involving over 100 participants.

He works on a wide area of interconnected topics, related to computational algebra and optimization, combinatorics, algorithm, symbolic computing, and game theory, and authored over 70 papers on these topics, several of them using Sage and/or its components, such as GAP.

He is an active Sage developer, and regularly contributes, himself or together with his undergraduate or graduate students, new or improved Sage interfaces to various mathematical packages and databases. He has good understanding of the overall Sage development process, as well as of development of other open-source sofware and databases, including their social/community aspects. ♠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]♠

Universit de Bordeaux

- INRIA, LaBRI, IMB
- journal de théorie des nombres
- Plafrim and a mesocentre Avakas
- Several softwares developped in Bordeaux: pari/GP, tulip, etc

Curriculum vitae

Vincent Delecroix CNRS researcher at the Laboratoire Bordelais de recherche en informatique, Vincent Delecroix is a junior researcher in Number theory, Combinatorics and Dynamical systems.

X publications. He has several international collaboration (England, Mexico, United-States).

- Bobo - cours Sage Bordeaux

Delecroix is a major developer of Sage in various components: - integers - combinatorics - dynamical systems

Karim Belabas Karim is One of the main pari developer.

Bill Allombert CNRS Ingénieur de Recherche. One of the main pari developer.

Publications, products, achievements

Some recent Publications :

1. Belabas, Karim; Friedman, Eduardo; Computing the residue of the Dedekind zeta function. Math. Comp. 84 (2015), no. 291, 357369.
2. The PARI Group; PARI/GP version 2.7.0, Bordeaux, 2014, http://pari.math.u-bordeaux.fr/.
3. Belabas, Karim et al. Explicit methods in number theory. Rational points and Diophantine equations, 179 pages, Panoramas et Synthses 36, 179p., 2012.
4. Allombert, Bill; Bilu, Yuri and Pizarro-Madariaga, Amalia;) CM-Points on Straight Lines , to appear in "Analytic Number Theory" (dedicated do H. Maier), Springer.

Previous projects or activities

Current grants:

1. ANR PEACE (2012-2015) Goal: The discrete logarithm problem on algebraic curves is one of the rare contact points between deep theoretical questions in arithmetic geometry and every day applications. On the one side it involves a better understanding, from an effective point of view, of moduli space of curves, of abelian varieties, the maps that link these spaces and the objects they classify. On the other side, new and efficient algorithms to compute the discrete logarithm problem would have dramatic consequences on the security and efficiency of already deployed cryptographic devices.
2. ERC starting grant ANTICS (2011-2016) Goal: "Rebuild algorithmic number theory on the firm grounds of theoretical computer science". Challenges: complexity (how fast can an algorithm be?), reliability (how correct should an algorithm be?), parallelisation.

Significant infrastructure

♠VD [**WRITE HERE**: this still needs to be done]♠

Two center for computations: Plafrim and Avakas.

Universit de Grenoble

♠JGD/CP [**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.

Clément Pernet Associate Professor at the joint Inria-LIG research group MOAIS, Clément Pernet is a junior researcher in Computer Algebra, parallel computing and coding theory with 16 papers published in international journals or refereed international conferences. He is associate editor of the ACM transactions on Mathematical Software and has co-organized 10 conferences, including 2 sage-days and the 2012 edition of ISSAC, the leading conference in computer algebra.

Since he was a post-doc at University of Washington, under the supervision of William Stein, head of the Sage project, he has had many contributions to Sage on the exact linear algebra and the symbolic computation tools. He co-authored the book “Calcul Mathmatique avec Sage” with the chapter on Linear algebra. Clément Pernet is the founder and lead developper of the fflas-ffpack library, kernel for dense linear algebra over a finite field, delivering high performance computation to LinBox and Sage. He is a core contributor to the LinBox library and contributed to the m4ri library.

Pierrick Brunet Junior Research and Development Engineer at INRIA Grenoble, Pierrick Brunet is working on compilation of C/C++ OpenMP program to C/C++ programs with calls to specific OpenMP runtimes.

With about 25% of commits in the Pythran project, Pierrick is one of the core devs of this project which compile a subset of the Python language to native Python modules.

Publications, products, achievements

Software projects

`fflas-ffpack`**:** An open-source `C++` library offering dense linear algebra kernels over a finite field. In the same spirit as the numerical `BLAS` (Basic Linear Algebra Subroutines), and `LAPACK` libraries, it delivers high performance for the most commonly used routines of scientific computing: matrix multiplication, solving linear systems, computing echelon forms, determinants, characteristic polynomials, etc. This library has set the standard approach for high performance exact dense linear algebra. It is currently used in `Sage`, and has inspired the design of similar routines in most commercial computer algebra softwares: `maple`, `magma`, etc.

`LinBox`**:** An open-source `C++` middleware library for exact linear algebra. It uses `fflas-ffpack` for its dense finite field linear algebra part and extends its functionalities to other computation domains (integers, rationals, polynomial rings) and type matrices (sparse and structures matrices, black-box matrices). `LinBox` is integrated in `Sage`.

`Pythran`**:** An open-source `Python`-to-`C++` optimizing compiler offering an high performance runtime for Scientific Python kernels. Dynamicity of the `Python` language is not compliant with static compilation. That's why only a subset of the `Python` language is supported by `Pythran`. Thanks to these restrictions, `Pythran` generate code up to 3000 faster than original module.

Selected Publications

1. Coauthoring of the open source book "Calcul Mathmatique avec Sage", the first of its kind comprehensive introduction to computational mathematics in Sage for education.
2. Parallel computation of echelon forms (with J-G. Dumas, T. Gautier and Z. Sultan). *In Proc. Euro-Par'14* (2014), LNCS 499–510. DOI: 10.1007/978-3-319-09873-9_42.
3. Pythran: Enabling static optimization of scientific python programs (Serge Guelton, Pierrick Brunet, Alan Raynaud, Adrien Merlini, and Mehdi Amini.) *Proceedings of the Python for Scientific Computing Conference (SciPy)* June 2013.
4. Fast Computation of Hermite Normal forms of random integer matrices (with W. Stein). *J. of Number Theory* **130.7** (2010), 1675–16833. DOI: 10.1016/j.jnt.2010.01.017
5. Dense Linear Algebra over Word-size Prime Fields (with J.-G. Dumas and P. Giorgi). *Trans. on Math. Software* **35.3** (2008), 1–42. DOI: 10.1145/1391989.1391992.
6. Faster Computation of the Characteristic Polynomial (with A. Storjohann). *In Proc. ISSAC'07* (2007), 307–314. DOI: 10.1145/1277548.

Previous projects or activities

1. Direction of the ANR program on High-Performance Algebraic Computations 2012-2015.
2. Participation to the NSF-Inria associate teams QOLAPS (with NCSU, USA)
3. Coordination of a CNRS PEPS grant (parallel computer algebra)
4. Organization of the ISSAC'12 conference, the main international conference in computer algebra, and of PASCO'15 a satelitte conference on parallel computer algebra.

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)
7. Mention Debian development experience

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. Five members of the Department are Fellows of the Royal Society, and one, Regius Professor Martin Hairer, was awarded a Fields Medal in 2014. 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). Nine members of the department currently hold ERC grants.

Curriculum vitae

John E. Cremona Professor of Mathematics. 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). 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 have been important for the successful management 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. 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). Of the group's members, two (Loeffler and Dokchitser) hold Royal Society Research Fellowships and one (Bartel) a Royal Commission 1851 Fellowship. Loeffler won a Leverhulme Foundation Prize jointly with Zerbes.

2. Several members of the Number Theory group at Warwick are `Sage` developers, including 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; David Loeffler, who has contributed substantially to the modular forms module in Sage; and postdoc Marc Masdeu, who has worked on the `Sage`-Flint interface.

Previous projects or activities

1. 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.

2. 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.

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.

Jacobs University Bremen

Jacobs University is a private Anglo-Saxon style research university. It opened in 2001 and has an international student body (1320 students from 115 nations as of 2011). The KWARC (KnoWledge Adaptation and Reasoning for Content [**KWARC:online**]) Group headed by *Prof. Dr. Michael Kohlhase* specializes in knowledge management for STEM. Formal logic, natural language semantics, and semantic web technology provide the foundations for the research of the group.

Curriculum vitae

Michael Kohlhase Dr. Michael Kohlhase is full professor for Computer Science at Jacobs University Bremen and an associate adjunct professor at Carnegie Mellon University.

He studied pure mathematics at the Universities of Tübingen and Bonn (1983 - 1989) and continued with computer science, in particular higher-order unification and automated theorem proving (Ph.D. 1994, Saarland University).

His current research interests include knowledge representation for mathematics, inference-based techniques for natural language processing, and computer-supported education. He has pursued these interests during extended visits to Carnegie Mellon University, SRI International, and the Universities of Amsterdam, Edinburgh, and Auckland.

Michael Kohlhase is recipient of the dissertation award of the Association of German Artificial Intelligence Institutes (AKI; 1995) and of a Heisenberg stipend of the German Research Council (DFG 2000-2003). He was a member of the Special Research Action 378 (Resource-Adaptive Cognitive Processes), leading projects on both automated theorem proving and computational linguistics. Michael Kohlhase was trustee of the Conference on Automated Deduction (CADE), Mathematical Knowledge Management (MKM), and the CALCULEMUS conference, he is a member of the W3C Math working group, president of the OpenMath Society, and the general secretary of the Conference on Intelligence Computer Mathematics (CICM).

Florian Rabe ***** to be written *****

Relevant previous experience: The KWARC group is the lead implementor of the OMDoc (Open Mathematical Document) format for representing mathematical knowledge [**Kohlhase:OMDoc1.2**] and redeveloped its formal core in the OMDoc/MMT format [**RabKoh:WSMSML13**]. The latter has been implemented in the MMT system [**MMTSVN:on**, **RabKoh:WSMSML13**] which provides efficient implementations of the computational primitives such as type checking, flattening, and presentation at a logic/foundation-independent level. The group has developed services powered by such semantically rich representations, different paths to obtaining them, as well as platforms that integrate both aspects. *Services* include the adaptive context-sensitive presentation framework provided by the MMT API and the semantic search engine MathWebSearch[**KohSuc:asemf06**, **ProKoh:mwssofse12**].

Semantic services can be integrated into the documents generated from OMDoc/MMT representations, making them into “active documents”, i.e. documents that are interactive and adaptive to the user and situatinon. For *obtaining* rich content, the group investigates assisted manual editing [**JucKoh:sidesc10:biblatex**] as well as automatic annotation using linguistic techniques [**GinJucAnc:alsaacl09**]. Finally, KWARC has developed the MathHub.info portal a community-based library and knoweldge management system for flexiformal libraries, which can be used for semantic publishing and eLearning [**KohDavGin:psewads11**, **MathHub:on**, **IanJucKoh:sdm14**].

The OMDoc/MMT knowledge representation format and the MathHub.info system will an important basis for the developments Work Packages 4 and 6.

Michael Kohlhase has initiated and led the CALCULEMUS! IHP-Research and Training Network and participated in the FP6 IST MoWGLI (Mathematics on the Web: Get it by Logic and Interfaces) project, the FP6 CSA Once-CS (Open Network of Centres of Excellence in Complex Systems), The FP7 EDC project WebALT (Web Advanced Learning Technologies).

Specific expertise:

- Modeling formal structures of mathematical knowledge in a web-scalable way
- Transforming large collections of legacy scientific publications to semantically structured markup
- Designing user interfaces for authoring and interacting with mathematical knowledge

University of Zürich

The University of Zurich consistently ranks among the top 15 research institutions in Europe. It is the largest university in Switzerland, with over 26000 students, and offers the most comprehensive academic program of the country. It has close to 600 professors and over 5000 academic staff.

Switzerland ranks high in innovation, competitiveness and research spending, and much of this is enthusiasm for research is concentrated around Zurich. UZH also benefits from synergies with the ETH Zurich.

The Mathematics Institute has 17 professors and around 60 PhD students, part of a graduate school run jointly with ETH Zurich. Also joint is a Computational Science program uniting 47 researchers, mostly in the sciences, who make use of computational methods.

Curriculum vitae

Paul-Olivier Dehaye Paul-Olivier Dehaye is a Swiss National Science Foundation Assistant Professor at the University of Zurich. After his Phd at Stanford (2006), he has also worked in Oxford, at the Institut des Hautes Etudes Scientifiques and at ETH Zurich. He currently has 13 papers published in international peer-reviewed journals. He is currently supervising three PhD students and one post-doc.

His main research is at the intersection of Number Theory and Combinatorics, and in particular in Random Matrix Theory conjectures. He has additional interests in FLOSS, semantic tools, massive online education and crowdsourcing, all with the view of enabling larger scale mathematical and scientific collaborations. He is also member of the program committee of CICM 2015 (Conference on Intelligent Computer Mathematics).

He is a contributor to the `Sage`, `LMFDB` and `OpenEdX` projects, and has organised two conferences relating to these projects. The first was held in 2013 in Edinburgh, and organised jointly with Nicolas Thiery. Its official title was *Online databases: from L-functions to combinatorics*, and it served as a precursor to some aspects of this grant, by bringing the `Sage-Combinat` and `LMFDB` communities together. The second was held in June 2014 in Zurich and organised jointly with Stanford. It aimed at building a community around the open source python-based MOOC platform `OpenEdX`, and opened a series of conferences now held twice annually.

Dehaye has also taught for two years now a python course using `OpenEdX`, which aims to bring first year students to the level of potential contributor to `Sage`. This course also has a project-based component. It is now run locally for a small audience, but could be scaled up in various ways.

Publications, products, achievements

1. Dehaye is editor for the LMFDB, and has contributed to the project since its inception (2007). His students are also contributors.

2. For several of his papers, Dehaye used extensive computer-assisted experimentation (using mostly the combinatorial components of `Sage`) to inform the formulation of the eventual theorems, including for instance:

    - *Combinatorics of lower order terms in the moments conjecture for the Riemann zeta function*, arXiv:1201.4478
    - *Integrality of hook ratios*, arXiv:1111.5959, in *Proceedings of the Formal Power Series and Algebraic Combinatorics 2012 (Nagoya) conference*.
    - *A multiset hook length formula and some applications*, with Guoniu Han, in *Discrete Mathematics*, (311) 23–24, pp. 2690–2702, 2011.
    - *A note on moments of derivatives of characteristic polynomials*, in *DMTCS Proc. Formal Power Series and Algebraic Combinatorics 2010*, vol. 12.
    - *Joint moments of derivatives of characteristic polynomials*, in *Algebra and Number Theory Journal* **2** (2008), no. 1, pp. 31–68.

3. Dehaye has been extensively involved in teaching `Python` and `Sage` at UZH, through an online platform called `OpenEdX`. This has led him to organise the first community-driven conference around this (open-source) software, and to develop (together with students) additional tools, such as `edx-presenter`.

Previous projects or activities

Swiss National Foundation PP00P2/138906: Combinatorics of partitions and number theoretic aspects
This grant covers research at the intersection of number theory and combinatorics. Some of its aims are to uncover combinatorial structures that lurk in complicated formulaes for moments of L-functions (such as the Riemann zeta function). As such, it is simultaneously a heavy user of numerical methods from analytic number theory and of combinatorial techniques implemented in `Sage`.

Significant infrastructure

1. The Faculty of Sciences of the UZH benefits from very strong specialized IT support in the form of the S3IT group. They operate for instance a research cloud and a local supercomputer, and provide further assistance for the design of hardware and software systems to further research. They have a pool of software engineers that can be hired on projects such as this one for shorter periods.

2. UZH has a stake in Piz Daint, currently the sixth largest (and most energy-efficient) supercomputer in the world. This supercomputer is now currently expanded.

Simula Research Laboratory

Dedicated to tackling scientific challenges with long-term impact and of genuine importance to real life, Simula Research Laboratory (Simula) offers an environment that emphasizes and promotes basic research. At the same time, we are deeply involved in research education and application-driven innovation and commercialisation.

Simula was established as a non-profit, limited company in 2001, and is fully owned by the Norwegian Ministry of Education and Research. Its research is funded through competitive grants from national funding agencies and the EC, research contracts with industry, and a basic allowance from the state. Simula?s operations are conducted in a seamless integration with the two subsidiaries Simula School of Research and Innovation and Simula Innovation.

At its outset, the laboratory was given the mandate of becoming an internationally leading research institution within select fields in information and communications technology. These fields are (i) communication systems, including cyber-security; (ii) scientific computing, aiming at fast and reliable solutions of mathematical models in biomedicine, geoscience, and renewable energy; and (iii) software engineering, focusing on testing and verification of mission-critical software systems, and on planning and cost estimation of large software development projects. Recent evaluations state that Simula has met its challenge and is an acknowledged contributor to top-level research in its focus areas. Specifically, in the 2012 national evaluation of ICT research organized by the Research Council of Norway and conducted by an international expert panel, Simula received the highest average score (4.67) on a 1-5 scale among all evaluated institutions. In comparison, the national average was 3.38. Only five of the 62 research groups evaluated were awarded the top grade (5), and two of these five groups are located at Simula.

Simula is currently hosting one Norwegian Centre of Excellence, Center for Biomedical Computing (2007-2017), and one Norwegian Centre for Research-based Innovation, Certus (2011-2018). In addition, we participate as research partner in another Centre for Research-based Innovation, Centre for Cardiological Innovation (2011-2018), hosted by Oslo University Hospital. These two center-oriented schemes are the most prestigious funding instruments offered by the Research Council of Norway.

Curriculum vitae

Hans Petter Langtangen is director of Center for Biomedical Computing at Simula Research Laboratory, a Norwegian Center of Excellence doing inter-disciplinary research in the intersection of mathematics, physics, computer science, geoscience and medicine. Langtangen is on 80% leave from a position as professor at the Department of Informatics, University of Oslo.

Langtangen received his PhD from the Department of Mathematics, University of Oslo, in 1989, and then worked at SINTEF before being hired as assistant professor at the University of Oslo in 1991. After being promoted to full professor of mechanics at the Department of Mathematics in 1998, he moved in 1999 to a professorship in computer science. In the period 1999-2002 he also held an adjunct professor position at the Department of Scientific Computing at Uppsala University in Sweden. The Simula Research Laboratory was formed in 2001, and Langtangen has since then worked with research and management at this laboratory. The scientific computing activity at Simula has been awarded the highest grade, Excellent, by five panels of top-ranked international scientists in the period 2001-2012.

Langtangen's research is inter-disciplinary and involves continuum mechanical modeling, applied mathematics, stochastic uncertainty quantification, and scientific computing, with applications to biomedicine and geoscience in particular. He has also been occupied with developing and distributing scientific software to make the research results more widely accessible and help accelerate research elsewhere. For over three decades he has been very active with teaching and supervision.

The scientific production consists of 4 authored books, 3 edited books, about 60 papers in international journals, about 60 peer-reviewed book chapters and conference papers, and over 130 scientific presentations. The publications cover fluid flow, elasticity, wave propagation, heat transfer, finite element methods, uncertainty quantification, and implementation techniques for scientific software. Langtangen is on the editorial board of 7 journals and serves as Editor-in-Chief of the leading SIAM Journal on Scientific Computing. He is also a member of the Norwegian Academy of Science and Letters.

Publications, products, achievements

1. ♠XXX [**WRITE HERE**: ...]♠

Previous projects or activities

1. The Centre for Biomedical Computing, a Norwegian Centre of Excellence, awarded by the Research Council of Norway. Duration: 2007-2017. Budget: 75 MNOK (10 MEUR).
2. The FEniCS Project (www.fenicsproject.org) Duration: 2007–on-going.

Significant infrastructure

The fully owned Simula subsidiary Simula Innovation handles pre-commercial innovation projects, creation and follow-up of company spin-offs, and general support for entrepreneurs.

University of Southampton

The University of Southampton (UoS) is one of the leading universities in the United Kingdom, was founded in 1952 and is a member of prestigious Russell Group of UK Universities. UoS has more than 19,000 undergraduate students and 4,000 postgraduates and is an excellent venue for conducting cutting-edge research and for providing high quality education. The university is truly international, drawing students from over 130 different countries and benefiting from a wide and varied culture. It is ranked in the top 1% of universities worldwide (QS world university rankings 2014-15) and in the top 15 of research led universities in the UK, and is participating in a high number of collaborative research projects and related initiatives. UoS has a successful track record of industrial collaborations and is at the centre of a cluster of local high technology companies. It has an enviable track record in the generation of patentable work, with a portfolio of over 350 patents. To ensure the impact of its research projects, University of Southamptons Research & Innovation Services (R&IS) is responsible for professional protection of IP and supporting commercial development with industry. R&IS has had considerable success, licensing annual revenue in excess of € 1million and launching twelve successful spin-out companies since 2000. UoS has a strong track record of working in European projects, especially within the Framework Programme. The EC 6th FP7 Monitoring Report ranked UoS 17th out of all higher and secondary education organisations for number of FP7 participations during 2007-2012. Throughout the FP7 UoS has received € 132M in research grants and has been involved in 319 projects, including 63 ICT and 8 INFRASTRUCTURES Collaborative Projects. In 2013/14 alone UoS has received over € 181.5M in research grants and contracts, including over € 16.3M from the European Commission.

The Faculty of Engineering and the Environment (FEE) is one of the lead engineering faculties in Europe, educating a range of professionals and generating research of the highest quality. In the 2008 Research Assessment Exercise(RAE), FEE was ranked second in the UK in terms of research power for both civil engineering and mechanical/aero/production engineering. FEE brings together a wide range of disciplines, offering undergraduate and postgraduate programmes in audiology and environmental science as well as acoustical, civil and environmental, mechanical, and aeronautical/astronautical engineering and ship science. It consists of 370 research postgraduate students and 340 academic and research staff. FEE also hosts the University Technology Centres and Research Framework Agreements with key partners including: Airbus, Rolls-Royce, Lloyds Register, Microsoft and Network Rail. FEE has a strong background in working on international research projects, including 84 EU FP7 projects worth over € 28M. In 2013/14 only FEE has received about € 50M in research grants and contracts, of which over € 1.7M from EU funding programmes.

Curriculum vitae

Principal Investigator Prof Hans Fangohr Hans Fangohr is Professor of Computational Modelling at the University of Southampton. He heads the University's interdisciplinary Computational Modelling Group (http://cmg.soton.ac.uk), and has more than 100 publications on applied computer simulation in magnetism as well as development of computational methods. He has attracted over € 12m as investigator and co-investigator for teaching, research and research infrastructure grants.

In 2013, he attracted € 5m from the UK's Engineering and Physical Sciences Research Council (EPSRC) to fund the only Centre for Doctoral Training in Next Generation Computational Modelling (ngcm.soton.ac.uk) in the UK. This flagship activity will train about 75 PhD students (10 to 15 starting every year, beginning in September 2014) in the state-of-the-art and best-practice in computational modelling, the programming of existing and emerging parallel hardware and to apply these skills and tools to PhD research projects across a range of topics from Science and Engineering. The centre has chosen IPython as a key tool to deliver this teaching, document and communicate computational exploration and drive reproducible computation to push for excellent computational science.

Hans Fangohr has led the development of the Open Source Nmag software (http://nmag.soton.ac.uk), which prodives a finite-element micromagnetic simulation suite to a community of material scientists, engineers and physicists who research magnetic nanostructures in academia and industry. He has designed the package in 2005 so that it has an IPython-compatible Python interface, to make the workflow of using the simulation package as accessible as possible to scientists without substantial computational background.

He has extensive experience in micromagnetic simulation use and development.

Publications, products, achievements

1. ♠XXX [**WRITE HERE**: ...]♠

Previous projects or activities

1. ♠XXX [**WRITE HERE**: ...]♠

Significant infrastructure

♠XXX [**WRITE HERE**: ...]♠

Universit de Versailles – Saint-Quentin-en-Yvelines

PRiSM Laboratory. The research teams of the PRiSM laboratory (Parallélisme, Rseaux, Systèmes et Modélisation) are involved in two main scientific themes of UVSQ: Mathematics and Computer science on one hand, “Design, Modelization and Implementation of Systems” on the other hand. These two directions are not separated from each other, as shown by many collaborations with other labs, and the participation of many PRiSM teams to both directions. Within the “Mathematics and Computer Science” theme, the PRiSM teams study cryptology and security, models for algorithms and operational research. All the teams also participate to the “Design, Modelization and Implementation of Systems” theme, with a particular focus on communication systems (networks and telecommunication), embedded systems, mobile systems, high speed networks, and database systems.

PRiSM is home to the “Cryptology and Information Security”. In its research activities, the cryptography team aims at widely covering the various themes of academic research in cryptology, public key and secret key cryptography, cryptanalysis, security of implementations, number theory, multivariate cryptography, hash functions, etc. The cryptology team brings its specificity in the computer science courses at UVSQ and, since several years, the university offers several teaching programs with a part devoted to cryptology and information security. In particular, the research graduate program “Applied Algebra” offers a full course in cryptology. It has been complemented by a professional graduate program, called SeCReTS (Security of Contents, Networks, Telecommunications and Systems). Many activities of the team, require the use of advanced computer algebra. For this, the team has a long history of using computer algebra systems (GAP, Pari, Maple, Magma, ...). In recent years, with the arrival of young researchers, and with the affirmation of Sage in research and teaching, the team has moved from a pure user perspective to a contributor one, taking active part in the development of computer algebra software.

Curriculum vitae of the investigators

Luca De Feo got his PhD in 2010 at Ecole Polytechnique. He was appointed Matre de Confrences at Versailles-St-Quentin-en-Yvelines University in 2011. His research interests cover Algorithmic Number Theory, Computer Algebra, Cryptology and Automated deduction, and he has already published 8 papers in international journals or refereed international conferences.

He is an active Sage contributor, with a dozen of tickets co-authored and about as much reviewed. He is also active in promoting the use of Sage for research and for teaching: most of his papers feature a publicly available Sage implementation, he teaches Sage to undergraduate and graduate students, he participates and organizes various events for the introduction of Sage to beginners and young researchers.

Publications, achievements

Recent publications:

1. L. De Feo, J. Doliskani, . Schost; Fast arithmetic for the algebraic closure of finite fields. ISSAC ’14. ACM, 2014. pp 122-129.

2. L. De Feo, . Schost; Transalpyne: a language for automatic transposition. ACM SIGSAM Bulletin, 2010, 44 (1/2), pp. 59-71.

Software:

1. FAAST, a C++ library for Fast Arithmetic in Artin-Schreier Towers. http://github.com/defeo/FAAST.

Previous projects or activities

Current grants:

1. ANR CLE (2013-2017): Cryptography from Learning with Errors.

2. DIGITEO project ARGC (2013-2016): “Fast arithmetic for geometry and cryptology”. The project explores fast algorithms and implementations for algebraic geometry and curve-based cryptography.
3. DIGITEO project IdealCodes (2014-2016): IdealCodes (http://idealcodes.github.io/) spans the three research areas of algebraic coding theory, cryptography, and computer algebra, by investigating the problem of lattice reduction.

♠MK [**WRITE HERE**: Description of University of Silesia]♠ ♠MK [**WRITE HERE**: PIC for University of Silesia]♠ ♠WS [**WRITE HERE**: Description of University of Washington, CV]♠

♠**TO DO**: *Reorder partners accordingly to the list on the front page?*♠

[**science-project**]

B.0.4.2 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?*♠

B.0.5 Ethics and Security

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

B.0.5.1 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♠

B.0.5.2 Security

Please indicate if your proposal will involve:

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