Deliverable: D7.2 – Mark-up modules

Partners: OU, KIT

Compiled by: David Morse (OU), Dauvit King (OU), Guido Sautter (KIT)

November 2012

WP7: Biodiversity literature access and data mining

Report on delivering D7.2 – Mark-up modules.

Due

Friday 30 November 2012 and completed ahead of schedule.

Milestones contributing towards deliverable

Four milestones contributed towards delivering D7.2, M7.11 and M7.13 from year one, and M7.16 and M7.17 from year two.

M7.11 Review of options for interactive mark-up tools within the Scratchpad infrastructure

Completed 17 June 2011.

See http://vbrant.eu/content/m711-review-options-interactive-mark-tools-within-scratchpad-infrastructure.

Delivered as a report[1] that concluded with the recommendation:

> "For maximum flexibility we prefer to implement web services. This will allow us to exploit the facilities of GoldenGate while providing a front end more in line with users' current expectations. The service will be accessible outside of Scratchpads too, so helping to establish a sustainable service through the potential for a larger user base."

This meant that to meet D7.2 we would deliver the existing mark-up services of GoldenGATE[2] rather than develop new services, but would have to rework the modules from then current monolithic workflow of the desktop implementation of GoldenGATE to make them available as web services. This defined the work to be done for M7.16.

M7.13 Review of options to use typographical information and other contextual clues

Completed 31 July 2011.

See http://vbrant.eu/content/m713-review-options-use-typographical-information-and-other-contextual-clues.

Delivered as two reports, a short overview report[3] of the issues and an in-depth review[4] of the modules within GoldenGATE that can use typographical information and other contextual cues.

This helped clarify the work required in year two to deliver relevant web-based mark-up services in M7.16 and to assess them in M7.17.

---

1   http://www.wiki.scratchpads.eu/w/M711report
2   http://www.plazi.org/?q=GoldenGATE
3   http://vbrant.eu/sites/vbrant.eu/files/M7-13_report.pdf
4   http://vbrant.eu/sites/vbrant.eu/files/Report%20-%20GoldenGATE%20Modules.doc

M7.16 Mark-up modules delivering outline mark-up

Completed 28 September 2012.

See http://vbrant.eu/content/m716-mark-modules-delivering-outline-mark.

The original delivery date for his milestone was 31 May 2012; however, owing to a change of priorities within WP7, agreed at Management Committee 6[5], the milestone was first deferred to 29 June 2012 and then again to its final autumn date to which it was successfully delivered.

This was the key practical milestone for meeting deliverable D7.2. In this milestone we delivered the individual GoldenGATE modules as web-based services. These services were developed locally and installed on the developers' machines.

In co-operation with Plazi a publicly available test interface was made available[6], though it is not widely advertised because it is intended purely for development and testing within the wider ViBRANT project. However, even this publicly available interface did not meet the requirement to make the services available within the Scratchpads infrastructure. To meet this requirement we worked in collaboration with WP5 to integrate the GoldenGATE web services into OBOE (Oxford Batch Operation Engine[7]), a component of the infrastructure specifically intended for the management of long-running tasks, such as the mark-up processes.

M7.17 Review of pilot mark-up processes within the Scratchpad infrastructure

Completed 23 November 2012.

See http://vbrant.eu/content/m717-review-pilot-mark-processes-within-scratchpad-infrastructure.

This milestone too was rescheduled from its original date of 31 July 2012 in line with the changes to M7.16 on which it depended. M7.16 was delivered 28 September 2012, but to allow time for a review of the mark-up processes around the other commitments, including attending TDWG conference[8] in Beijing, the milestone was finally moved to 23 November 2012 at the end of three one-day workshops[9] that week in which mark-up processes and their integration with the Scratchpads was among the topics discussed. The milestone was successfully completed to this revised schedule following a satisfactory review in the ViBRANT workshop held on 21 November 2012[10].

The review covered two possible workflows.

The first workflow was the full GoldenGATE workflow, as required by the desktop version of GoldenGATE, but using the online web services. This was successfully completed to prove that all GoldenGATE modules are available as web services should they be required. As noted in M7.11 though, this workflow is not practical for every day use being too time consuming.

The second workflow was involved selected services only, such as Date Tagging and Geolocation referencing, that could be added to a user's existing workflow or integrated into new, focused workflows. This review used GoldenGATE web services delivered through OBOE to confirm their

---

5   http://vbrant.eu/content/man-comm-6
6   http://vbrant.ipd.uka.de:8080/GgOE/markup
7   https://oboe.oerc.ox.ac.uk/
8   http://www.tdwg.org/conference2012/
9   http://vbrant.eu/content/three-days-workshops
10   http://vbrant.eu/content/notes-oxford-workshop-21-november-2012

use within the Scratchpad infrastructure. This was successfully completed to prove that the de-coupling across GoldenGATE modules had been achieved, as well as their integration within the Scratchpad infrastructure.

Summary

In summary:

- in M7.11 we identified GoldenGATE as a suitable tool to meet this deliverable

- in M7.13 we confirmed the available options for mark-up within this deliverable

- in M7.16 we delivered each GoldenGATE module as a discrete web service

- in M7.17 we integrated GoldenGATE web services through OBOE into the Scratchpad infrastructure.

For more detail on each of theses steps on the road towards delivering D7.2, rather than repeat the contents of each milestone's report, please see refer to the reports themselves. The reports include a sample script to invoke GoldenGATE web services and screenshots of their integration through OBOE with the Scratchpad infrastructure.

Progress in year three

Though this formal deliverable is complete, we recognise that to gain maximum value from it requires additional work. We now want to exploit the services in more workflows. To do this we will liaise with WP8 and their development of ViBRANT workflows. We will also look to provide more example use cases, supported by demonstration scripts, to show how the web services can be used in other workflows. Finally, there remains the packaging of the tools and their being made publicly available through the Scratchpads hosted GIT repositories.