

# The Czech Republic and ELIXIR

Pavel Fibich, et al.

Czech NGI

7. 11. 2012



Luděk Matyska (CESNET, **CERIT-SC**)

Miroslav Ruda (**CESNET**, CERIT-SC)

Ivana Křenková (CESNET, CERIT-SC, **Czech NIL in EGI**)

Aleš Křenek (CESNET, CERIT-SC)

Michal Jaroš (CERIT-SC)

Pavel Fibich (CESNET, **chair of VT ELIXIR**)

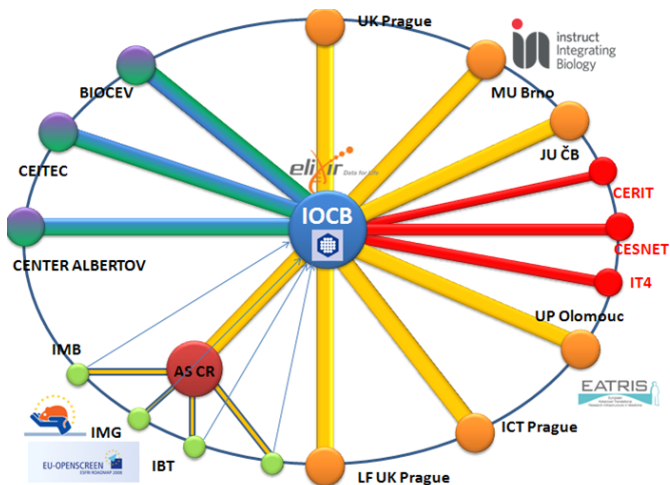
- ▶ postdoc in Botany
- ▶ user and middleware support in NGI/EGI

CESNET (NREN, NGI),

CERIT-SC (Resource provider/experiments)

- ▶ a way how to organize Czech institutions involved in ELIXIR
  - ▶ Life science and e-infrastructure organizations
- ▶ Headquarters at IOCB
  - ▶ current lead representative Jiří Vondrášek
- ▶ currently informal consortium, formal setup under way
  - ▶ consortial agreement under preparation
  - ▶ expected before end of the year
- ▶ Ministry of Education already **signed MoU** with ELIXIR on behalf of Czech Republic (and ELIXIR\_CZ)

e-infrastructures, universities, projects, academy of science



- ▶ Organizational aspects
  - ▶ covered through ELIXIR\_CZ consortium
  - ▶ expected an equal representation by member organization
  - ▶ other details under discussion
  - ▶ ELIXIR\_CZ meetings primary on formal aspects of the collaboration
- ▶ Technical aspects
  - ▶ national node proposal preparation and submission
  - ▶ implementation of the national node
- ▶ Event
  - ▶ A seminar of ELIXIR\_CZ members on collaboration with e-Infrastructures
    - ▶ Planned late November/early December
    - ▶ to discuss more technical aspects of the collaboration
    - ▶ compute and data requirements and architecture discussion

- ▶ Proposal prepared by 7 ELIXIR\_CZ consortium members
  - ▶ application submitted to ELIXIR HQ
  - ▶ both Life science and e-infrastructure organizations involved
    - ▶ e-infrastructures: CESNET and CERIT-SC
    - ▶ universities and projects: CEITEC, UPOL, ICT, JU, and IOCB
- ▶ ELIXIR\_CZ members with strong contribution to the ELIXIR node and immediate international ELIXIR-related collaboration

- ▶ at IOCB, small compute and data node dedicated to ELIXIR\_CZ
- ▶ hardware provided by CESNET
  - ▶ front-end (24 cores, 128GB RAM, 2x600GB disks)
  - ▶ compute node (48 cores, 512GB RAM, 2x600GB disks)
  - ▶ storage node (16 cores, 128GB RAM, 100TB in raid5)
- ▶ system administration and advanced IT support provided by CESNET and CERIT-SC
- ▶ additional flexible (cloud) compute and data resources are available both at MetaCentrum and CERIT-SC

Ongoing technical collaboration.

- ▶ Work on further refinement of ELIXIR\_CZ node IT architecture
  - ▶ Bioinformaticians and e-infrastructure experts jointly involved
  - ▶ Extending the node facilities through grid and cloud environment
  - ▶ Cloning the successful node configuration on the cloud
- ▶ Helping to shape the national e-infrastructure landscape
  - ▶ All major players are involved
  - ▶ Preparation for full data flow expected in 18 months



Three parts, together covering all aspects

- ▶ CESNET
  - ▶ NREN, backbone of academic computer network
  - ▶ NGI MetaCentrum, Grid coordination, also resource provider
  - ▶ data, very new aspects, large scale (>10 PB) data repositories
- ▶ CERIT-SC
  - ▶ Grid resource provider
  - ▶ In silico experimental facility
  - ▶ compute&data in novel configurations
- ▶ IT4Innovations
  - ▶ PRACE partner, supercomputing centre
  - ▶ High end national compute facilities
  - ▶ Under construction, first resources 1H 2013

- ▶ Human-genome data focused on haematological and oncological malignancies
- ▶ Genome data on plants, humans and other organisms
- ▶ Tools for human-genome data analysis
- ▶ Tools for protein structural motif analysis (MOLE, MOLEonline)
- ▶ Non-coding genome sequences and genetic network tools (Repeat Explorer)
- ▶ Tools for protein-stability prediction and protein-mutation analysis
- ▶ Tools for pump-probe simulations

The data proposed to contribute to the community

- ▶ database of small molecules developed for therapy
- ▶ human and plant genome data from next-gen sequencing (NGS) projects
- ▶ a database of biologically important protein structural motifs

The tools proposed to contribute

- ▶ tools for the location, comparison and characterization of biologically important protein structural motifs
- ▶ an interaction Energy Matrix (IEM) tool for protein-stability prediction and analysis
- ▶ a pump-probe simulation server
- ▶ Repeat Explorer for analysis of transposable-element protein-coding domains
- ▶ tools for NGS data analysis

## Technical cooperation

- ▶ Potential for some pilot experiments
  - ▶ among two or more interested VT ELIXIR members
  - ▶ demonstrate technical issues common for several countries/teams
- ▶ Suggestions coming from national presentations
  - ▶ try to mention potentially interesting subjects from your side
- ▶ We can offer out human resources and know-how for selected experiment
  - ▶ will help with pilot experiment(s) formulation

▶ EGI VT ELIXIR

▶ [https://wiki.egi.eu/wiki/VT\\_ELIXIR](https://wiki.egi.eu/wiki/VT_ELIXIR)

▶ CESNET

▶ <http://www.ces.net>

▶ MetaCentrum

▶ <http://www.metacentrum.cz/>

▶ CERIT-SC

▶ <http://www.cerit-sc.cz/en/>

▶ ELIXIR\_CZ

▶ <http://www.elixircz.cz>

▶ IOCB

▶ <http://www.uochb.cz/>