ToolDog — generating tool descriptors from the ELIXIR tool registry

Kenzo-Hugo Hillion, Ivan Kuzmin, Hedi Peterson, Jon Ison, Hervé Ménager

Galaxy Community Conference, 30th June 2017, Montpellier, France
Current issues

Tool
Current issues

Registry or Repository

- Increase visibility and discoverability
- Usually simple and done by tool developer
- Examples:
  - Pypi (Python)
  - Bioconductor (R)
  - Bioconda (any tool)
  - bio.tools (any tool)
Current issues

**Registry or Repository**
- Increase visibility and discoverability
- Usually simple and done by tool developer
- Examples:
  - Pypi (Python)
  - Bioconductor (R)
  - Bioconda (any tool)
  - bio.tools (any tool)

**Workbench environments**
- Increase availability and reproducibility
- Workflow construction and sharing
- **Time consuming and not straightforward**
- Done either by tool developers or workbench environments administrators
- Examples:
  - Galaxy
  - Environments using CWL
Current issues

Registation

Tool

Tool

Integration

Tool

Registry or Repository

- Increase visibility and discoverability
- Usually simple and done by tool developer
- Examples:
  - Pypi (Python)
  - Bioconductor (R)
  - Bioconda (any tool)
  - bio.tools (any tool)

Workbench environments

- Increase availability and reproducibility
- Workflow construction and sharing
- **Time consuming and not straightforward**
- Done either by tool developers or workbench environments administrators
- Examples:
  - Galaxy
  - Environments using CWL
Current issues

Registry or Repository
- Increase visibility and discoverability
- Usually simple and done by tool developer
- Examples:
  - Pypi (Python)
  - Bioconductor (R)
  - Bioconda (any tool)
  - bio.tools (any tool)

Workbench environments
- Increase availability and reproducibility
- Workflow construction and sharing
- Time consuming and not straightforward
- Done either by tool developers or workbench environments administrators
- Examples:
  - Galaxy
  - Environments using CWL
bio.tools registry

- Discovery portal for a wide range of key bioinformatics resources
bio.tools registry

- Discovery portal for a wide range of key bioinformatics resources
ToolDog

Web browser

ToolDog

Source code

Get EDAM mapping to datatypes

Galaxy server

bio.tools tool card

Resource description

html

json

django REST framework
ToolDog

**Source code**

Code inspection

**Docker**

- **Python**
  - argparse2tool*: Analysis of Python's code using argparse

*: [http://github.com/erasche/argparse2tool](http://github.com/erashe/argparse2tool)

**Covered**

**Not covered yet**
ToolDog

- Web browser
- django REST framework
- web tools tool card
- Resource description
- html
- json

Source code
Code inspection
Get EDAM mapping to datatypes
Galaxy server
CWL
Galaxy
Tool descriptions
Generated
Annotated
Perspectives

- Build a new common Tool model
- Restructure Python SDK for Galaxy XML (galaxyxml: https://github.com/erasche/galaxyxml) and CWL (python-cwlgen: https://github.com/common-workflow-language/python-cwlgen)
- Integrate ToolDog annotation part to Planemo
- Web interface for ToolDog
Contact & Acknowledgment

Please contact us at kehillio@pasteur.fr and hmenager@pasteur.fr

Github: https://github.com/bio-tools/tooldog
Contact & Acknowledgment

Please contact us at kehillio@pasteur.fr and hmenager@pasteur.fr

Github: https://github.com/bio-tools/tooldog

- ELIXIR-EE
  - Ivan Kuzmin
  - Hedi Peterson
- WINTER group at Institut Pasteur:
  - Hervé Ménager
  - Fabien Mareuil
  - Olivia Doppelt-Azeroual
- bio.tools dev team
  - Jon Ison
- Bioinformatics and Biostatistics Hub
- Collaborators:
  - Eric Rasche
  - John Chilton
  - Michael R. Crusoe
  - Daniel Blankenberg