

# Architektura dużych projektów bioinformatycznych

Bartek Wilczyński

[bartek@mimuw.edu.pl](mailto:bartek@mimuw.edu.pl)

<http://regulomics.mimuw.edu.pl>

**Wykład 2. - Projekty BioX**

10. X. 2016

# Plan na dziś

- Bioperl i trochę historii
- Open Bioinformatics Foundation
- Software for Public Interest
- Biopython, Biojava, bioruby
- BioSQL, BioDAS, EMBOSS
- Bioinformatics Open Source Conference
- Google Summer of Code
- Przegląd aktywności projektów
- Biopython trochę bardziej z bliska

# BioPerl

- Założony około roku 1996
- Pierwsze “wydanie” 0.7 – 2000
- Wersja 1.0 – 2002
- 2002 - The Bioperl Toolkit: Perl Modules for the Life Sciences JE. Stajich et al. *Genome Res*
- 2002 – powstaje Open Bioinformatics Foundation
- Do 2010 niesłabnąca popularność

# Open Bioinformatics Foundation

- Zarządzanie projektami, hosting, storage, backups, etc.
- Uformowana w 2001/2002 na potrzeby projektów bioperl/piopython/biojava
- W 2005 zyskała formalną “osobowość” jako fundacja not-for-profit
- W 2012 przekazano prawa do zarządzania fundacją na rzecz Software in the public interest

# Software in the public interest

- Organizacja Not for profit
- Zajmuje się projektami open source, nie tylko bioinformatycznymi (Haskell, ffmpeg, Debian,...)
- Zbiera wpłaty, podpisuje umowy, oferuje pomoc prawną
- Demokratyczne zarządzanie
- Pomaga organizować sprzęt i usługi dla utrzymania projektów
- “Freedom to leave”

# Różne projekty BioX

- BioPerl
- BioPython
- BioJava
- BioSQL
- BioRuby
- BioHaskell
- BioJS
- ....

# Rozwój BioPerl'a

GitHub

This repository Search

Explore Features Enterprise Blog

Sign up

Sign in



bioperl / bioperl-live

★ Star 131

Fork 81

Contributors

Commits

Code frequency

Punch card

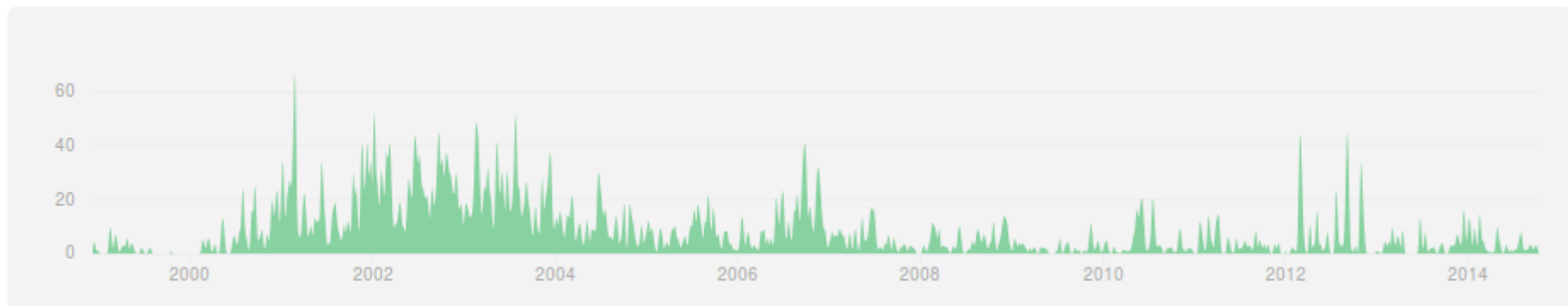
Network

Members

Dec 6, 1998 – Oct 13, 2014

Contributions: **Commits**

Contributions to master, excluding merge commits



# Użycie BioPython'a

GitHub

This repository Search

Explore Features Enterprise Blog

Sign up

Sign in

biopython / biopython

★ Star 505

🍴 Fork 300

Contributors

Commits

Code frequency

Punch card

Network

Members

Dec 5, 1999 – Oct 13, 2014

Contributions: **Commits** ▾

Contributions to master, excluding merge commits





# Użycie Biojava

GitHub

This repository Search

Explore Features Enterprise Blog

Sign up

Sign in

biojava / biojava

★ Star 46

Fork 68

Contributors

Commits

Code frequency

Punch card

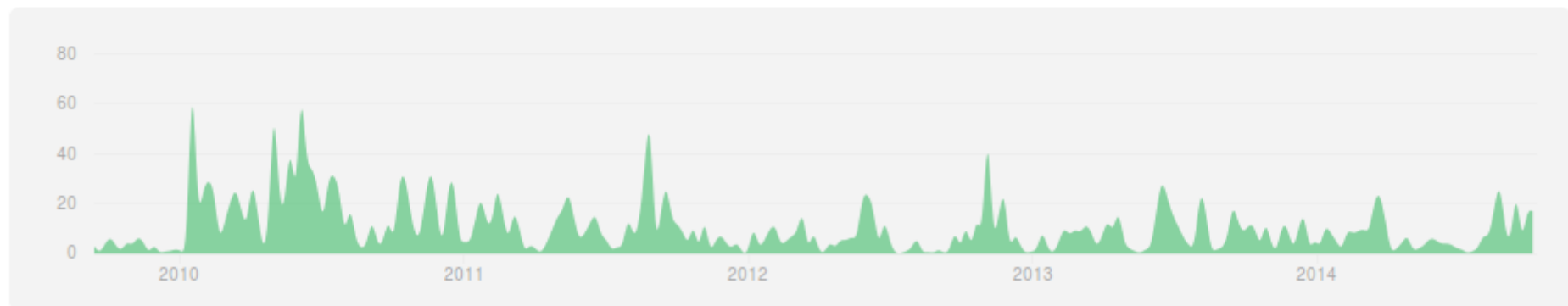
Network

Members

Sep 13, 2009 – Oct 11, 2014

Contributions: **Commits** ▾

Contributions to master, excluding merge commits



# Bioruby...

GitHub

This repository Search

Explore Features Enterprise Blog

Sign up

Sign in



bioruby / bioruby

★ Star 248

🍴 Fork 100

Contributors

Commits

Code frequency

Punch card

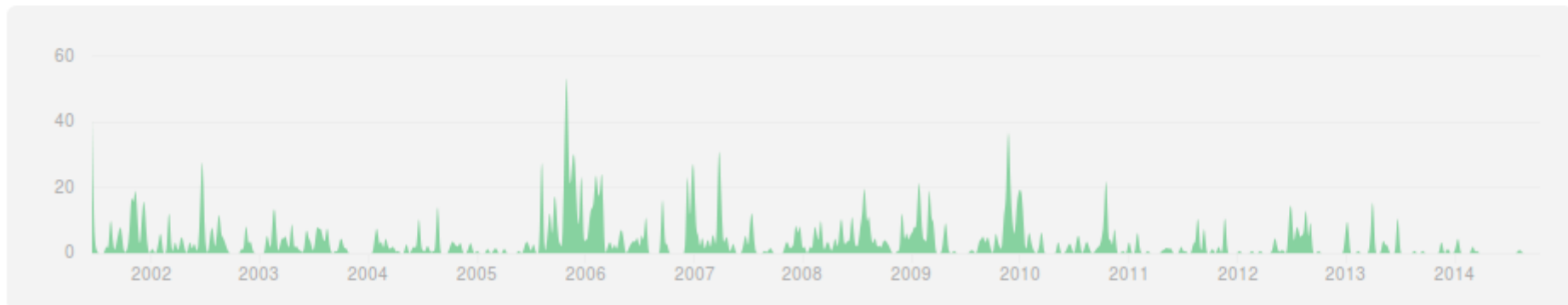
Network

Members

Jun 17, 2001 – Oct 13, 2014

Contributions: **Commits** ▾

Contributions to master, excluding merge commits



# Inne projekty BioX

- Bio-DAS – Distributed Annotation System
- Bio-SQL – serialization in SQL
- Bio-MOBY – index of webservices
- GMOD – generic model organism database

# Bioinformatics Open Source Convention

[Log in / create account with Open](#)

O|B|F

Page [Discussion](#)

[Read](#)

[View source](#)

[View history](#)

## BOSC 2014



The 15th Annual Bioinformatics Open Source Conference (BOSC 2014) was held at the Hynes Convention Center in Boston on July 11-12, right before [ISMB 2014](#) as one of the official ISMB SIG satellite meetings. Feedback welcome via email [bosc@open-bio.org](mailto:bosc@open-bio.org) or [@OBF\\_BOSC on Twitter](#).

[Complete Schedule](#) (with links to slides and videos of the talks)

### Media

Storify Tweet archives for [Day One](#), [Day Two](#) and [during ISMB](#).

[Brad Chapman's blog posts](#)

[BioInform article about Philip Bourne's keynote talk](#) (posted with permission)

[James Malone's blog post](#)

### Important Dates

- July 9-10, 2014: [Codefest 2014](#), Boston, MA
- July 11-12, 2014: BOSC 2014, Boston, MA
- July 12-16, 2014: [ISMB 2014](#), Boston, MA
- September 18-19, 2014: [EU-Codefest 2014](#), Cambridge, UK

### Overview

The Bioinformatics Open Source Conference (BOSC) is run as a two-day meeting before the annual ISMB conference. It is sponsored by the Open Bioinformatics Foundation (O|B|F), a non-profit group dedicated to promoting the practice and philosophy of Open Source software development and Open Science within the biological research community.

Open Source software has flourished in the bioinformatics community for well over a decade. When the first BOSC (Bioinformatics Open Source Conference) was held in 2000, there were already a number of popular open source bioinformatics packages, and the number and range of these projects has increased dramatically since then. covers the wide range of open source bioinformatics software being developed, and encompasses the growing movement of Open Science, with its focus on transparency, reproducibility, and data provenance. We welcome submissions relating to all aspects of bioinformatics and open science software, including new computational methods, reusable software components, visualization, interoperability, and other approaches that help to advance research in the biomolecular sciences. Two full days of talks, posters, panel discussions, and informal discussion groups will enable BOSC attendees to interact with other developers and share ideas and code, as well as learning about some of the latest developments in the field of open source bioinformatics.

Please spread the word about BOSC--all are welcome. On Twitter, follow [@OBF\\_BOSC](#) and use hash tag [#bosc2014](#).

#### Navigation

[Main Page](#)  
[Projects](#)  
[Newsletters](#)  
[News](#)  
[OBF Board](#)  
[Participate!](#)  
[Current events](#)  
[Recent changes](#)  
[Random page](#)  
[Help](#)

#### Toolbox

[What links here](#)  
[Related changes](#)  
[Special pages](#)  
[Printable version](#)  
[Permanent link](#)

# Google summer of Code

- Organizowany przez google od 2005 roku
- Zbiera potencjalne projekty, mentorów i oferuje studentom niewielkie wynagrodzenie za zaangażowanie w projekt open Source
- W sumie ponad 8500 uczestników, 55 milionów linii kodu
- OBF czynnym uczestnikiem od 2009

# Biopython

- Biopython.org
- Źródła na githubie
- Narzędzia do:
  - analizy sekwencji
  - Serializacji,
  - PDB,
  - text-mining,
  - Grafika,
  - ...

# Ważne kwestie dot. Biopythona

- Jak kod projektu jest zarządzany na githubie
- Contributing howto
- Git-usage
- Testing framework
- Mailing lists (biopython, biopython-dev)
- Documentation guidelines
- licenses