



PCN - Horizon2020

HORIZON 2020

LE PROGRAMME DE RECHERCHE ET
D'INNOVATION DE L'UNION EUROPÉENNE

Big Data 2017 calls

ICT-14; ICT-15 ; ICT-16, ICT-17 and others

January 06th, 2017, Paris

Brokerage

General information

*Aix-Marseille Université (**AMU**)*

Cristinel Diaconu (Centre de Physique des Particules de Marseille)

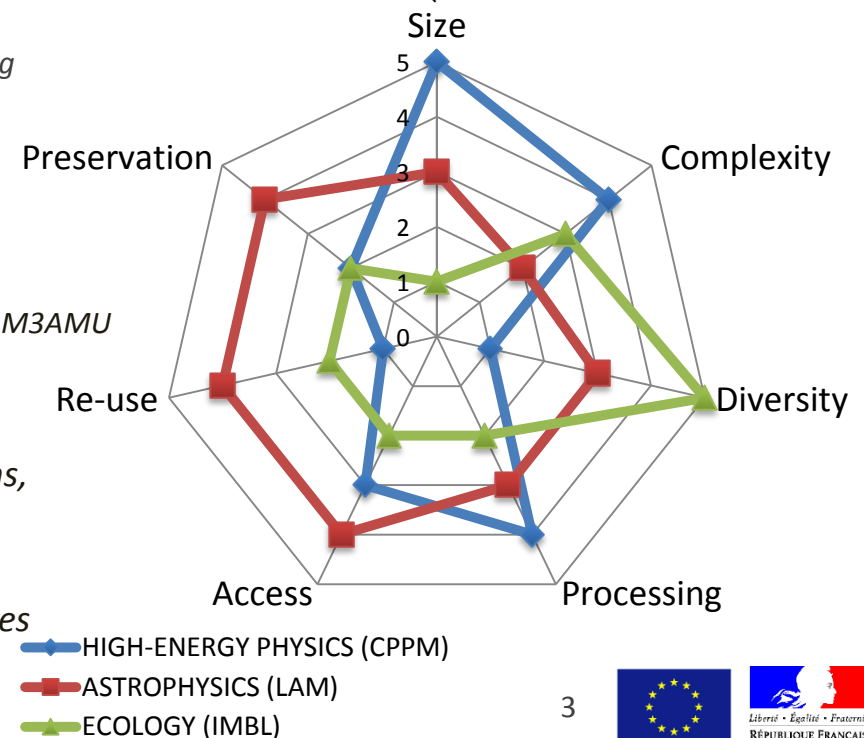
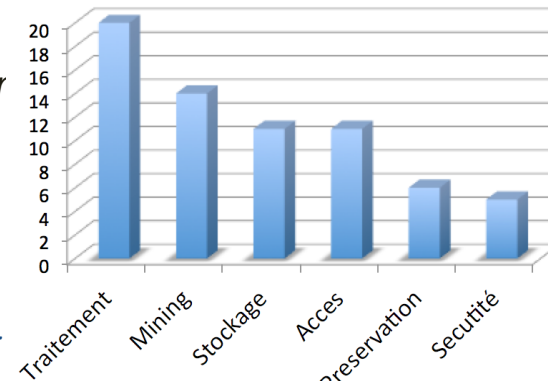
diaconu@cppm.in2p3.fr

+33 4 92 82 76 70 / +33 6 72 69 28 80

Area of interest	Choose Y or N
ICT-14: Big Data PPP: Cross-sectorial and cross-lingual data integration and experimentation A) Data integration activities	Y
ICT-14: Big Data PPP: Cross-sectorial and cross-lingual data integration and experimentation B) Data experimentation incubators	Y
ICT-15-2016-2017: Big Data PPP: Large Scale Pilot actions in sectors best benefitting from data-driven innovation Large Scale Pilot actions	Y
ICT-16-2017: Big data PPP: research addressing main technology challenges of the data economy cross-sector and cross-border problems or opportunities of clear industrial significance.	Y
ICT-17-2016-2017: Big data PPP: Support, industrial skills, benchmarking and evaluation Benchmarking action	Y
Others:	Y

Competencies

- **Organisation competencies:**
 - 30 laboratories involved in a transverse coordination around “big data”
 - Large spectrum of projects concerning large/complex amounts of data
 - Contact hub to future projects for AMU
- **Organisation experience in the European project**
 - Structural projects in big data
 - EGI European Grid, Distributed/HPC computing
 - Rare Genetics
 - Social science, Language, text analysis
 - + many more....
- **The skills you can bring:**
 - Large computing infrastructures: Data center/ M3AMU
 - Big data analytics:
 - specialised teams and large pools of projects
 - Very large panel of scientific data, algorithms, applications in many scientific fields
 - Cross-fertilisation theory/data/applications
 - Links with local enterprises/competitvity poles



Strengths

**Data
Mining**

Theoretical basis

Large/complex data sets

Tools and frameworks
Complementarity

Data preservation

Data-bases in many
disciplines

Massive and long-term
storage

**Access
Preservation**

**Usage,
ethics,
legal issues**

Society impact of open/big
data

Epistemology and ethics

Media studies

Large infrastructures

Massive computing

Cutting edge technology
(virtualisation, cloud...) etc.)

Infrastructure

Needs

Data Mining

Trans/Inter disciplinary cooperation

Advanced and Student training programs

Common centrally managed resources

Multi-disciplinary approach

Common Platforms for expertise exchange and cooperation

Access Preservation

Usage, ethics, legal issues

SHS expertise mutualisation

Interdisciplinary cooperation

Training

New topics: security psycho-social and use impact

Inter-connectivité, distributed ressources , basic infrastructure

Training

Experts in data managment

« Big data »-like expertise

Infrastructure

Opportunities

Data Mining

Collaboration on methods and algorithms across disciplines

Common projects on themes across disciplines (examples: e-pub, biodiversity, health, social networks, urban analytics etc.)

Enhance the links with regional authorities

Common databases and their usage (language, bioinformatics, etc.)

Develop common frameworks

Scientific data preservation

Heterogeneous data integration

Access Preservation

Usage, ethics, legal issues

Impact on management

Ethics and society impact of big data
(example: personal genomics)

Legal issues of big data: medial data, security, intellectual property

AMU scientific data infrastructure

Infrastructure

M³AMU

Mésocentre Multi-Modalités in AMU

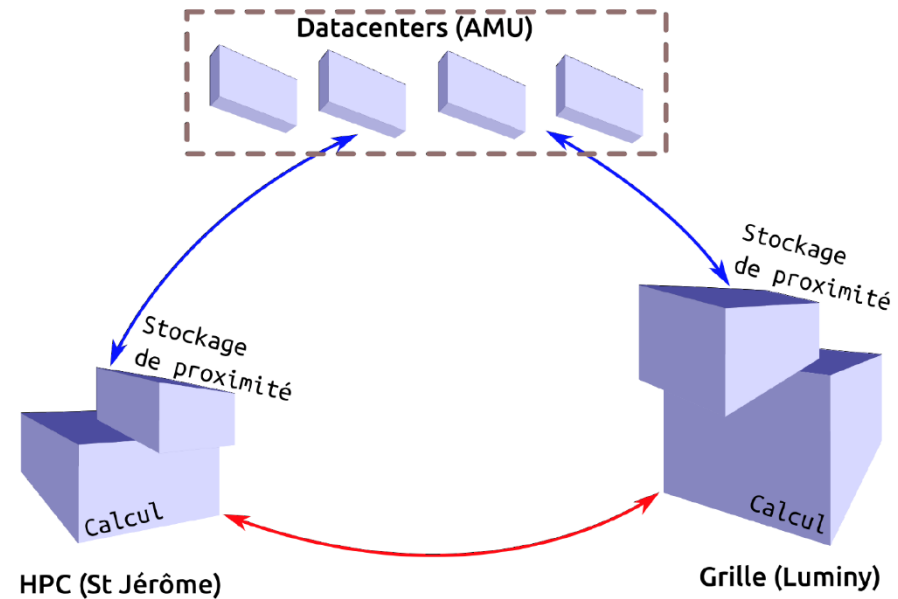
HPC, grid, cloud, storage for scientific computing and Big Data

Project M3AMU funded by CPER/FEDER/CD13 (2016 – 1018) :

- Update + extension of the resources: +50 % more powerful
- Permanent storage close to the computing nodes
- Interface to the CPPM grid
- Cloud computing



Choosing an infrastructure for massive data: storage & processing?



RHETICUS > 20 Tflops
1300 cores – 2.3 Tb RAM
+3.5 Tb of shared RAM



1958 cores on 98 servers
900 Tb stored on 18 servers
10 Gb/s network: LHCONE

Project idea

- *Describe your project idea*
 - *AMU “Big Data” : hub to large number pf competencies*

- *List of the complementary skills you need for your consortium*
 - *Collaborations with academic/industrial partners on specific projects to be developed*