

NBIC Galaxy to Strengthen the Bioinformatics Community in the Netherlands

Hailiang Mei

(hailiang.mei@nbic.nl)

David van Enckevort

(david.van.enckevort@nbic.nl)



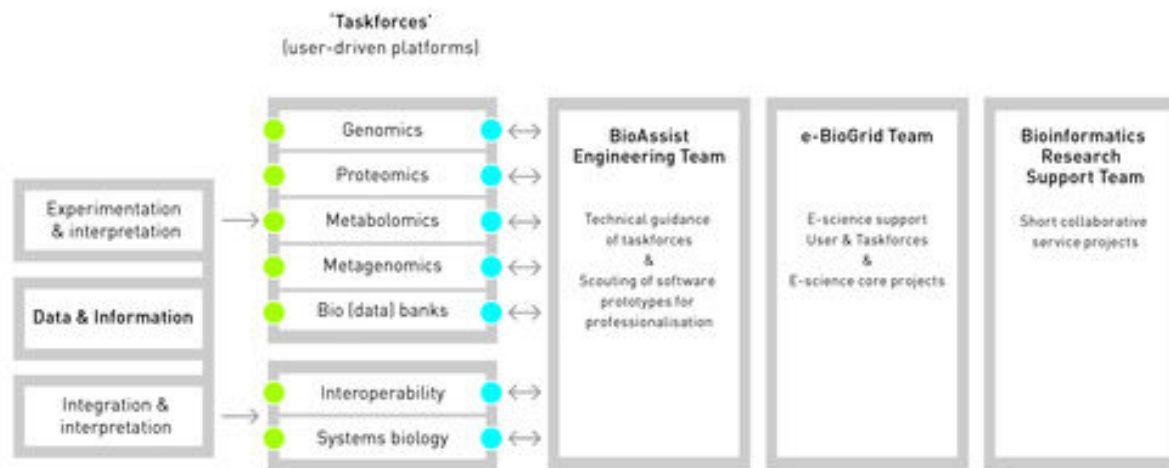
Introduction

- Working for NBIC since 2010
 - Scientific programmer
 - Technical project leader for the NBIC Biobanks Task Force
- Worked for 11 years as developer and system administrator at several companies
- Classical Archaeologist by education



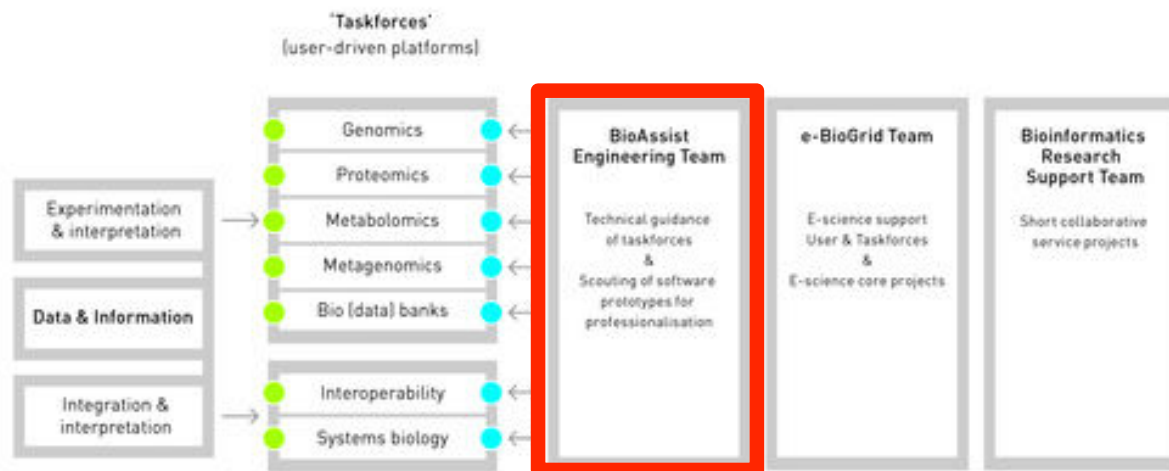
About NBIC BioAssist

- One of the 3 pillars of NBIC (next to BioWise (education) and BioRange (research)).
- Goal is to *make other people's data work*



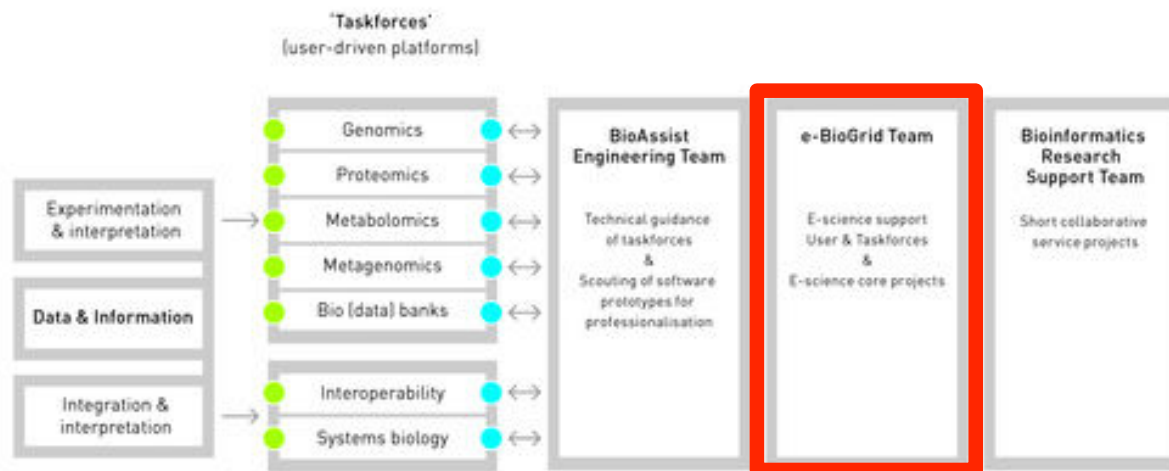
About NBIC BioAssist

- One of the 3 pillars of NBIC (next to BioWise (education) and BioRange (research)).
- Goal is to *make other people's data work*
- Consists of a **Central Engineering Team**



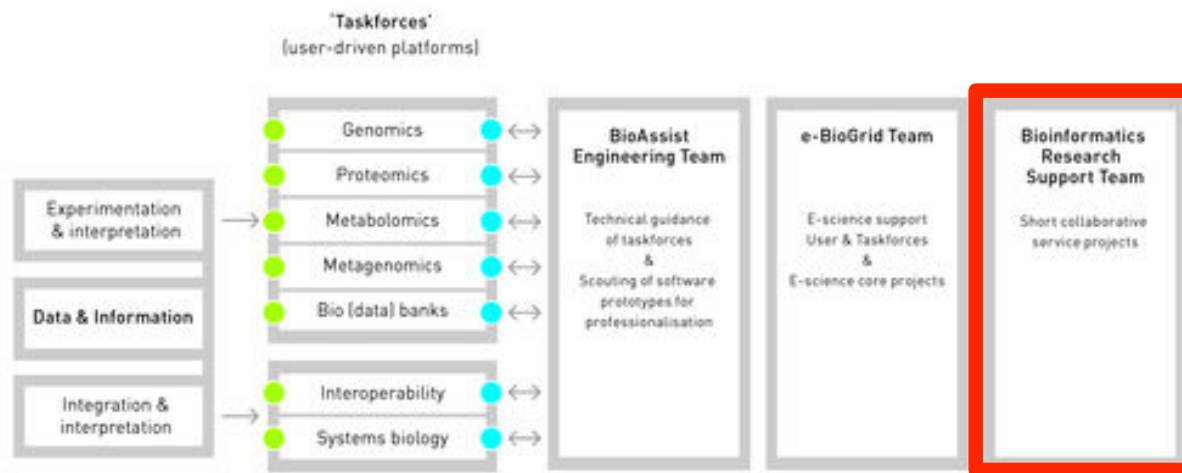
About NBIC BioAssist

- One of the 3 pillars of NBIC (next to BioWise (education) and BioRange (research)).
- Goal is to *make other people's data work*
- Consists of a Central Engineering Team, **eBioGrid**



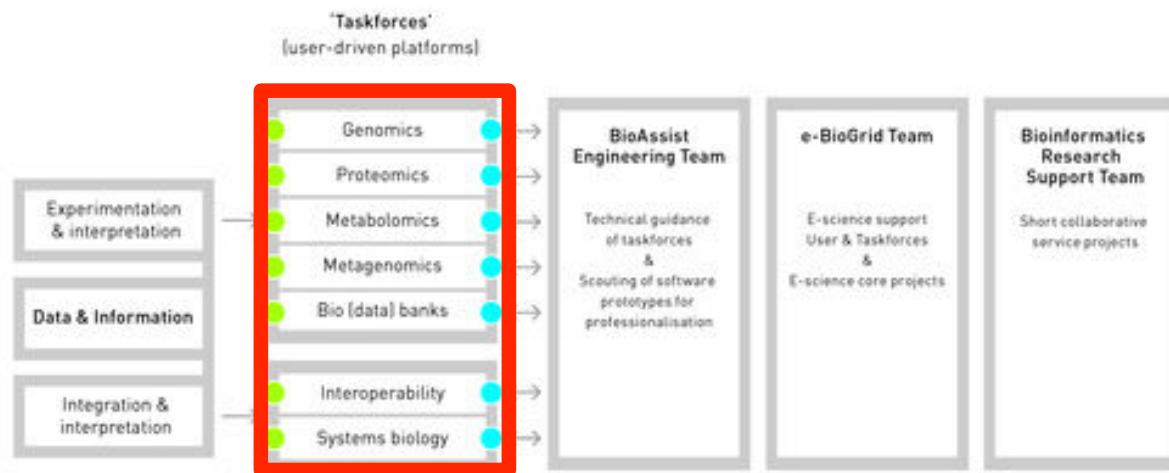
About NBIC BioAssist

- One of the 3 pillars of NBIC (next to BioWise (education) and BioRange (research)).
- Goal is to *make other people's data work*
- Consists of a Central Engineering Team, eBioGrid, **Bioinformatics Research Support**



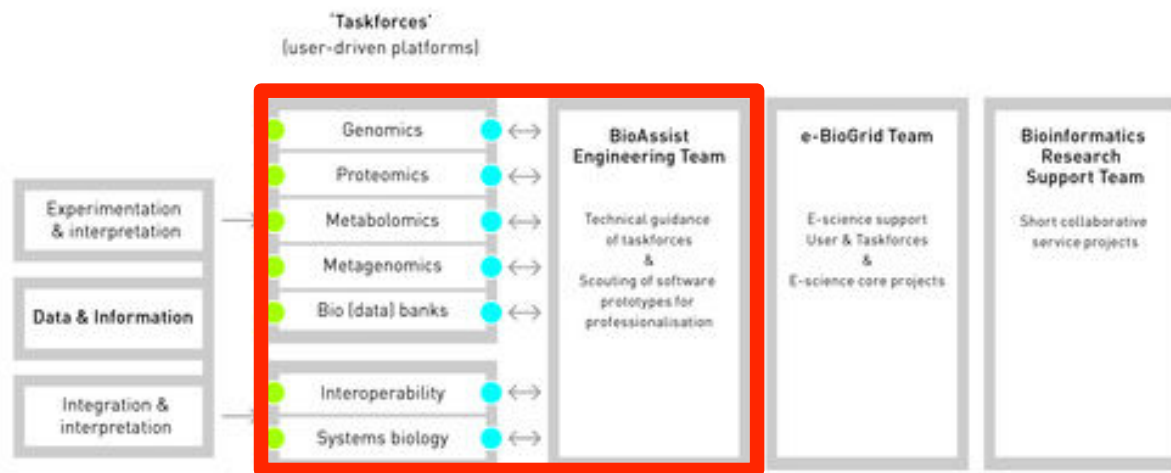
About NBIC BioAssist

- One of the 3 pillars of NBIC (next to BioWise (education) and BioRange (research)).
- Goal is to *make other people's data work*
- Consists of a Central Engineering Team, eBioGrid, Bioinformatics Research Support and **Task Forces**



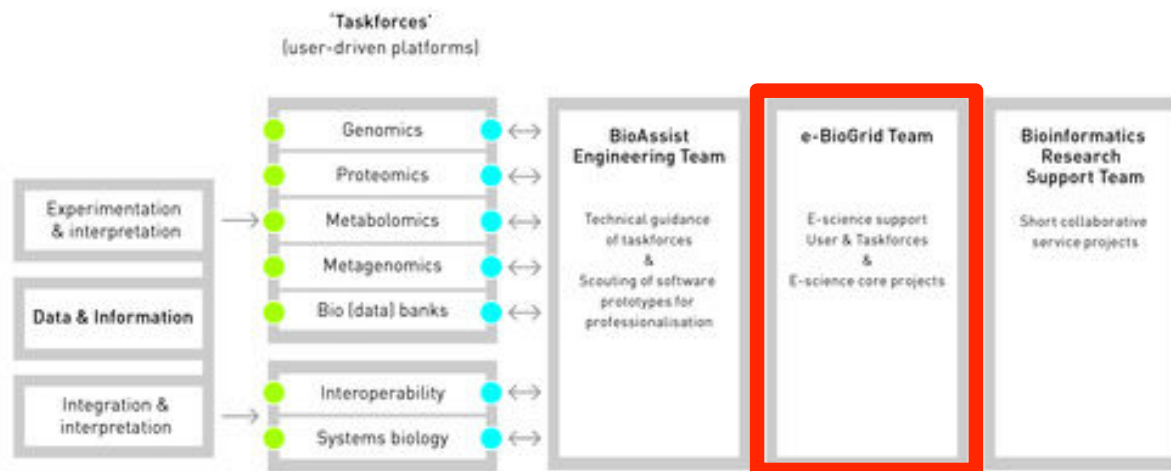
Where does Galaxy fit in?

- Training and demonstration
- Collaboration in several consortia (e.g. CTMM TraIT)
- Sharing of NBIC developed tools and pipelines



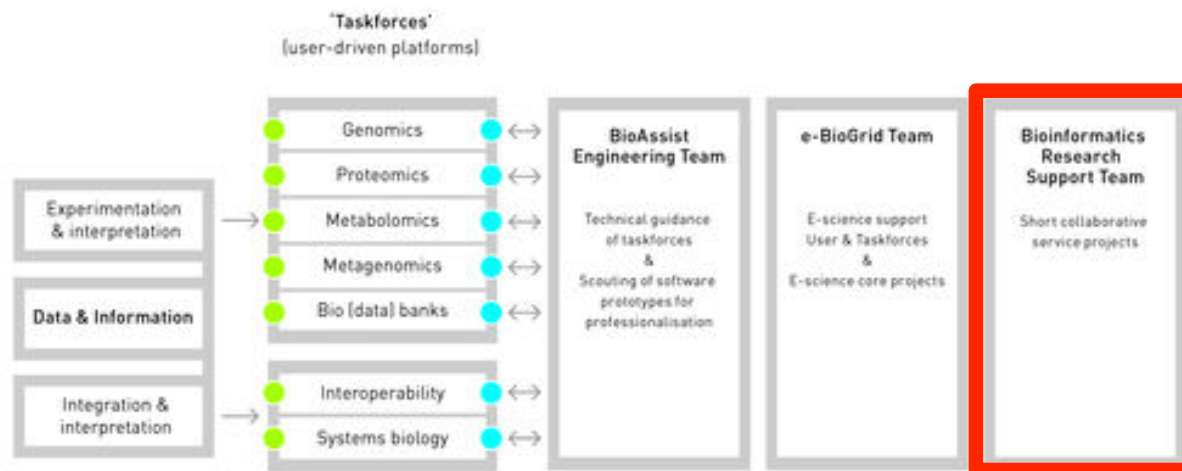
Where does Galaxy fit in?

- Enabling scientists to run analysis on national HPC infrastructures.



Where does Galaxy fit in?

- Provisioning of tools developed by BRS
- Persistence of tools and publications



What do we have?

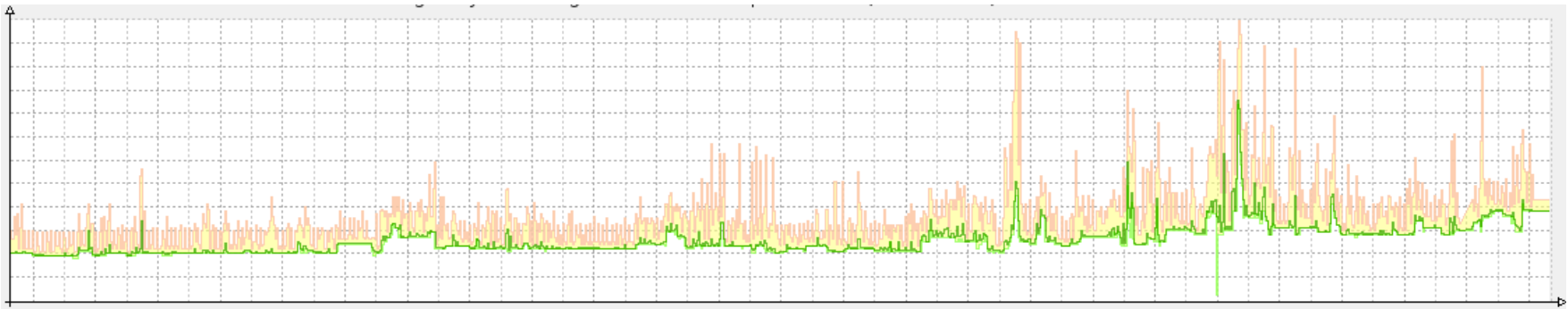
- Running several galaxy instances (Task forces, BRS)
- Running a public instance (<http://galaxy.nbic.nl/>)
- Host a public subversion repository with several Galaxy related tools (<https://trac.nbic.nl/>)
- Galaxy VM (https://wiki.nbic.nl/index.php/Galaxy_VM)
- Wiki (https://wiki.nbic.nl/index.php/Category:NBIC_Galaxy)

Galaxy.nbic.nl

- 8 cores, 24GB RAM, 1.5TB disk
- 10GB quota for users, extendable upon request
- Pipelines and Tools from NBIC & partners
 - NGS, Proteomics, etc.
- Flexible access:
 - FTP, API access
- Extended from demo to production level in 2011

Usage

- 451 registered users
- 900GB data
- Steady growth in usage



(# processes over the last year)

Some of the NBIC Tools

- Benchmarking Wrapper – benchmarks several alignment and denovo assembly pipelines (Bowtie, BWA, Abyss, Velvet)
- MS compare – proteomics benchmarking
- Taverna workflows – run Taverna workflows from within Galaxy
- Genome of the Netherlands Alignment pipeline (planned)
- And many more...

Available at <https://trac.nbic.nl/>

More

- CLI-Mate – generate tool wrappers (LUMC)
(<http://cli-mate.lumc.nl/>)
- Courses, e.g.:
 - GAPSS tutorial (LUMC)
 - RNA-Seq data analysis (NBIC BioWise)
- Galaxy hackathon (together with the main Galaxy Team – Sorry, I should have brought some stroopwafels for you guys)
- Organized the 2011 GCC

Near Future: Galaxy at the SARA HPC Cloud

- HPC Cloud: 608 cores 4.75TB RAM, 400TB shared resource provided by SARA
- Galaxy installed with CloudMan and the NBIC Galaxy Tools
- Auto scaling
- BigGrid project in collaboration with SARA (Niek Bosch), Netherlands Institute for Ecology (Mattias de Hollander)

Expected to be available Mid 2012

A graphic consisting of several interlocking gears of various colors (blue, green, yellow, pink, purple) arranged in a cluster, with a few small dots scattered around them.

BiG Grid

the dutch e-science grid

nbic Netherlands
Bioinformatics
Centre

Acknowledgements

NBIC

Leon Mei

Freek de Bruijn

Rob Hooft

Mark van Driel

Netherlands Institute of Ecology (NIOO-KNAW)

Mattias de Hollander

Erasmus MC

Saskia Hiltemann

Galaxy Admin Team

Pieter Neerincx (UMCG)

Jeroen Laros (LUMC)

Ishtiaq Ahmad (AMC/RUG)

SARA

Niek Bosch

Machiel Jansen

eBio-Grid

BigGrid

CTTM TraIT