Integrating Taverna Player into Scratchpads

www.taverna.org.uk | scratchpads.eu

Robert Haines[+], Simon Rycroft[*],

Vince Smith[*], Carole Goble[+]

[+]*University of Manchester, UK;* [*]*Natural History Museum, UK*

robert.haines@manchester.ac.uk

Scratchpads and Taverna Player

- What are Scratchpads?
- Taverna and Taverna Player
- Why integrate the two?
- Lightweight integration (embed)
- Tight integration (Web Service)

- Scratchpads developed in ViBRANT
- Taverna Player developed in BioVeL

Scratchpads
Virtual Research Environments

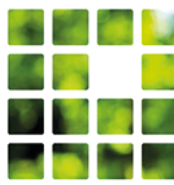*Making taxonomy digital, open & linked*

What are Scratchpads?

- Virtual Research Environments

- Hosted websites for biodiversity data

- Virtual research & publication platform

- Curated data and analysis

- Completely open access & open source

- Modular & flexible

What are Scratchpads?

facilitate

development of online research communities
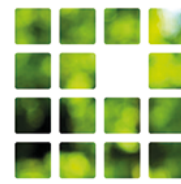
through

standardized environment of entering and curating data

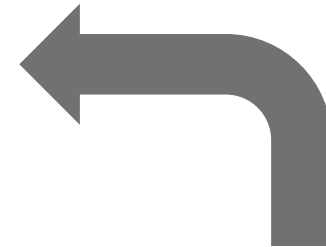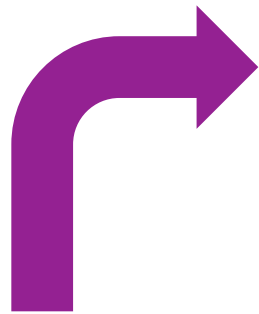**that allow**

sharing and interlinking

and

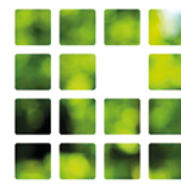dissemination of research products

The Scratchpads concept

A Scratchpad is a website that holds data for you and your community

Your data

External data & services

The Scratchpads concept

Scratchpads details

- Drupal CMS (7.26) with both custom and Drupal.org contributed modules.

- Over 500 Scratchpads sites managed by Aegir
  - www.aegirproject.org

- Hosted on two application servers
  - Load balancing and caching performed by Varnish

- 2 MySQL database servers
  - master-master configuration
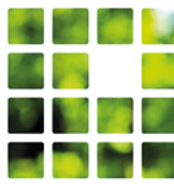
- 1 Apache Solr search server

Taverna

- Scientific Workflow Management System
- Workbench
  - Desktop application
- Command-line tool
  - Batch
- Server
  - Multi-user
  - Secure separation of workflow runs
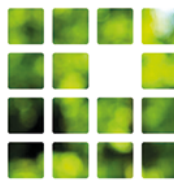 - REST and SOAP interfaces

Taverna Player

- A Ruby on Rails plugin library
  - Hooks into host application's
    - Workflow model
    - Authentication and authorization system
  - Provides a REST interface
- Talks to Taverna Server's REST interface
  - Uploads the workflow, sets inputs
  - Presents workflow interactions to the user
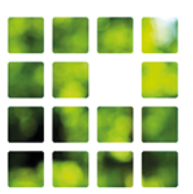 - Retrieves results, logs and provenance data
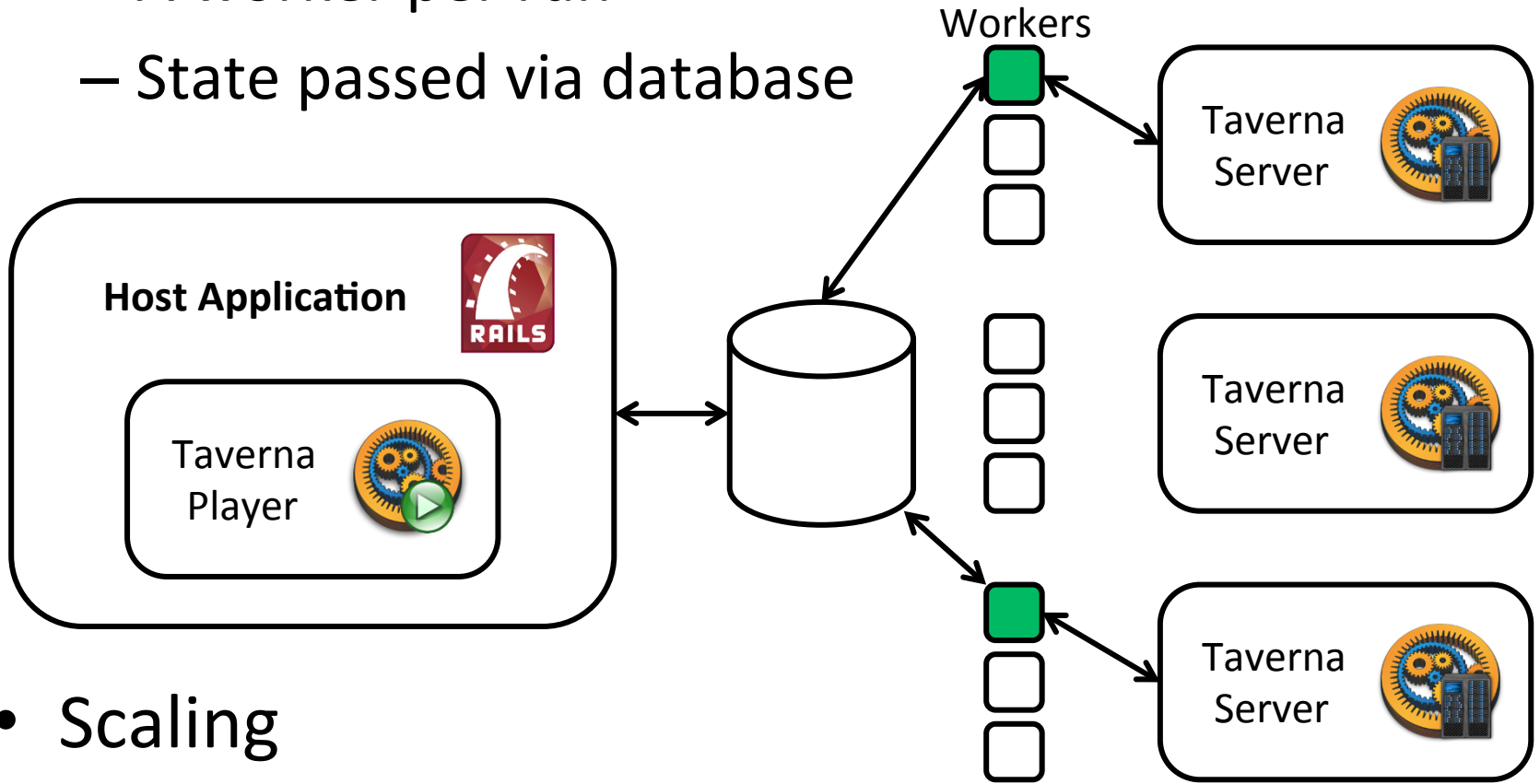
Taverna Player

- Surfaces a workflow run in three ways:
  - As a Web interface in the browser
    - In the host application
  - As an embeddable widget
    - In any Web page (c.f. YouTube videos)
  - As a REST-based Web Service
- All look-and-feel and styling is derived from the host application
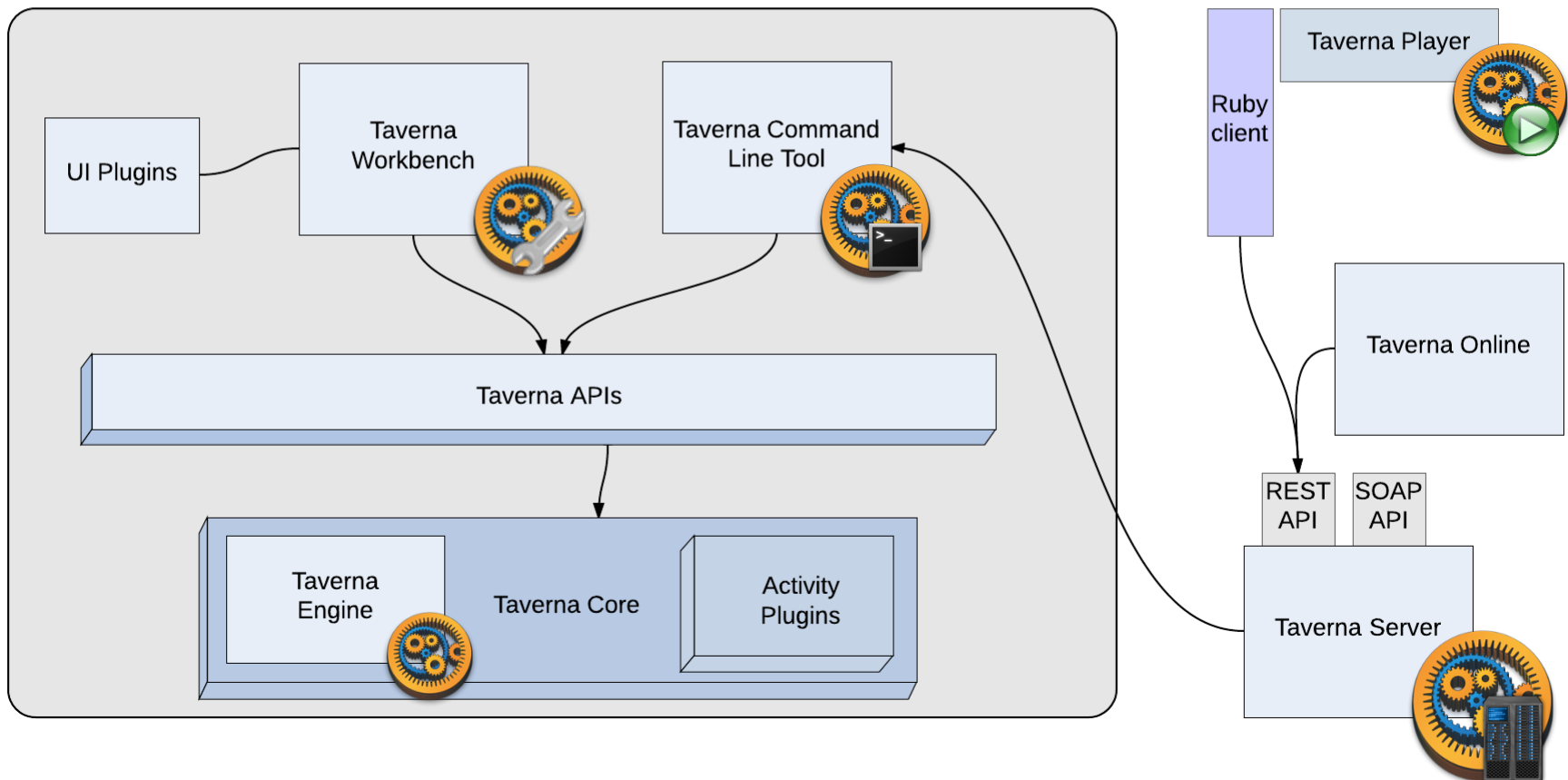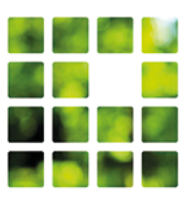 - Rails's hierarchical layouts and views

Taverna Player

- Total workflow run isolation
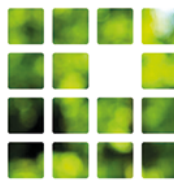 - A worker per run
  - State passed via database

- Scaling

Taverna all together

Taverna Player details

- Ruby on Rails
  - Version 3.2 (released) and 4.x (testing)
  - Plugin

- Delayed Job for workflow run isolation
  - Manages workflow run queues
  - Start workers to match Taverna Server capacity
  - Loss of a workflow run will not affect any others

Player in the BioVeL Portal
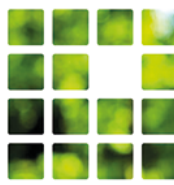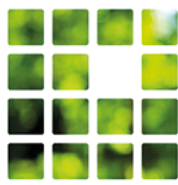
Why integrate?

- Join two communities

- BioVeL Portal
  - Good for the "day job", collaboration with others in, or close to, the project.

- Scratchpads
  - Dissemination; wide reach but focussed area
  - Move science into the public domain
  - Lots of data compatible with BioVeL pipelines

Workflows in Scratchpads I

- Lightweight embedding of a workflow

- Scratchpads updated to expose public data
  - As CSV
  - For each individual taxa
- Workflow is run as "guest" user
  - Embedding only available for "public" workflows
- Results stay in Taverna Player
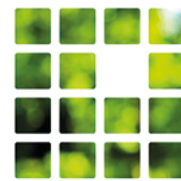
Workflows in Scratchpads I

- Embed like a YouTube video

- Embedded workflow is passed the URI of data

```
<iframe src="http://portal.org/runs/new?
                embedded=true&
                workflow_id=1&
                input_uri=http://scratchpad.org/taxa/1234/data"
>
</iframe>
```

- This level of integration is lightweight
  - Science showcases
  - One off analyses

https://github.com/myGrid/taverna-player/wiki/Embedding

Workflows in Scratchpads I

Workflows in Scratchpads II

- Tighter integration of analysis pipelines
  - Scratchpads directly controls Taverna Player

- Scratchpads has 'offsite computation' modules
  - Large scale batch operations, etc
  - Workflow runs added to this group

- Scratchpads uses the Taverna Player REST API
  - Within the host application/portal

Workflows in Scratchpads II
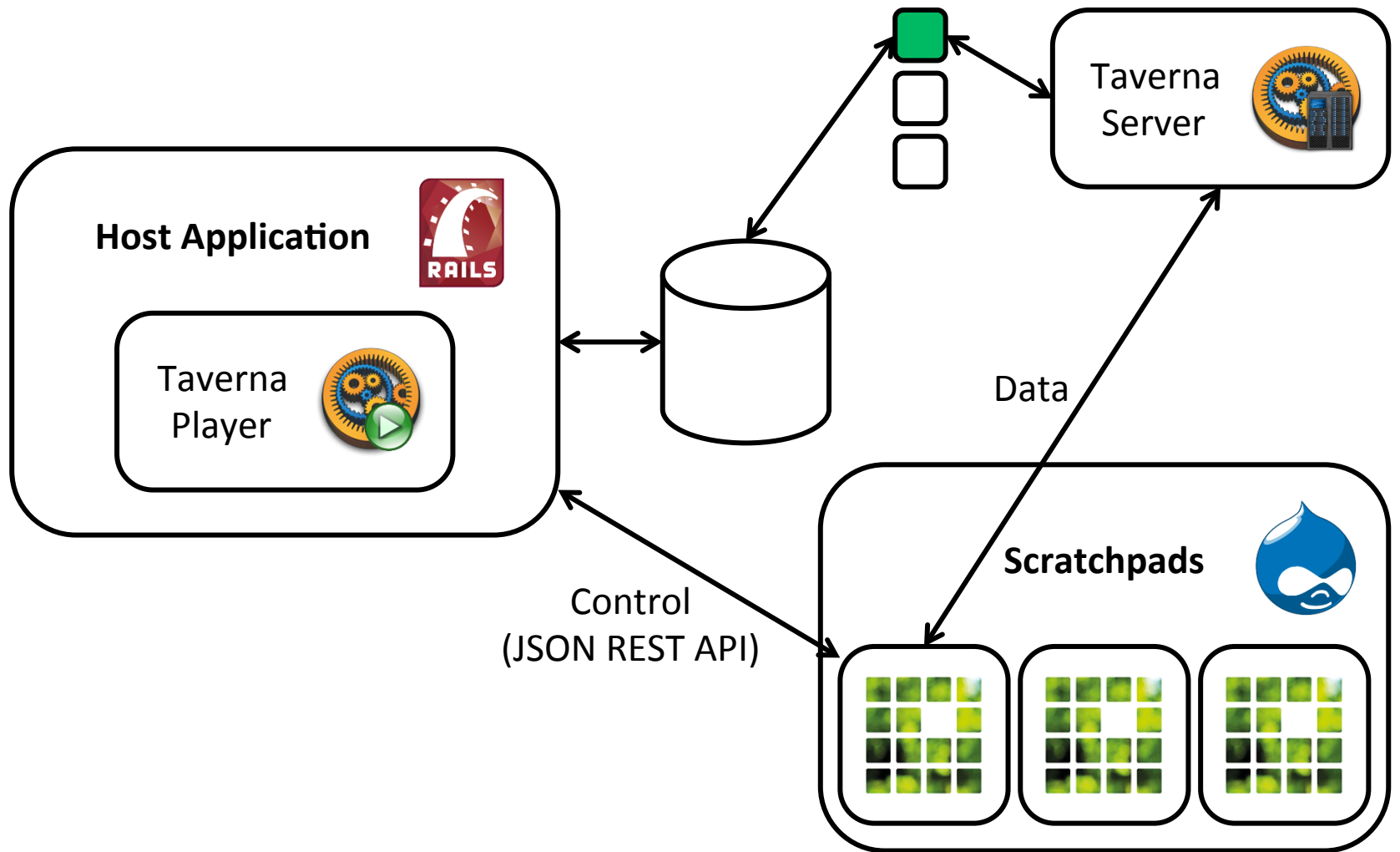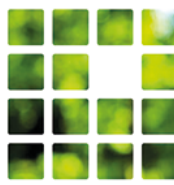
Workflows in Scratchpads II

- Scratchpads can
  - Authenticate to Taverna Player
  - Get a list of workflows and show these to the user
  - Set up the workflow run, inputs, etc
  - Present interactions to the user
  - Retrieve results for further analysis
- This level of integration is more suited to
  - Long-running workflows
  - Larger, repeated studies

Workflows in Scratchpads II

Workflows in Scratchpads II

Simple Interaction

| View | Delete |

Job name	Simple Interaction
Job type	Multi-interaction
Created	Thu, 2014-07-10 15:31
Changed	Thu, 2014-07-10 15:31

Id	1103
Name	Multi-interaction (Multi-interaction (Thu, 2014-07-10 15:31))
Workflow id	52
State	running
Start time	2014-07-10T15:31:55Z
Finish time	
Status message	Running
Inputs	

Workflows in Scratchpads II

Workflows in Scratchpads II

Simple Interaction

| View | Delete |

Job name	Simple Interaction
Job type	Multi-interaction
Created	Thu, 2014-07-10 15:31
Changed	Thu, 2014-07-10 15:34
Output file	taverna-1103.zip

Id	1103
Name	Multi-interaction (Multi-interaction (Thu, 2014-07-10 15:31))
Workflow id	52
State	finished
Start time	2014-07-10T15:31:55Z
Finish time	2014-07-10T15:34:32Z
Status message	Finished
Inputs	
Log	/runs/1103/download/log

Integration comparison

Lightweight embedding

- Run a specified workflow
  - Chosen by the Scratchpads owner
- Results are not stored in the Scratchpads itself
- Workflow run retains host app look and feel

Tight integration

- Run any workflow
  - That the Scratchpads is authorized to see
- Results are available for further analysis
- Workflow appears as part of the Scratchpads

Common

- Workflows are run within Taverna Player in the host app
- Interactions are presented to the user
- Results can be downloaded
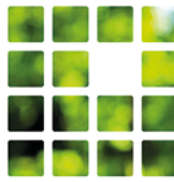
Tool Statistics

Taverna

- Workbench
  - Over 90,000 downloads
    - 1,000 in month after release
  - Over 380 sites use/used it
- Server
  - ?
- Player
  - Over 2000 downloads

Scratchpads

- Over 500 Scratchpads
- Over 900,000 pages
- Over 6,000 users

Thank you!

- Taverna Player
  - Robert Haines: robert.haines@manchester.ac.uk
  - Web: www.taverna.org.uk
  - Code: github.com/myGrid/taverna-player
  - Licence: BSD
- Scratchpads
  - Simon Rycroft: s.rycroft@nhm.ac.uk
  - Web: scratchpads.eu
  - Code: git.scratchpads.eu/git/scratchpads-2.0.git
  - Licence: GPL2