Version Number: 01-2020

U.S. Department of Commerce
U.S. Patent and Trademark Office

Privacy Threshold Analysis
for the
Reed Tech Patent Data Capture (PDCap)

U.S. Department of Commerce Privacy Threshold Analysis

USPTO Reed Tech PDCap

Unique Project Identifier: PTOC-013-00

Introduction: This Privacy Threshold Analysis (PTA) is a questionnaire to assist with determining if a Privacy Impact Assessment (PIA) is necessary for this IT system. This PTA is primarily based from the Office of Management and Budget (OMB) privacy guidance and the Department of Commerce (DOC) IT security/privacy policy. If questions arise or further guidance is needed in order to complete this PTA, please contact your Bureau Chief Privacy Officer (BCPO).

Description of the information system and its purpose: *Provide a general description system (in a way that a non-technical person can understand) of the information system that addresses the following elements:*

The E-Government Act of 2002 defines "information system" by reference to the definition section of Title 44 of the United States Code. The following is a summary of the definition: "Information system" means a discrete set of information resources organized for the collection, processing, maintenance, use, sharing, dissemination, or disposition of information. See: 44. U.S.C. § 3502(8).

The Reed Tech Patent Data Capture (PDCap) system is designed to process, transmit and store data and images to support the data-capture and conversion requirements of the USPTO patent application process. Patent applications are typically submitted to USPTO on paper (hard copy) or in electronic format. Under the Patent Data Capture contract, Reed Tech hosts and manages the PDCap system and is required to convert the paper applications into an electronic format, including all text, graphics, artwork, drawings, etc. Once converted to electronic data, each patent is composed and formatted to USPTO specifications for delivery back to USPTO. The

Reed Tech Published Application Alert Service (PAAS) is a service offered by the USPTO to allow the public to configure queries and alerts for key words in pre-grant published patent applications.

a) Whether it is a general support system, major application, or other type of system
Reed Tech PDCap is a major application and a contractor system.

b) ***System location***
Primary location: 7 Walnut Grove Drive, Horsham, PA 19044
Secondary location: 2331 Mill Road, Suite 300, Alexandria, VA 22314

c) Whether it is a standalone system or interconnects with other systems (identifying and describing any other systems to which it interconnects)

It is not a standalone system. A point-to-point DS3 provides connectivity between the Horsham, PA, location and the Alexandria, VA, location. A point-to-point DS3 is a wide area network (WAN) connection to USPTO. This connection is owned and managed by USPTO, as is the equipment hosting the connection. It is an extension of the IFW network. There is a fail-over VPN connection between the two sites that provides fault tolerance for the point-to point connection. Additionally Reed Tech has subcontractors supporting PDCap and transfer USPTO data to the subcontractor locations. Traffic permitted into the network from the internet includes the following traffic to servers in the De-Militarized Zone (DMZ):

- Encrypted FTP traffic to public FTP server
- Remote access VPN
- Public email
- Backup point-to-point VPN (Internet Protocol Security [IPSEC])
- Alexandria location (only used when primary link fails).

d) The purpose that the system is designed to serve

- The Reed Tech Patent Data Capture (PDCap) system is designed to process, transmit and store data and images to support the data-capture and conversion requirements of the USPTO to support the USPTO patent application process.
- The Reed Tech Published Application Alert Service (PAAS) is a service offered by the USPTO to allow the public to configure queries and alerts for key words in pre-grant published patent applications.

e) The way the system operates to achieve the purpose

- PDCap – When both hardcopy and electronic patent applications are initially received at the USPTO, the documents are scanned/uploaded respectively into the Image File Wrapper (IFW) system. Applications are electronically exported to the Reed Tech PDCap system via a USPTO-managed interconnection. Once received by Reed Tech PDCap, every application is then examined by a Reed Tech proprietary application which breaks down each page into separate sections, such as graphics and text. Each section is then sent to separate directories on the Reed Tech PDCap network for manipulation. The sections of the application are processed by separate Reed Tech PDCap departments, with departments dedicated to text, headers, and complex work units, such as math, chemistry, and drawings. These departments use a combination of proprietary and commercial software to complete their work on each section. When all the sections have been completed, a queue reassembles the file and it is forwarded to the Composition Department. The Composition Department is responsible for the final formatting, layout, and any remaining error corrections before the file is delivered back to USPTO. There are several phases to the overall process: PreGrant Publication (PGPub), Initial Data Capture (IDC), File Maintenance (FM), and Final Data Capture (FDC).

- PAAS – The system operates to achieve its purpose as follows. After a logged-in user creates a keyword search, it will be executed on a weekly basis against only the most recent pre-grant published patent applications. The queries will be executed at the date and time of the publication of the data by the USPTO. The data that will be used for searching will be copied out of the main PDCap system onto a file system on or attached to the backend server. The queries will be run against the data on that file system and not within the main PDCap file system. After the queries are executed, the data for that week's pre-grant published patent applications will be deleted from the file system on or attached to the backend server. After the queries are executed, an email alert will be sent to the user's email address, which will be part of the profile created during registration. Queries against patent applications older than the most recent publication date will not be possible, as prior publication data is removed from the system after the weekly search is executed.

f) A general description of the type of information collected, maintained, used, or disseminated by the system

- PDCap – The type of information collected, maintained, used, or disseminated by the system are Patent applications data that includes text, graphics, artwork, drawings, math equations, chemistry equations, etc.
- PAAS – The types of information collected, maintained, used, or disseminated by the system are configured queries and alerts for specific key words in pre-grant published patent applications.

g) Identify individuals who have access to information on the system

- PDCap – Users include full and part-time Reed Tech employees. Reed Tech also utilizes several subcontractors to support portions of the patent data capture process. Subcontractors connect to a Secure Shell (SSH) File Transfer Protocol (SFTP) server, using SSH client software.
- PAAS - PAAS information consists of the published PreGrant data which is public information. Majority of the users who use this information are public users and other identified users who support the application.

h) How information in the system is retrieved by the user

- PDCap - Information in the system is retrieved by the user after the patent applications are electronically exported to the Reed Tech PDCap system via a USPTO-managed-interconnection. Every application is then examined by a Reed Tech proprietary application which breaks down each page into separate sections, such as graphics and text. Each section is then sent to separate directories on the Reed Tech PDCap network

for manipulation by the different departments dedicated to text, headers, and complex work units such as math, chemistry, and drawings.

- PAAS – Information in the system is retrieved by the user after a logged-in user creates a keyword search, which will be executed on a weekly basis against only the most recent pre-grant published patent application.

i) How information is transmitted to and from the system

Patent applications are sent to and from the PDCap system via Secured File Transfer Protocol (SFTP).

Questionnaire:

1. Status of the Information System

1a. What is the status of this information system?

☐ This is a new information system. *Continue to answer questions and complete certification.*

☐ This is an existing information system with changes that create new privacy risks. *Complete chart below, continue to answer questions, and complete certification.*

Changes That Create New Privacy Risks (CTCNPR)					
a. Conversions	☐	d. Significant Merging	☐	g. New Interagency Uses	☐
b. Anonymous to Non-Anonymous	☐	e. New Public Access	☐	h. Internal Flow or Collection	☐
c. Significant System Management Changes	☐	f. Commercial Sources	☐	i. Alteration in Character of Data	☐
j. Other changes that create new privacy risks (specify):					

☐ This is an existing information system in which changes do not create new privacy risks, and there is not a SAOP approved Privacy Impact Assessment. *Continue to answer questions and complete certification.*

☒ This is an existing information system in which changes do not create new privacy risks, and there is a SAOP approved Privacy Impact Assessment (version 01-2015 or 01-2017). *Continue to answer questions and complete certification.*

☐ This is an existing information system in which changes do not create new privacy risks, and there is a SAOP approved Privacy Impact Assessment (version 01-2019 or later). *Skip questions and complete certification.*

1b. Has an IT Compliance in Acquisitions Checklist been completed with the appropriate signatures?

☐ Yes. This is a new information system.

☐ Yes. This is an existing information system for which an amended contract is needed.

☐ No. The IT Compliance in Acquisitions Checklist is not required for the acquisition of equipment for specialized Research and Development or scientific purposes that are not a National Security System.

☒ No. This is not a new information system.

2. Is the IT system or its information used to support any activity which may raise privacy concerns?

NIST Special Publication 800-53 Revision 4, Appendix J, states "Organizations may also engage in activities that do not involve the collection and use of PII, but may nevertheless raise privacy concerns and associated risk. The privacy controls are equally applicable to those activities and can be used to analyze the privacy risk and mitigate such risk when necessary." Examples include, but are not limited to, audio recordings, video surveillance, building entry readers, and electronic purchase transactions.

☐ Yes. *(Check all that apply.)*

Activities			
Audio recordings	☐	Building entry readers	☐
Video surveillance	☐	Electronic purchase transactions	☐
Other (specify):			

☒ No

3. Does the IT system collect, maintain, or disseminate business identifiable information (BII)?

As per DOC Privacy Policy: "For the purpose of this policy, business identifiable information consists of (a) information that is defined in the Freedom of Information Act (FOIA) as "trade secrets and commercial or financial information obtained from a person [that is] privileged or confidential." (5 U.S.C.552(b)(4)). This information is exempt from automatic release under the (b)(4) FOIA exemption. "Commercial" is not confined to records that reveal basic commercial operations" but includes any records [or information] in which the submitter has a commercial interest" and can include information submitted by a nonprofit entity, or (b) commercial or other information that, although it may not be exempt from release under FOIA, is exempt from disclosure by law (e.g., 13 U.S.C.)."

☒ Yes, the IT system collects, maintains, or disseminates BII.

☐ No, this IT system does not collect any BII.

4. Personally Identifiable Information (PII)

4a. Does the IT system collect, maintain, or disseminate personally identifiable information (PII)?

As per OMB 17-12: "The term PII refers to information that can be used to distinguish or trace an individual's identity either alone or when combined with other information that is linked or linkable to a specific individual."

☒ Yes, the IT system collects, maintains, or disseminates PII about: *(Check all that apply.)*

- ☐ DOC employees
- ☐ Contractors working on behalf of DOC
- ☐ Other Federal Government personnel
- ☒ Members of the public

☐ No, this IT system does not collect any PII.

If the answer is "yes" to question 4a, please respond to the following questions.

4b. Does the IT system collect, maintain, or disseminate Social Security numbers (SSNs), including truncated form?

☐ Yes, the IT system collects, maintains, or disseminates SSNs, including truncated form.

Provide an explanation for the business need requiring the collection of SSNs, including truncated form.
Provide the legal authority which permits the collection of SSNs, including truncated form.

☒ No, the IT system does not collect, maintain, or disseminate SSNs, including truncated form.

4c. Does the IT system collect, maintain, or disseminate PII other than user ID?

☒ Yes, the IT system collects, maintains, or disseminates PII other than user ID.

☐ No, the user ID is the only PII collected, maintained, or disseminated by the IT system.

4d. Will the purpose for which the PII is collected, stored, used, processed, disclosed, or disseminated (context of use) cause the assignment of a higher PII confidentiality impact level?

Examples of context of use include, but are not limited to, law enforcement investigations, administration of benefits, contagious disease treatments, etc.

☐ Yes, the context of use will cause the assignment of a higher PII confidentiality impact level.

☒ No, the context of use will not cause the assignment of a higher PII confidentiality impact level.

If any of the answers to questions 2, 3, 4b, 4c, and/or 4d are "Yes," a Privacy Impact Assessment (PIA) must be completed for the IT system. This PTA and the approved PIA must be a part of the IT system's Assessment and Authorization Package.

CERTIFICATION

☒ I certify the criteria implied by one or more of the questions above **apply** to the Reed Tech PDCap and as a consequence of this applicability, I will perform and document a PIA for this IT system.

☐ I certify the criteria implied by the questions above **do not apply** to the Reed Tech PDCap and as a consequence of this non-applicability, a PIA for this IT system is not necessary.

System Owner Name: Susan Scanlon Office: Office of Data Management Phone: (703) 756-1561 Email: Susan.Scanlon@uspto.gov Signature: Users, Scanlon, Susan Digitally signed by Users, Scanlon, Susan Date: 2021.01.27 12:34:45 -05'00' Date signed: _____	**Chief Information Security Officer** Name: Don Watson Office: Office of the Chief Information Officer (OCIO) Phone: (571) 272-8130 Email: Don.Watson@uspto.gov Signature: DON R Watson Digitally signed by DON R Watson Date: 2021.01.28 12:55:00 -05'00' Date signed: _____
Privacy Act Officer Name: John Heaton Office: Office of General Law (O/GL) Phone: (571) 270-7420 Email: Ricou.Heaton@upsto.gov Signature: Users, Heaton, John (Ricou) Digitally signed by Users, Heaton, John (Ricou) Date: 2021.01.26 14:24:48 -05'00' Date signed: _____	**Bureau Chief Privacy Officer and Co-Authorizing Official** Name: Henry J. Holcombe Office: Office of the Chief Information Officer (OCIO) Phone: (571) 272-9400 Email: Jamie.Holcombe@uspto.gov Signature: Users, Holcombe, Henry Digitally signed by Users, Holcombe, Henry Date: 2021.01.28 14:33:46 -05'00' Date signed: _____
Co-Authorizing Official Name: Andrew Faile Office: Office of the Commissioner for Patents Phone: (571) 272-8800 Email: Andrew.Faile@uspto.gov Signature: Users, Faile, Andrew Digitally signed by Users, Faile, Andrew Date: 2021.02.02 11:29:56 -05'00' Date signed: _____	