Web of Science Data Access Policy

FULL POLICY CONTENTS

Effective: *9/26/2016*
Last Updated: *2/6/2018*

Responsible Office:
Indiana University Network Science Institute

Responsible Administrator
Patricia Mabry, Executive Director, IUNI

Policy Contact:
Matthew Hutchinson iuniwosd@indiana.edu

Scope

All agents of the university who wish to access the WoS dataset stored on IUNI's data enclave.

Policy Statement

IUNI has entered into a contract with Clarivate Analytics to receive a copy of the 'Web of Science' (WoS) dataset that contains over sixty-one million records documenting scholarly research from the end of the nineteenth century up to and including 2015. IUNI has been tasked with making this data accessible while simultaneously protecting the proprietary nature of the dataset and honoring the agreement with Clarivate Analytics. To this end, the data will be stored on a discrete enclave on a dedicated Data Intensive Node within the Karst computing cluster where access can be managed. This document describes the rules Indiana University employees must follow in order to access the WoS dataset.

- The data is accessible via three distinct modes all of which are available on the enclave:
    - Raw XML data
    - A relational PostgreSQL database created by parsing the raw XML data
    - A browser based GUI to be developed within the enclave

Reason for Policy

IUNI wants to make the 'Web of Science' (WoS) data, purchased from Clarivate Analytics available to all eligible employees (as defined in DM-01) while adhering to the terms and conditions of the agreement with Clarivate Analytics.

In consultation with the Research Data Steward, the dataset has been categorized as **'Restricted'** as defined by DM-01. The protection of this proprietary data is a key priority for the institute and any violation may result in a termination of the contract with Clarivate Analytics; resulting in significant challenges for research using publication data throughout the University.

The goal of the data access policy is to maximize accessibility while simultaneously ensuring adequate controls over the data and remaining in compliance with the Clarivate Analytics' licensing agreement.

Procedures

1.1. The data enclave will be accessible for all eligible employees of Indiana University (as defined in DM-01). Except where there is a compelling reason not to do so, it is the intention of IUNI to approve all applications from individuals who meet these criteria and who also complete the actions required to comply with this policy.

1.2. Students who are not employed by the University are not permitted to access or work with the data.

1.3. The Office of the Vice-President and General Counsel has final authority regarding interpretation of the license agreement and whether the proposed uses of the data comport with the agreement.

1.4. Analysis and research can be conducted within the enclave using applications already installed on the Karst cluster. Additional applications requested by users can be installed on the enclave at the Research Data Steward's discretion (as required by DM-02). A request for new software should be submitted to the IUNI Data Manager for consideration. The Data Manager will work with the Research Data Steward and UITS to determine whether the request can be honored. All persons submitting a request will receive a response regarding the disposition of the request in a timely manner.

1.5. Analysis and research can also be conducted on the indexed data in the PostgreSQL database by either submitting direct SQL queries or by using the browser based GUI from inside the enclave.

1.6. Analysis and research requiring more computational resources than the current Data Enclave provides, can be performed by the IUNI Data Manager on users' behalf using the University's advanced computational systems (e.g. Karst, Big Red II, Mason, etc.). A request must be submitted to the Data Manager and will be reviewed before the request is executed. In some cases, it may not be possible

to fulfill the request or it may only be possible to partially fulfill a request. The results of these jobs are subject to review by the Data Manager.

1.7. To gain access to the data enclave, individuals must submit an application to the IUNI Data Manager via the IUNI website. As part of the application, individuals must select the level of access they require. There are two levels of access to the WoS data: General Data User and Data Custodian:

1.7.1. General Data User:

1.7.1.a. The General Data User classification represents the majority of individuals accessing the data.

1.7.1.b. General Data Users cannot themselves export any material from the enclave; they are limited to conducting research on the data using the tools installed on Karst. Users with this classification may request the export of research products such as charts and summary tables, provided these products do not represent a more than insubstantial part of the content.

1.7.1.c. Requests to export any material from the enclave are submitted to the IUNI Data Manager. Requests will be reviewed to make sure they do not represent a violation of the terms of use. This review may involve consultation with, but is not limited to, the Indiana University Office of the Vice President and General Counsel, the Committee of Data Stewards and Clarivate Analytics before determining whether or not a request is approved.

1.7.1.d. If the Data Manager is not available, the IUNI Director of IT will review the request. If the request is approved, the UITS System Administrator will extract the data.

1.7.1.e. An individual may appeal the judgement of the Data Manager if they believe their request falls within the terms of the agreement with Clarivate Analytics. The individual should contact the Research Data Steward, Heather Coates (hcoates@iupui.edu). The request will be reviewed by Indiana University's Committee of Data Stewards who will provide an opinion regarding the appeal. If needed, the opinion of the IU Office of the Vice-President and General Counsel will be obtained for a final decision.

1.7.2. Data Custodian:

1.7.2.a. IUNI recognizes that the limits of the data enclave make it unsuitable for certain kinds of research. For users for whom the data enclave is insufficient, IUNI has created a secondary user classification: Data Custodian.

1.7.2.b. Data Custodians are permitted to export research products directly from the enclave, in addition to the privileges granted to a General Access User. IUNI expects that research teams who require this level of access will only need one Data Custodian for their group.

1.7.2.c. Any data removed from the enclave by a Data Custodian is still subject to IU's information policies regarding **Restricted** data (DM-01) and the terms of the licensing agreement with Clarivate Analytics. For the full list of restrictions please contact the IUNI Data Manager.

1.7.2.d. An individual applying for Data Custodian classification must submit to IUNI the following documentation:

1.7.2.d.i. A detailed description explaining why the General User Role is unsuitable for the proposed research project and how the applicant is qualified to administer the proprietary WoS data.

1.7.2.d.ii. A written statement agreeing to comply with all terms and restrictions required by Clarivate Analytics and the Data Use Agreement written by IUNI.

1.7.2.d.iii. If the Data Custodian intends to store their copy of the WoS data on a system that is not part of the Indiana University Data Center, they must submit a completed copy of the Institutional Data Standards checklist describing the security and administration of the server on which the data will be stored. If the applicant is not the System Administrator for the target server, the individual named in the System Administrator Agreement (see 1.7.2.d.iv) should complete the checklist. The checklist can be found at the following address: https://datamgmt.iu.edu/tools/checklist-for-system-providers.php

1.7.2.d.iv. If the System Administrator of the server on which the WoS data will be stored is not the applicant for Data Custodian status, the System Administrator must submit the 'System Administrator Agreement' form on the IUNI website.

1.7.2.e. The Data Custodian must also agree to the following:

1.7.2.e.i. Adhere to all rules outlined in this policy and its subsequent revisions. An up-to-date policy is always available at http://www.iuni.iu.edu/resources/wos_policy.html

1.7.2.e.ii. Maintain log files recording what they have exported from the enclave. These files must be maintained in the enclave in a location that can be accessed by the IUNI Data Manager whenever necessary.

1.7.2.e.iii. Notify the IUNI Data Manager and it-incident@iu.edu of any suspected systems breach or data exposure as well as following the steps outlined in REPORTING SUSPECTED SENSITIVE DATA EXPOSURES policy at https://protect.iu.edu/online-safety/report-incident/sensitive-data-breaches.html

1.7.2.e.iv. Store any data removed from the enclave on a system that is compliant with the University's standards for storage of **Restricted** data.

1.7.2.e.v. Failure to perform any of the tasks agreed upon in this section may result in the revocation of Data Custodian status.

2.1. All users are expected to notify the IUNI Data Manager about any use of the data or derived results in publications or funding proposal submissions and include a citation for the IUNI/Clarivate Analytics data. Format for the citation will be posted to the IUNI WoS data website.

2.2. An individual or group that uses the XML version of the data as it was purchased from Clarivate Analytics should include the following acknowledgement in the

published work: *"This work uses Web of Science data by Clarivate Analytics provided by the Indiana University Network Science Institute."*

2.3. An individual or group that uses the database in the enclave should include the following acknowledgement in the published work: *"This work uses Web of Science data by Clarivate Analytics provided by the Indiana University Network Science Institute and the Cyberinfrastructure for Network Science Center at Indiana University."*

Definitions

- **IUNI**: The Indiana University Network Science Institute (iuni.iu.edu)
- **Eligible Employees**: Persons with a professional affiliation with Indiana University as either faculty or staff (as defined in DM-01).
- **Web of Science or WoS**: The dataset and all future updates that have been provided to IUNI by Clarivate Analytics.
- **Data Enclave or 'The enclave'**: A dedicated data intensive node in the Karst network that stores the WoS dataset and is maintained by UITS on behalf of IUNI.
- **Data Manager**: An employee of IUNI who conducts day-to-day administration of the WoS dataset and supports researchers in accessing the data. The Data Manager serves as the single point of contact for all issues regarding the enclave and the WoS data (See https://datamgmt.iu.edu/governance/structure.php for more details).
- **Data Custodian**: A special class of user permitted to remove material from the enclave (See https://datamgmt.iu.edu/governance/structure.php for more details).
- **General Data User**: An individual granted access to the enclave to conduct research

Sanctions

Indiana University will handle reports of misuse and abuse of information and information technology resources in accordance with existing policies and procedures issued by appropriate authorities. Depending on the individual and circumstances involved this could include the offices of Human Resources, Vice Provost or Vice Chancellor of Faculties (or campus equivalent), Dean of Students (or campus equivalent), Office of the General Counsel, and/or appropriate law enforcement agencies. See policy IT-02, Misuse and Abuse of Information Technology Resources for more detail.

Failure to comply with Indiana University information technology policies may result in sanctions relating to the individual's use of information technology resources (such as suspension or termination of access, or removal of online material); the individual's

employment (up to and including immediate termination of employment in accordance with applicable university policy); the individual's studies within the university (such as student discipline in accordance with applicable university policy); civil or criminal liability; or any combination of these.

Additional Contacts

Subject	*Contact*	*Phone*	*Email*
Web of Science	**Matthew Hutchinson**	**(812) 855-1404**	**iuniwosd@indiana.edu**

Forms

If you are interested in accessing the WoS data as a General Data User please complete the application form at the following address:
http://www.iuni.iu.edu/forms/f/form/wos_data_user?casticket=ST-870917-DGAus63pTLalw2oQ6U4ncasprd01

If you are interested in accessing the WoS data as a Data Custodian, please complete the application form at the following address:
http://www.iuni.iu.edu/forms/f/form/wos_data_custodian

If you are a system administrator who will be responsible for managing a copy of the WoS data but you are not applying for Local Data Custodian status, please complete the application form at the following address:

History

This policy replaces the ‘Web of Science - Interim Data Access Policy’.