# TEXTUS JISC Final Report

<table>
<thead>
<tr>
<th>Project Information</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Identifier</td>
<td>To be completed by JISC</td>
<td></td>
</tr>
<tr>
<td>Project Title</td>
<td>TEXTUS</td>
<td></td>
</tr>
<tr>
<td>Project Hashtag</td>
<td>@TEXTUSProject</td>
<td></td>
</tr>
<tr>
<td>Start Date</td>
<td>01/02/2012</td>
<td></td>
</tr>
<tr>
<td>End Date</td>
<td>31/09/2012</td>
<td></td>
</tr>
<tr>
<td>Lead Institution</td>
<td>Open Knowledge Foundation</td>
<td></td>
</tr>
<tr>
<td>Project Director</td>
<td>Sam Leon</td>
<td></td>
</tr>
<tr>
<td>Project Manager</td>
<td>Sam Leon</td>
<td></td>
</tr>
<tr>
<td>Contact email</td>
<td><a href="mailto:sam.leon@okfn.org">sam.leon@okfn.org</a></td>
<td></td>
</tr>
<tr>
<td>Partner Institutions</td>
<td>Goldsmiths University London</td>
<td></td>
</tr>
</tbody>
</table>
| Project Web URL           | [http://textusproject.org](http://textusproject.org)  
                             [http://beta.openphilosophy.org](http://beta.openphilosophy.org) |
| Programme Name            | TEXTUS |
| Programme Manager         | Christopher Brown  
                             Christopher Brown  
                             <c.brown@jisc.ac.uk> |

<table>
<thead>
<tr>
<th>Document Information</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Jillian Mathews &amp; Sam Leon</td>
<td></td>
</tr>
<tr>
<td>Project Role(s)</td>
<td>Project Managers</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>23/10/2012</td>
<td></td>
</tr>
<tr>
<td>Filename</td>
<td></td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Access</td>
<td>This report is for general dissemination</td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgements

TEXTUS was funded by JISC and is part of the Research Tools program of projects. The two project partners were Goldsmiths, University of London and the Open Knowledge Foundation, Cambridge.

We would like to thank the project team Tom Oinn, Anca Pusca and Aya Hino as well as Christopher Brown at JISC for his support. We would also like to thank the Open Philosophy Advisory Board for their input and those that attended the user requirements workshops in London and Cambridge for their invaluable feedback and comments.

Project Summary

The combination of freely available digital copies of public domain works, open bibliographic data and open source tools has the potential to revolutionise research in the humanities. However, there are currently numerous obstacles which mean that they are often under utilised by scholars and students in teaching and research.
From classic literary and cultural works, to letters, drafts, notes, and other historical documents, there is a huge amount of freely available public domain material that is highly relevant to scholars and students engaged in research in the humanities. But these works can be difficult to find, difficult to work with, and works by a given author may be scattered in a variety of locations. Search results may be confusing or unclear. Automated Optical Character Recognition of texts may be inaccurate or incomplete.

There are a growing number of open source tools for transcribing, translating and annotating texts. However many of these are one-off projects and it may not be clear how to deploy the tools in relation to a given text or collection of texts. Academic awareness of these tools and their potential benefits to their research is not great.

The overall goal of TEXTUS is to provide an open-source platform through which scholars and students are able to re-use the vast and expanding amount of digitised cultural heritage material now available through portals such as the Internet Archive, Wikisource and Project Gutenberg. The aim is to enable scholarly communities to easily establish their own instances of TEXTUS catering to their specific community and to enable developers to easily extend the platform itself and develop other apps and services for it.

The figure below illustrates this idea:

TEXTUS aims to make it easy for scholars and research students to work collaboratively around public domain texts. The JISC project has created an exemplary instance of the TEXTUS platform on beta.openphilosophy.org with students and researchers at Goldsmiths University in order to kick-start the uptake of the TEXTUS platform within the humanities academic community.
The core functionality that has been developed allows users to:

**Search**

Nicolò Machiavelli, *The Prince*, 1532

Karl Marx, *On the Jewish Question*, 1842

Thomas Hobbes, *De Cive*, 1642

John Stuart Mill, *On Liberty*, 1859

Sigmund Freud, *Reflections on War and Death*, 1918

Immanuel Kant, *Perpetual Peace*, 1795


**Read**

**Annotate and share**
Reference and cite using stable URLs

Text

For, in that sense, it would be a mere nuisance, or a suspension of hostilities, and not a Peace. A Peace properly signifies the end of all hostilities, and to qualify it by the addition of the epithet 'permanent', or 'internal' is pernicious and abusive. All existing causes for a future war—although they were perhaps unknown to the contracting parties at the time—to be regarded as entirely removed, or assented to by the hands of Peace, even if they might be pointed out by the interest or an acute interpretation from the terms of documents in the public Almanacs. There may be a mental reservation of all pretensions or claims with the view of according them at a future time; of which, however, neither party makes any mention for the present between them. They have not consented to contract for war, while these reserves the law will to the exigency of the time, for the present; but this is an entire suspension of hostilities, or declaration of Peace. In this case, the combatants may be said to be at war; but the Peace is merely the suspension of hostilities, with this condition, that it is to be broken by a subsequent agreement for a future state of hostilities. In no case can it be said that they have made a Peace, unless it be with the reservation of all their pretensions, and the right of making a future War, as is absolutely necessary for the protection of the public peace. While this is the case, the combatants may be said to be at war, but the Peace is merely an agreement to suspend hostilities, and the right of making a future War, as is absolutely necessary for the protection of the public peace.

A State is not to be regarded as a property or a patrimony, but as a moral entity, on which no one but itself has the right to take an interest. To give the sword of a State to a new owner, and to establish it in a group of subjects, is to destroy the freedom on which a State is founded. The sword is to be used for the protection of the State, and for the enforcement of its laws; not to be used for the enrichment of its rulers. The sword is to be used for the protection of the State, and for the enforcement of its laws; not to be used for the enrichment of its rulers. The sword is to be used for the protection of the State, and for the enforcement of its laws; not to be used for the enrichment of its rulers. The sword is to be used for the protection of the State, and for the enforcement of its laws; not to be used for the enrichment of its rulers.

References

Immanuel Kant, "Perpetual Peace", 1795
## Main Body of Report

### Project Outputs and Outcomes

<table>
<thead>
<tr>
<th>Output / Outcome Type</th>
<th>Brief Description and URLs (where applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXTUS platform and beta.openphilosophy.org</td>
<td>First deployment of TEXTUS with reading, annotation and citation functionality populated with public domain philosophy texts and annotations.</td>
</tr>
<tr>
<td></td>
<td>Open Philosophy instance: <a href="http://beta.openphilosophy.org">http://beta.openphilosophy.org</a></td>
</tr>
<tr>
<td></td>
<td>Code repository: <a href="https://github.com/okfn/textus">https://github.com/okfn/textus</a></td>
</tr>
<tr>
<td></td>
<td>Documentation: <a href="https://github.com/okfn/textus/tree/master/docs">https://github.com/okfn/textus/tree/master/docs</a></td>
</tr>
<tr>
<td>TEXTUS user requirements workshop 1</td>
<td>A user requirements workshop was held in month 1 in London with researchers and students from the humanities.</td>
</tr>
<tr>
<td>TEXTUS user requirements workshop 2</td>
<td>The second user requirements workshop was held in month 2 of the project to get feedback on wire-frames and mock-ups from students and researchers at Goldsmiths University.</td>
</tr>
<tr>
<td>User testing workshop, Goldsmiths University of London</td>
<td>In month 4 of the project the alpha version of the platform was tested by the group at Goldsmiths University.</td>
</tr>
<tr>
<td></td>
<td>Blog post: <a href="http://textusproject.org/2012/05/22/testing-textus/">http://textusproject.org/2012/05/22/testing-textus/</a></td>
</tr>
<tr>
<td>User testing workshop at Cambridge University, Philosophy Department</td>
<td>User testing workshop at Cambridge University, Philosophy Department with feedback on alpha version of platform.</td>
</tr>
<tr>
<td></td>
<td>Blog post: <a href="http://textusproject.org/2012/06/27/user-testing-workshop-at-cambridge-university/">http://textusproject.org/2012/06/27/user-testing-workshop-at-cambridge-university/</a></td>
</tr>
<tr>
<td>User testing workshop at Kings College London, Digital Humanities Department</td>
<td>User testing workshop at Kings College London, Digital Humanities Department.</td>
</tr>
<tr>
<td></td>
<td>Blog post: <a href="http://textusproject.org/2012/06/26/user-testing-workshop-at-kings-college-london/">http://textusproject.org/2012/06/26/user-testing-workshop-at-kings-college-london/</a></td>
</tr>
<tr>
<td>Gathering feedback from users via mailing lists</td>
<td>Two main mailing lists were used and grown for discussion around TEXTUS.</td>
</tr>
<tr>
<td></td>
<td>(1) Open Humanities Discussion List (222 members)</td>
</tr>
<tr>
<td></td>
<td>List Archives: <a href="http://lists.okfn.org/pipermail/open-humanities/">http://lists.okfn.org/pipermail/open-humanities/</a></td>
</tr>
<tr>
<td></td>
<td>(2) Humanities Developer List (42 members)</td>
</tr>
<tr>
<td></td>
<td>List Archives: <a href="http://lists.okfn.org/pipermail/humanities-dev/">http://lists.okfn.org/pipermail/humanities-dev/</a></td>
</tr>
<tr>
<td>Project blog with regular updates</td>
<td>The project blog was updated on a monthly basis for the duration of the project.</td>
</tr>
<tr>
<td></td>
<td>Blog URL:</td>
</tr>
</tbody>
</table>
Discussion of TEXTUS amongst open-source developer community

Github and the Humanities Developer list hosted by the Open Knowledge Foundation (see above) were used for engaging open-source developers and gathering input and feedback from them.

Engagement with users via social media channel

The main Twitter handle used to for TEXTUS was @TEXTUSproject and the #JISCrestools hash-tag.

TEXTUS hackdays

Held at Bibliohack in collaboration with the JISC Funded Open Biblio 2 project and at OKFestival in Helsinki, September 2012.

Philosophy reading groups at Goldsmiths University

A series of 8 reading groups taking place between October and December 2012 at Goldsmiths University at which scholars presents texts uploaded into TEXTUS and use the platform as an aid for annotation and discussion. A write-up of the first three meetings and the lessons learned so far can be found here.

Final report

NA

How did you go about achieving your outputs / outcomes?

Overview of work undertaken and how it was completed

The work undertaken as part of TEXTUS was divided over four main areas of work or Work Packages: development (Work Package 1), use testing and evaluation (Work Package 2), community building and dissemination (Work Package 3) and project management (Work Package 4).

The main body of work was done over 6 months, but there was also a two-month no-cost extension for Work Package 2 to allow for further dissemination work and the delivery of the reading groups at Goldsmiths University of London.

Work Package 1: Software development

Aims and objectives

- Develop an open and extendible online platform for collaborating around texts
- Integrate existing open-source tools such as BibServer

Methodology

Software development on the TEXTUS project followed user-centric principles and agile methodologies.

User requirements were gathered in Work Package 2 in month 1 of the project and an initial evaluation of the existing open-source tools and components was made to see if existing tools could meets the needs of users. As outlined below, it was decided that certain components, such as the Annotator, which we initially set out to integrate were deemed at this stage not to meet the needs of the particular community of users we were addressing.

Iterations consisted of two-week cycles and the platform was tested by students and researchers at Goldsmiths University at the very earliest stages informing later iterations.
The first two months were spent building the annotation rendering engine, the second two months building the bibliographic management component and then the final two months implementing citation functionality and iterating on the user interface and its appearance.

The project benefited from the input and feedback of the open-source development community engaged through Work Package 3 via the Open Knowledge Foundation’s Humanities Developer List.

**Deliverables (from table above)**

- TEXTUS platform (https://github.com/okfn/textus) and Open Philosophy instance (http://beta.openphilosophy.org)

**Work Package 2: User testing and evaluation**

**Aims and objectives**

- Gathering user requirements
- Testing aspects of the system
- Using the system and populating it with content

**Methodology**

The user testing and evaluation of the TEXTUS platform focused on feedback from groups of researchers and students from the humanities. Requirements gathering took place over two half-day sessions with a group from Goldsmiths University.

User testing sessions were held in the second phase of the project at the Cambridge University Philosophy Department, Kings College London Digital Humanities Department and Goldsmiths University of London.

Feedback was collected verbally and all participants were given the opportunity to fill in evaluation forms that were collated by the Project Manager and used to inform the development work.

Students were incentivized to attend the user testing workshops by being offered a token gift of an £20 Amazon voucher.

**Deliverables (from table above)**

- User requirements workshop 1
- User requirements workshop 2
- User testing workshop at Goldsmiths University
- User testing workshop at Cambridge University Philosophy Department
- User testing workshop at Kings College London
- Open Philosophy reading groups at Goldsmiths University
Work Package 3: Community building and dissemination

**Aims and Objectives**

- Promote the TEXTUS platform to academics and students who could use it in their teaching and research
- Engage the open-source developer community around the TEXTUS platform

**Methodology**

Scholarly engagement around the platform was achieved through engaging the high-profile [Open Philosophy Advisory Board](#) and inviting them to attend TEXTUS events organised at their institution, as well as keeping them informed about the development of the project via email. User requirements, user testing workshops and reading groups also proved important mechanisms for raising awareness of the platform within academic institutions.

The developer community was engaged through the [Humanities Developer List](#) hosted by the Open Knowledge Foundation and through [Github](#) via push and pull requests and issues. We also ran two TEXTUS hackdays directly targeted at developers who had an interest in TEXTUS and annotation platforms for the humanities.

The wider-community, including JISC, was engaged through the dedicated blog, the [Humanities Discussion List](#) and Twitter using the dedicated [@TEXTUSproject](#) handle and the #JISCresttools hashtag.

**Deliverables (from table above)**

- Project blog
- Use and expansion of Open Knowledge Foundation humanities mailing lists
  - [Humanities Discussion List](#) (222 members)
  - [Humanities Developer List](#) (42 members)
- TEXTUS Hack at Bibliohack, London
- TEXTUS Hack at Open Science and Culture Hackday at OKFestival, Helsinki
- Philosophy reading groups at Goldsmiths University

Workpackage 4: Project management

**Aims and objectives**

- Financial management of project
- Ensure deliverables are met
- Engage Advisory Board
- Report project progress to JISC
Methodology

The Open Knowledge Foundation was tasked with the project management. The Foundation oversaw management of the project budget, arranged project meeting and liaised with JISC and the Open Philosophy Advisory Board.

An initial meeting of the whole project team was convened in the first month of the project. Weekly meetings took place between the Project Manager and the developer to track progress towards deliverables and keep track of the iteration plan. The Open Philosophy Advisory Board was updated via email and engaged at events held at the members’ institutions.

Deliverables (from table above)

- Final project report

Deviations from original description of work

While the vast majority of work unfolded according to plan there were a few areas in which we deviated from the plan set out in the original proposal. All but one deviations were down to the core development taking longer than initially anticipated:

<table>
<thead>
<tr>
<th>Proposed work</th>
<th>Deviation</th>
<th>Reason for deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Release of full version of TEXTUS on Open Philosophy in month 5 of project</td>
<td>Delayed until September 2012</td>
<td>TEXTUS was originally proposed as a pure integration project with its annotation functionality being built on top of the Annotator. The user requirements sessions revealed that the Annotator was not fit to deliver the functionality requested by students and scholars and an annotation engine needed to be developed from scratch taking considerable developer time. More detail on this can be found in the TEXTUS documentation and on the TEXTUS blog</td>
</tr>
<tr>
<td>Use of Walter Benjamin texts on Open Philosophy and in scholarly case study</td>
<td>Walter Benjamin’s texts not used, alternative public domain texts found</td>
<td>Lack of out-of-copyright English translations of any of Benjamin’s works</td>
</tr>
<tr>
<td>Reading group series to start in month 2</td>
<td>Delayed until October 2012</td>
<td>Due to delay in platform development (see above)</td>
</tr>
<tr>
<td>Development of transcription functionality</td>
<td>Not developed</td>
<td>After an initial survey of the existing open-source tools and services for the transcription of digitised material it was deemed that it would be better to build TEXTUS in such a way that it</td>
</tr>
</tbody>
</table>
could feed into existing transcription tools such as Wikisource Scriptorium rather than building a new feature from scratch.

<table>
<thead>
<tr>
<th>Development of reading list curation functionality</th>
<th>Partially developed</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXTUS can manage bibliographic references through the integration of BibServer. However, there was not sufficient developer time to implement the full range of reading list curation functionalities. This has now been tagged as an issue for the developer community around TEXTUS to address in future iterations.</td>
<td></td>
</tr>
</tbody>
</table>

Lessons Learned

Development activities

- Sources for digital public domain texts, which TEXTUS feeds off such as Wikisource and Project Gutenberg, do not adhere to clear standards. A lack of standards in this domain for digital transcriptions makes the development of import scripts very difficult and is a big obstacle to the re-use of digital public domain texts in platforms such as TEXTUS;
- Negotiations around coding language caused hold-ups. The platform was initially to be coded in Python but was later coded in Node JS. Starting in Node JS given the developer’s skill set would have been more appropriate;
- An event should have been planned in the early part of the project bringing together developers from TEXTUS, Annotator and the OpenBiblio projects allowing for integration and collaboration at an earlier stage.

User requirements, testing and evaluation

- There is a clear demand for a tool that makes the reading, referencing and annotating of digital public domain text easier within universities. This can be seen for instance in that the initial list of sign-ups for remote user testing of Open Philosophy was well above 100;
- The group from which user requirements were gathered was strongly skewed towards students. To have had more teachers involved at this stage would have allowed for more refined work to be done on the application of TEXTUS in teaching contexts and might have increased uptake of the platform towards the end of the project.

Dissemination activities

- Syncing up the project with the academic year, the final phase of development was undertaken in a period in which many universities were not teaching. This made disseminating information about Open Philosophy more difficult during the project, leading to the request for a no cost extension.
Project management

- Certain aspects of the project management were hindered by project staff from the Open Knowledge Foundation and Goldsmiths University having to connect remotely with little opportunity for face-to-face meetings. Face-to-face meetings would likely have helped forge a closer project team, more quickly able to respond to the obstacles faced during the lifetime of the project.

Other

- There is still a distinct lack of clarity with respect to the copyright status of certain digitised public domain works and in many cases a lack of authoritative digital transcripts of digital texts. This was clearly demonstrated by the lack of any openly licensed digital copies of Walter Benjamin’s works which it was the projects initial intention to use in OpenPhilosophy.org

Immediate Impact

Three academic institutions across the country (Goldsmiths, Cambridge and Kings College London) have now used the TEXTUS platform. In Goldsmiths plans are afoot to use the platform to teach certain humanities courses allowing students from other disciplines and institutions to engage in the material taught on that course.

The wider community has benefited from TEXTUS in three main ways:

- There is a set of openly licensed digitised texts on beta.openphilosophy.org that the wider community can augment and annotate;
- There is an active community focused around the Open Knowledge Foundation’s Open Humanities mailing list devoted to searching for openly licensed digital works for inclusion in future instances of TEXTUS. The crowd-sourced spreadsheet of philosophical works for upload into a future instance of TEXTUS is evidence of this;
- There is an open platform that can be implemented for communities with specific needs and further customised.

TEXTUS has changed the attitudes of stakeholders in three ways:

- Researchers: It has demonstrated to certain parts of the research community that digital editions of texts can be scholarly resources, provided tools like TEXTUS exist to create reliable citations;
- Teachers: It has demonstrated the value of using online platforms for running seminars and reading groups to a number academic institutions within the South East;
- Open content community: Major stakeholders within the open content community such as Wikimedia, the Internet Archive and Distributed Proofreaders are aware of TEXTUS and the value of such tools for catalysing active reuse of the openly licensed material that they host.
Future Impact

- We hope that through the ongoing dissemination work of the Open Knowledge Foundation we can continue to engage developers and those within the humanities to continue building TEXTUS functionality and establishing more instances of the platform catering for groups beyond the philosophy community. We will be tracing this impact via mailing list discussion analytics and Github push, pull and issue requests;
- In the interim period the volunteer developer community at the Open Knowledge Foundation is continuing to work on TEXTUS and implementing further instances of it on OpenLiterature.net with a planned migration of the OpenShakespeare.org content to the new instance of TEXTUS. The road map for this community work can be found here.
- The Open Knowledge Foundation developer community has started to develop tools such as the Guttenizer that will make it easier to import texts into TEXTUS that are available through Project Guttenberg addressing the need (outlined above) to store digitised public domain texts in a common format.

Conclusions

The JISC TEXTUS project has developed a robust and innovative platform for collaborating around digitised public domain texts and has promoted the platform to academic communities within the UK and further-a-field.

On beta.openphilosophy.org groups of students and researchers from Goldsmiths University of London have read and annotated an exemplary set of material that they have curated and uploaded into the platform.

There remains work to be done on further developing the existing features particularly with respect to the management of bibliographic references, but also in encouraging the further uptake and establishment of TEXTUS instances beyond beta.OpenPhilosophy.org that cater for different scholarly communities.

There is clearly a demand for collaborative online tools such as TEXTUS in teaching and research and more should be done to continue to engage scholars and promote the use of TEXTUS in seminars, lectures and classrooms.

Recommendations

- Goldsmiths University continue to run the series of TEXTUS reading groups currently taking place;
- Goldsmiths University continue to support use of the platform in teaching environments;
- The Open Knowledge Foundation continue to engage developers through hackdays and mailing lists for the development of further iterations of the platform;
- Resource permitting JISC should continue to support further feature development of TEXTUS and the establishment of further instances for humanities scholarly communities;
- The Open Knowledge Foundation continue to work closely with galleries, libraries, archives and museums through its OpenGLAM initiative to secure the release of more openly licensed digital content and metadata in machine readable formats that can be re-used via TEXTUS and other similar platforms for the use of cultural heritage data;
The Open Knowledge Foundation solicit further funding for the further development and promotion of TEXTUS.

Implications for the future

- Humanities scholars, Digital Humanists and the open-source developer community now have a robust, extendable, open-source platform for annotating, reading and citing digital public domain texts;
- Further suggested features and improvements are already banked in the TEXTUS Issues Github repository and are currently being addressed by the developer community at the Open Knowledge Foundation made possible by the open-source nature of the code itself.
- For the platform to have greater uptake, the system will need to be further improved. The Open Knowledge Foundation is currently seeking European funding for the further development of TEXTUS with a large pan-European Consortium;
- OpenPhilosophy.org will continue to be hosted and maintained by the Open Knowledge Foundation as well as the other instances of TEXTUS such as OpenLiterature.org currently under development in the working group on the humanities at the Open Knowledge Foundation.
- The long term contact for the project is Sam Leon <sam.leon@okfn.org> who will be responsible for pursuing further funding leads and engaging the scholarly and developer community around the existing platform.