The ALife Zoo: cross-browser, platform-agnostic hosting of Artificial Life simulations

Simon Hickinbotham, Michael Weeks & James Austin

University of York, Heslington, York YO1 5DD, UK
email: sjh518@york.ac.uk

August, 2013
Sharing / Distributing our simulations is good for science
Better to collect simulations in one place: “ALife Zoos”
**Web Resource:** https://portal.youshare.ac.uk
“Publish your computer code: it is good enough” (Nick Barnes, Nature, 2010)

Excuses for not doing so:

- **It is not common practice.**
  - This must change. Some disciplines, such as bioinformatics, are already changing.

- **People will pick holes and demand support and bug fixes.**
  - Publishing may see you accused of sloppiness. Not publishing can draw allegations of fraud. Which is worse?
  - Nobody is entitled to demand technical support for freely provided code: if the feedback is unhelpful, ignore it.

- **It is valuable intellectual property & belongs to my institution.**
  - Really, that little MATLAB routine to calculate a two-part fit is worth money? Frankly, I doubt it.
  - Some code may have long-term commercial potential, but almost all the value lies in your expertise.
  - My industry has a name for code not backed by skilled experts: abandonware.
  - Institutions should support publishing; those who refuse are blocking progress.

- **It is too much work to polish the code.**
  - For scientists, the word publication is totemic, and signifies perfectionism.
  - But your papers need not include meticulous pages of Fortran;
  - the original code can be published as supplementary information.
Some practicalities

- Often software is distributed to the community via personal, community or commercial websites.
- Compile-time errors, dependency errors and run-time errors aren’t unusual.
- Can we maintain source code or binaries that run on a range of operating systems?
  - the developer is often only experienced in writing software for personal use.
  - considerable expertise is required to compile binaries or install (development) software.
  - the code base becomes increasingly unwieldy.
If mishandled...

- It can promote the independent development of code bases.
- It can reduce the impact of the research.
- It can overburden a finite resource, namely the researcher’s time.
- If the research is interdisciplinary (e.g. ALife):
  - fewer skilled programmers than a pure computer science project
  - fewer domain experts than a pure biological project
  - research that uses more than one simulator is rare
Summary

- Sharing / Distributing our simulations is good for science
- Better to collect simulations in one place: “ALife Zoos”
- **Web Resource:** [https://portal.youshare.ac.uk](https://portal.youshare.ac.uk)
YouShare

- YouShare is an online collaborative facility that allows users to upload data, and software in the form of services.
- An attached execution environment allows services to be run over a heterogeneous cluster of compute nodes,
- the service infrastructure guarantees that the service will be executed in the correct environment, and provide consistent results.
- A workflow facility allows multiple services to be connected together to create more flexible evaluations.
YouShare Architecture

Distributed Storage System

Database

Web Server

Java Servlets

Compute Farm
Current prerequisites

- Command line application (non-interactive)*
- Outputs indicated with xml tags:
  - `<output>file.txt</output>`
- ANY OS (that can run Java)
- ANY Language
Specify service, inputs and outputs via web browser
job submitted to the “run service” servlets
servlet selects VM/Physical device
servlet stages service and inputs on device, starts process
...
On termination, outputs are stored in database with appropriate metadata
Deploying ALife Services

- Tierra version 6.02 (.C)
- Avida version 2.12.4 (.cpp)
- Stringmol version 0.2 (.cpp)
- Evolutionary Activity (.R)
Strategy

- wrap executable in a bash script to:
  - deploy a job as a single zip file (.tgz, .tar, .zip ...)
  - move files to specific directories
  - check errors
  - xml tags on output
- remove any user interaction (such as “press any key to continue”)
DEMO
Welcome to the YouShare portal - please view using the latest version of Internet Explorer, Firefox, Chrome, Safari or Opera.

The youShare VL allows users to share data and services with other users. It provides a secure online environment for:

- Storage of data.
- Storage and execution of analysis code.
- Storage of the metadata associated with data and code using standard descriptions.
- Storage of experimental protocols.
- Analysis of data.
- Controlled and secure sharing of data, analysis code, analysis results and experimental protocols, etc. with other chosen users e.g. sharing other members of a research group, sharing between different research groups, etc.
- Use of standard data exchange formats.
- Curation of data, analysis code and experimental protocols.
- Making new contacts in similar research areas.

First time users of the system can find a more detailed introduction on how to use the system here.

The youShare VL is funded by the UK HEFCE University Modernisation Fund.

Please report any problems to the support team.

Please read the Legal Statements.

How to use the system

The system provides access to data and services some of which are publicly available to view and download. If you are not "logged in" you will not be able to see any data, services or workflows.

Latest news

Workflows are introduced from April 2013.

more
Click on a tag to search the Data:

tom config research verbs aspergillus personas edwards subset jackson labelled basic metabolism jpg msc verb case sat pairs dataset ukwac semantics al evaluation choice tierra pair avida_ga_lim_res test parts surrogate web tao data carbon format niger edf model persona one face subject/object submitted experiments image lle qmsemantics word nouns project
Click on a tag to search the Data:

tom config research verbs aspergillus personas edwards subset jackson labelled basic metabolism jpg msc verb case sat pairs dataset ukwac semantics al evaluation choice tierra pair avida_ga_lim_res test parts surrogate web tao data carbon format niger edf model persona one face subject/object submitted experiments image lle qmsemantics word
nouns project
Service Name: Tierra Artificial Life Simulator with popdy outputs

Description: This version of tierra produces output in the popdy format in addition to the conventional tierra outputs. Tierra is a computer simulation developed by ecologist Thomas S. Ray in the early 1990s in which computer programs compete for central processing unit (CPU) time and access to main memory. In this context, the computer programs in Tierra are considered to be evolvable and can mutate, self-replicate and recombine. The basic Tierra model has been used to experimentally explore in silico the basic processes of evolutionary and ecological dynamics. Processes such as the dynamics of punctuated equilibrium, host-parasite co-evolution and density-dependent natural selection are amenable to investigation within the Tierra framework. A notable difference between Tierra and more conventional models of evolutionary computation, such as genetic algorithms, is that there is no explicit, or exogenous fitness function built into the model. Often in such models there is the notion of a function being optimized; in the case of Tierra, the fitness function is endogenous: there is simply survival and death. This is Tierra version 6.02, with the patch from Matthias Rav, http://tierra.lolwh.at/ The service takes a zipped folder as its input and produces a zipped folder as output. The input folder should contain: 1: a soup config file, specifying a subfolder where the remaining input files can be found 2: a subfolder containing, an opcode file (e.g. opcode.map from the tierra installation), and genome files for all genomes listed at the end of the soup file

Service Identifier: 235

You may add annotations here...
Execute Service > Service Inputs Output Location Run

Your service has successfully been submitted. Please see the service log for progress and results.

Done
The Tierra Artificial Life Simulator with popdy outputs selected service requires the following parameters setting - defaults are shown where provided.

Start: 2013-09-03 09:20:25.0, End: 2013-09-03 09:20:24.0
File Name: gb0.tgz
File Identifier: 350402
Uploaded: 2013-06-12 12:02:46.0
File Size: 92 KB

This item can be accessed externally using the following permanent link (provided sharing is set to 'Public' and the accessor has or creates an account):

The current sharing settings for the selected item are shown below.

Experimental data (system identifier: 350402 / LSID: URN:LSID:portal.carmen.org.uk:file:350333) sharing settings:
You may set the data visibility to other users below. Note that the existence of the data will always be visible under the experiment name (for example in search results) but you may control the viewing of the actual data below.

- Apply single sharing settings to data and metadata.

Metadata (system identifier: 34287 / LSID: URN:LSID:portal.carmen.org.uk:metadata:32012) sharing settings:
Here you may control the sharing of your metadata. Note that the existence of the data will always be visible under the experiment name (for example in search results) but you may control the viewing of the actual data below.

- Private
- Public
- Shared

(Only you will be able to view the metadata.)
(Anyone, even public users will be able to view the metadata.)
(Only you and other users you choose will be able to view the metadata.)
The current sharing settings for the selected item are shown below.

Experimental data (system identifier: 350402 / LSID: URN:LSID:porta.carmen.org.uk:file:350333) sharing settings:
You may set the data visibility to other users below. Note that the existence of the data will always be visible under the experiment name (for example in search results), but you may control the viewing of the actual data below.

- Apply single sharing settings to all data associated with metadata.
- Create full path to shared data files.
- Allow sharing users to add annotations to the data.

Metadata (system identifier: 34287 / LSID: URN:LSID:porta.carmen.org.uk:metadata:32012) sharing settings:
Here you control the sharing of your metadata. Note that the existence of the data will always be visible under the experiment name (for example in search results), but you may control the viewing of the actual data below.

- [ ] Private
  - Only you will be able to view the metadata.

- [ ] Public
  - (anyone, even public users will be able to view the metadata.)

- [ ] Shared
  - (Only you and other users you choose will be able to view the metadata.)

Type/suggest select users/groups

Shared with: [ALifeZoo [Artificial Life Simulator]]
[Alison Marshall [amarshall@carmen.ac.uk]]
[Alisdair Edwards [alis@cs.york.ac.uk]]

Portal version 3 Revision: 3046 Date: 2013-07-08 13:49:01 -0100 (Mon, 08 Jul 2013)
Some icons used under Creative Commons from http://flavanproject.com/.
An ALife Workflow
ALife workflow implemented in YouShare
Evolutionary Activity outputs
Future work

- on-line visualisation
- Deploying on HPC
- Link to external web server
- Expose services to other web apps
Strategic aims of the ZSL

1. To undertake and promote relevant high quality zoological and conservation research, to help us achieve our conservation objectives and to inform and influence conservation policy.

2. To encourage and motivate all our stakeholders to support and engage in conservation.

3. To implement and achieve effective and appropriate in situ and ex situ conservation programmes for priority species and habitats.

4. To lead and deliver exemplary Zoos in order to advance ZSLs mission.

5. To further ZSLs mission by maximising opportunities to generate funds.
Strategic aims of the ALZoo

1. To undertake and promote relevant high quality ALife research, to help us achieve our research objectives and to inform and influence research policy.

2. To encourage and motivate all our stakeholders to support and engage in ALife research.

3. To implement and achieve effective and appropriate in situ and ex situ ALife programmes for priority simulation examples.

4. To lead and deliver exemplary ALife Zoos in order to advance ISALs mission.

5. To further ISALs mission by maximising opportunities to generate funds.
Summary

- Sharing / Distributing our simulations is good for science
- Better to collect simulations in one place: “ALife Zoos”
- **Web Resource:** [https://portal.youshare.ac.uk](https://portal.youshare.ac.uk)
The ALife Zoo: cross-browser, platform-agnostic hosting of Artificial Life simulations

Simon Hickinbotham, Michael Weeks & James Austin

University of York, Heslington, York YO1 5DD, UK
email: sjh518@york.ac.uk

August, 2013