

TEXT AND META DATA SHARING AGREEMENT

THIS TEXT AND METADATA SHARING AGREEMENT (this “Agreement”) is entered into by and between PROQUEST LLC, a Delaware Limited Liability Company (“PROQUEST”), with offices at 789 e. Eisenhower Parkway, Ann Arbor, MI 48106-1346 and Networked Interface for Nineteenth-century Electronic Scholarship, a not-for-profit endeavor from Texas A&M University (“ARC” representing “REKN/ 18THCONNECT”) as of the Contract Date specified below (the “Effective Date”).

WHEREAS, PROQUEST’S mission is to create indispensable research solutions that connect researchers with information and empower them to use, create and share content, including but not limited to its EEBO Database which includes page images of works originating in 14th to 17th century; and

WHEREAS, REKN/18THCONNECT seeks to enable REKN and 18thConnect users to explore 14th to 17th-century works currently available in digital form by providing scholars with the ability to conduct research within federated digital environments, as well as to produce personal annotated collections and exhibits using an integrated tool.

NOW, THEREFORE, in consideration of the premises and the mutual promises contained herein and other good and valuable consideration, the receipt and sufficiency of which are hereby acknowledged, the parties agree as follows:

1. DEFINITIONS

For purposes of this Agreement, the following definitions shall apply:

- 1.1 “Typed Plain Text” shall mean the text that has been manually re-keyed, reviewed by the Text Creation Partnership (TCP), and that the TCP has authorized REKN/18THCONNECT to use.
- 1.2 “OCR Plain Text” shall mean the text files derived from page images (1.6) by REKN/18THCONNECT using open-source Optical Character Recognition Programs (1.8), and any corrections to that text provided by REKN’s/18THCONNECT’s membership using the Crowd-sourced Editing Platform (Appendix 1).
- 1.3 “EEBO Collection” shall mean the online database published by PROQUEST called Early English Books Online (EEBO).
- 1.4 “Metadata” shall mean bibliographic information describing the objects contained in the page images reproduced in the EEBO Collection. Metadata regarding these objects has been provided to REKN/18THCONNECT by the British Library of the English Short Title Catalogue (ESTC).
- 1.5 “Project” shall mean REKN’s/18THCONNECT’s use of the Typed Plain Text as given to 18THCONNECT by the TCP, OCR Results and OCR Plain Text as created by

18THCONNECT and its members, and Metadata as given to REKN/18THCONNECT by the British Library for indexing and enabling scholars to search and locate content in connection with REKN's/18THCONNECT's integrated environment for aggregated, peer-reviewed research and online scholarship of 16th- through 19th-century texts .

- 1.6 "Page Images" shall mean the digitized page images of the texts in the EBBO Collection.
- 1.7 "OCR Plain Text" shall mean the text of a complete work or portion of a work in the EBBO Collection that has been OCR'd and corrected by users of REKN/18THCONNECT.
- 1.8 "OCR Results" shall refer to the OCR xml as required by PROQUEST and the TCP to allow full-text searching in EBBO and word coordinates generated by the Open-source Optical Character Recognition Programs – specifically Tesseract, Gamera, and OCRopus. The form and format to be returned to PROQUEST shall be mutually agreed upon.
- 1.9 "Open-source Optical Character Recognition Programs" shall refer to the programs used by the 18THCONNECT team to generate "OCR Results," currently but not limited to Tesseract, Gamera and OCRopus.
- 1.10 "Open Source Libraries" shall refer to the modifications made to open-source Optical Character Recognition Programs by the 18THCONNECT team as it attempts to derive better results from the Optical Character Recognition Process.

2 GRANT OF LICENSE

- 2.1 PROQUEST hereby grants REKN/18THCONNECT a limited, nonexclusive, royalty-free right to create and use the OCR Results and OCR Plain Text (as submitted by 18THCONNECT and as peer reviewed by REKN/18THCONNECT) solely in connection with the Project. REKN/18THCONNECT may store the OCR Results on its servers for the purpose of indexing and text correction using the TypeWright tool. and shall display such OCR Results to users only in short, query-dependent texts or data ("Snippets"). Snippets may include Metadata as well as a few lines of query-dependent Full Text (as taken from either the Typed Full Text or OCR Results) and are not likely to exceed three (3) lines at a reasonable font size. REKN/18THCONNECT will also display as part of the Project a link to a page on the PROQUEST website for the EBBO Collection, <http://eebo.chadwyck.com/home> where PROQUEST will make the content referenced by the Snippet available to users authorized under license to access the EBBO Collection. For avoidance of doubt, this Grant of License does not permit any person or party to access any product (for example, EBBO) unless such person or party is an authorized user as determined by appropriate license and purchase agreements.
- 2.2 REKN/18THCONNECT shall not under any circumstances redistribute the Typed Plain Text, OCR Results, the OCR Plain Text or Metadata to any third party, agent, or affiliate except as set forth in Article 2.1 herein and integrated into REKN's/18THCONNECT's search-and-browsing interface that is part of the Project, unless such redistribution is explicitly in writing authorized by PROQUEST REKN/18THCONNECT's license to use

and/or incorporate the Metadata is covered under separate license from the British Library as part of ESTC. REKN/18THCONNECT's license to use and/or incorporate the Typed Plain Text is covered under separate license from the Text Creation Partnership (TCP)

- 2.3 REKN/18THCONNECT will use the Metadata as supplied by the ESTC, the Typed Plain Text as supplied by the TCP and 18THCONNECT will extract OCR Results and OCR Plain Text from page images of the EEBO Collection or parts thereof as delivered to the NCSA (National Center for Supercomputer Applications) by PROQUEST for 18THCONNECT following the execution of this Agreement.
- 2.4 REKN/18THCONNECT will return the OCR Results and OCR Plain Text, to PROQUEST, per an agreed-upon schedule but not less than annually, for use within ProQuest's products and services, including the EBBO Collection, at no cost/obligation to PROQUEST, should PROQUEST choose to use the OCR Results or the OCR Plain Text. In addition, REKN/18THCONNECT shall return any custom dictionaries and/or taxonomies created as part of the Project for use within ProQuest's products and services. The right to use the OCR Results, the OCR Plain Text and any custom dictionaries and taxonomies as granted above is a perpetual, unrestricted right in any format, whether in existence now or developed in the future, and the right shall explicitly survive termination of this Agreement.
- 2.5 REKN/18THCONNECT and PROQUEST expressly agree to allow text correction by scholars of OCR Results in order to create OCR Plain Text under the provisions of the attached Appendix 1.
- 2.6 PROQUEST may request removal of the Typed Plain Text, the EEBO Collection page images, the OCR Results, the Plain Text or Metadata or any part thereof from the Project at any time and for any reason whatsoever, by sending a request to the REKN/18THCONNECT Notice Contact including the specific URLs to be removed. On receipt of such a request, REKN/18THCONNECT will make reasonable efforts to remove and cease the use of the specific EEBO Collection page images, the OCR Results, the Plain Text or Metadata – including the display of the OCR Plain Text pertaining to the URLs within sixty (60) days. Notwithstanding the foregoing, PROQUEST may make an emergency request to REKN/18THCONNECT for the removal of specific EEBO Collection page images, the OCR Results, the Metadata and/or OCR Plain Text or any part thereof from the Project. In the event of an emergency request, REKN/18THCONNECT shall make reasonable efforts to remove and discontinue the display of the any EEBO Collection page images, the OCR Results, the Metadata and/or any OCR Plain Text pertaining to the URL's specified within five (5) business days.
- 2.7 During the term of this Agreement, representatives of PROQUEST and REKN/18THCONNECT shall cooperate in discussing the progress of the Project on an as-needed basis.

3. PAYMENTS AND FEES

Nothing in this Agreement shall be construed as providing for compensation to either party for its services under this Agreement, and each party shall bear the costs of performing its obligations hereunder.

4. NON-DISCLOSURE

Except as set forth in Section 2.1 herein, in connection with the creation and/or use of OCR Results, the OCR Plain Text, EEBO Collection page images, the Typed Plain Text and /or Metadata for purposes of conducting the Project, REKN/18THCONNECT shall not, during the term of this Agreement or any time thereafter, disclose, or permit any of their employees, agents, or assignees to disclose, to any other person or entity any Confidential Information (as defined below) of the other party. Confidential information does not include information which (a) was or becomes generally available to the public other than as a result of a disclosure by the receiving party or its representatives or (b) was or becomes available to the receiving party on a non-confidential basis from a source other than the disclosing party or its advisers, provided that such source was not known by the receiving party to be bound by any agreement to keep such information confidential, or otherwise prohibited from transmitting the information to the receiving party by a contractual, legal or fiduciary obligation.

5. PUBLICITY

Either party may issue a press release regarding the existence of this Agreement, but not without prior written authorization from the other party. Such authorization shall not be unreasonably withheld, and the parties shall make best efforts to reply to such requests within ten business days.

6. INTELLECTUAL PROPERTY

Subject to the intellectual property rights of third parties, PROQUEST shall retain all rights to the EEBO Collection, the EEBO Collection page images, and the OCR Results. Further, PROQUEST shall have the right to determine the use of the EEBO Collection page images, the Typed Plain Text, the OCR Results, the OCR Plain Text and Metadata, and will have unlimited access to the Open Source Libraries. Except as expressly granted herein, under no circumstances shall anything in this Agreement be construed as granting, by implication, estoppel, or otherwise, to one party a right or license to the other party's name, logo, design marks, trade names, or service marks (collectively, the "Trademarks"), issued patents and patent applications, copyrights and copyrights registration and applications, rights in ideas, designs, works of authorship, derivative works, or any other rights or license relating to the other party's intellectual property or services.

7. TERM AND TERMINATION

7.1 This Agreement shall continue in effect for the length of the Project, unless otherwise terminated in writing by a party or by the parties mutually, with or without cause.

7.2 Upon termination of this Agreement, all rights granted hereunder shall be terminated, except that the rights obligations set forth in Sections 2.4, 2.5, 4, 5, 6 and 9 shall survive termination, as shall any other provisions which by their terms are intended to survive termination. On termination REKN/18THCONNECT immediately shall remove or deactivate all links within its control built using the Metadata, cease using PROQUEST Trademarks in association with REKN's/18THCONNECT's product, and forthwith provide a written statement that the aforementioned deactivation has taken place.

8. USE OF TRADEMARK

REKN/18THCONNECT and PROQUEST agree to adhere to standards consistent with the high level of quality associated with one another's Trademarks. Neither party shall use any Trademarks or other database content to make it appear that the other party is endorsing, sponsoring, or recommending the information, goods, or services in its web site. Neither party shall suggest the other party is affiliated with it, its advertisers, or other entities to which it is linked, and neither party will otherwise misrepresent its relationship with the other party or present false or misleading information about the other party's products or services. Neither party shall use the Trademarks, or other materials, in a manner that is likely to cause confusion with, dilute, or damage the reputation of the other party or its database. Neither party shall place the other party's web pages in a "frame" within that party's web site, use other techniques that alter or obstruct the visual or other presentation of the other party's database in whole or in part, or otherwise impose editorial comment or commercial material or any other type of identification on or in proximity to content displayed in the other party's database without written permission from an authorized representative of the other party. Neither party shall use the name or logo of the other party, its participants, or content in any "metatag" without the other party's express written permission.

9. NOTICES

All notices given pursuant to this Agreement shall be in writing and sent to the Notice Address specified below. These addresses also are applicable for general contact between the parties. Notices may be delivered by hand, by overnight carrier, or shall be deemed to be received within five (5) business days after mailing if sent by registered or certified mail, postage prepaid. If any notice is sent by facsimile or email, then confirmation copies must be sent as specified above. Either party may from time to time change its Notice Address by written notice to the other party.

If to PROQUEST :

Mary Sauer-Games
Vice President, Publishing
ProQuest LLC
789 E. Eisenhower Parkway
Ann Arbor, MI 48108
Tel: 734-761-4700
Email: Mary.Sauer-Games@proquest.com

With cc: to General Counsel

If to REKN/18THCONNECT:

Laura Mandell
Director of ARC (organization overseeing REKN/18THCONNECT)
Director, Initiative for Digital Humanities, Media, and Culture
TAMU 4227
Texas A&M University
College Station, TX 77843-4227 U.S.A.
(tel): (979) 845-8345
(fax): (979) 862-2292
(email): mandell@tamu.edu

10. REPRESENTATIONS AND WARRANTIES

10.1 Each party hereby represents and warrants that it is duly organized and validly subsisting and has full authority and all rights necessary to enter into and perform this Agreement and to bind the party to the terms and conditions herein. Each party further represents and warrants that it has caused this Agreement to be executed by a duly authorized representative.

10.2 OTHER THAN THE EXPRESS WARRANTIES STATED IN THIS SECTION 10, USE OF THE OCR RESULTS, THE OCR PLAIN TEXT, THE METADATA , AND EEBO COLLECTION, ARE PROVIDED ON AN “AS IS” AND “AS AVAILABLE” BASIS. EACH PARTY DISCLAIMS ANY AND ALL OTHER WARRANTIES, CONDITIONS, OR REPRESENTATIONS (EXPRESS, IMPLIED, ORAL OR WRITTEN),RELATING THE METADATA , THE OCR RESULTS, THE OCR PLAIN TEXTAND EEBO COLLECTION OR ANY PART THEREOF, INCLUDING, WITHOUT LIMITATION, ANY AND ALL IMPLIED WARRANTIES OF QUALITY, PERFORMANCE, COMPATIBILITY, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. NEITHER PARTY WARRANTS THE USE OF THEIR PRODUCTS , SOFTWARE OR SYSTEMS WILL BE UNINTERRUPTED OR ERROR-FREE, N OR DOES EITHER PARTY MAKE ANY WARRANTY AS TO THE AVAILABILITY OF THEIR SERVICES OR SOFTWARE THE ACCURACY, TIMELINESS, CORRECTNESS, RELIABILITY, CURRENCY OR COMPLETENESS OF THE INFORMATION OR THE RESULTS OF USE OF THE SERVICES.

NEITHER PARTY MAKES ANY WARRANTIES RESPECTING IMPROPER USAGE OF ITS RESPECTIVE DATABASE BY USERS..

10.3 PROQUEST shall not be liable for any loss, injury, claim, liability or damage of any kind resulting from the unavailability of its database, interruption of the services provided hereunder, or arising out of or in connection with the its linking activities.

- 10.4 Neither party shall be liable for any indirect, special, incidental, punitive or consequential damages, including but not limited to loss of data, business interruption, or loss of profits, even if advised of the possibility of a claim. Neither party shall have any responsibility or liability for any content appearing in the other party's database or for any link to or from third party web sites.

11. MISCELLANEOUS

11.1 This Agreement constitutes the entire agreement of the parties and supersedes all prior communications, understandings and agreements relating to the subject matter hereof, whether oral or written. No modification or claimed waiver of any provision of this Agreement shall be valid except by written amendment signed by authorized representatives of PROQUEST and Licensee.

11.2 This Agreement and any amendments may be executed in one or more counterparts, each of which shall be deemed an original, but all of which together shall constitute one agreement.

11.3 Nothing contained herein shall be deemed to create an agency, joint venture, or partnership relationship between the parties. Neither party shall use or attempt to exercise apparent authority concerning the other party in its dealings with third parties.

11.4 Neither party shall be liable in damages or have the right to cancel this Agreement for any delay or default in performing hereunder if such delay or default is caused by conditions beyond its control including, but not limited to Acts of God, Government restrictions (including the denial or cancellation of any export or other necessary license), wars, insurrections, strikes, fires, floods, work stoppages, unavailability of materials, carriers or communications facilities, and/or any other cause beyond the reasonable control of the party whose performance is affected.

ARC on behalf of REKN /
18THCONNECT

PROQUEST LLC

By: Lana Mandell

By: Mary Sue Barnes

Title: Director, ARC

Title: Vice President

Date: September 5, 2012

Date: September 4, 2012

Rev. 10Aug2012/JD

Appendix 1

CREATION AND USE OF OCR PLAIN TEXT BY 18THCONNECT MEMBERS

A.1 DEFINITIONS

A.1.1 “Crowd-sourced Editing Platform” shall mean any open-source interface designed to allow users to see a word unreadable in OCR Results in the context of one line on either side, currently but not limited to TypeWright.

A.1.2 “Editor” shall mean anyone appointed by REKN/18THCONNECT to approve and accept text corrections proposed by users of REKN/18THCONNECT through its Crowd-Sourced Editing Platform.

A.1.3 “Users” shall mean any registered user of REKN/18THCONNECT.

A.1.4. “OCR Plain Text” shall mean the text of a complete work or portion of a work in the EBBO Collection that has been OCR’d and corrected by users of REKN/18THCONNECT.

A.1.5. “Full Editions” shall mean any and all editorial and historical content added through a textual apparatus to the “OCR Plain Text” by an editor after having fully corrected the document.

A.2 GRANT OF LICENSE

A.2.1 PROQUEST allows REKN/18THCONNECT to display OCR Results and/or the Typed Plain Text along with the scanned EBBO Collection images in its “Crowd-sourced Editing Platform.”

A.2.2 Text corrected in the “Crowd-sourced Editing Platform” and approved by an Editor will be returned by REKN/18THCONNECT to PROQUEST per an agreed-upon schedule but not less than annually, for use within ProQuest’s products and services, including the EBBO database at no cost/obligation to PROQUEST, should PROQUEST choose to use the OCR Plain Text (OCR Results + text-correction by users). The right to use the OCR Plain Text as granted above is a perpetual, unrestricted right in any format, whether in existence now or developed in the future, and the right shall explicitly survive termination of this Agreement.

A.2.3 PROQUEST grants to Users the right to correct OCR Results of EBBO Collection texts in exchange for such User obtaining a copy of the Full Text of agreed-upon documents.

A.2.4. Users who have corrected OCR Results of images from the EBBO Collection and thereby created OCR Plain Text of such images or objects may publish electronic versions of those OCR Plain Text s by submitting them to peer review at REKN/18THCONNECT.

A.2.4 PROQUEST shall have first right of refusal to publish in print ☐ on ☐ demand or electronic form the Full Editions of corrected texts that have been peer ☐ reviewed by

REKN/18THCONNECT, and may nonetheless have a non-exclusive right to publish at any time the OCR Plain Text that had been corrected by that editor.