U.S. Securities and Exchange Commission

Enterprise Human Capital Repository (EHCR)
PRIVACY IMPACT ASSESSMENT (PIA)

August 22, 2017

Office of Human Resources

Section 1: System Overview

1.1 Name of Project or System

Enterprise Human Capital Repository (EHCR)

1.2 Is the system internally or externally hosted?

☒ Internally Hosted (SEC) — Office of Human Resources sponsors EHCR

☐ Externally Hosted (Contractor or other agency/organization) — If the system is externally hosted, please list the Division or Office.

1.3 Reason for completing PIA

☐ New project or system

☒ This is an existing system undergoing an update

First developed: 5/2015

Last updated: 12/2016 Click here to enter a date.

Description of update: In Release III, beginning 7/1/2016, new data was ingested to augment worker data and EHCR data was published formally to other data sources. Release III was deployed 12/2016.

1.4 Does the system or program employ any of the following technologies?

☒ Electronic Data Warehouse (EDW)

☐ Social Media

☐ Mobile Application (or GPS)

☐ Cloud Computing Services

☐ www.sec.gov Web Portal

☐ None of the Above

Section 2: Authority and Purpose of Collection

2.1 Describe the project and its purpose or function in the SEC's IT environment

The Enterprise Human Capital Repository (EHCR) is a single centralized repository for personnel data for the SEC workforce (employees, contractors, interns, Intergovernmental Personnel Act (IPA) employees, and fellows) and other individuals that may support the Commission. The repository resides on the Enterprise Data Warehouse (EDW) platform and is a part of the EDW initiative. EHCR is a "system of truth"[1] and will not be the system of record for any data. Source systems will remain as the system of record for the relevant data.

The data are ingested from source systems using SEC's enterprise Extract, Transform and Load (ETL) tool, Infosphere Information Server (IIS). Source Systems and data include the following:

- US Access: Worker Information including enrollment ID and status. Will provide initial identity record for all employees and contractors.
- Workforce Tracking and Transformation System (WTTS): Worker Information including pay position, grade, salary, clearances, and department. Data is retrieved from Department of Interior (DOI) data-mart.
- DOI, Federal Personnel/Payroll System (FPPS): Worker Information including work schedule, security clearances, position information and work schedule.
- Lead, Engage, Achieve and Perform (LEAP) learning platform: Contains information about training courses taken and completed by workers and contractors.

[1] Single System of Truth: The practice of structuring information models and associated schemata such that every data element is stored exactly once. Any possible linkages to this data element (possibly in other areas of the relational schema or even in distant federated databases) are by reference only. Because all other locations of the data just refer back to the primary "source of truth" location, updates to the data element in the primary location propagate to the entire system without the possibility of a duplicate value somewhere being forgotten.

- Active Directory (AD): Contains additional worker related information such as the username AD ID, which is used for system authentication.
- Contractor Personnel (CP) List: Repository for Contractor related information. Data includes name, email, office location, and phone number for CP and contractor project manager or POC.
- Archibus: Contains employee and contractor workplace location information. Data includes name, building, office/cube assignment, and other related fields.
- HR4ME: Contains employee information for various SEC benefits programs and information on employee certifications, including Certified Public Accountants (CPA) and Bar memberships. Also includes telework agreement details. Data includes full name, work schedule, telework agreement status, and organization name, among others.
- WebTA: Employee supervisor information. Data includes employee and supervisor name and organization.

The Office of Human Resources (OHR) relies on multiple external system feeds that populate several internal databases and applications. Additionally, there are a number of manual processes used to add employee and contractor data to these internal databases and applications, resulting in copies of personnel data residing in multiple locations. The EHCR system is used as a reference repository of all personnel records and maintains a copy of the information that resides in the underlying OHR systems and applications. The underlying systems and applications will maintain the official records.

2.2 What specific legal authorities, arrangements, and/or agreements allow the information to be collected?

All legal authorities and agreements associated with the data maintained in the source system follows the data collected, used, maintained, retrieved, and disseminated within EHCR, to include the following:
15 U.S. C 77a et seq., 78a et seq., 80a-1 et seq., and 80b-1 et seq.; 5 U.S.C. 1302, 2951, 3301, 3372, 4103, 4113, and 4118; and 5 CFR part 410; 5 CFR parts 213, 293, 302, and 335 and Office of Personnel Management Regulations promulgated thereunder; and 5 CFR, parts 213, 293, 302, and 335; 5 U.S.C. 3109 and Civil Service Regulations promulgated thereunder.

2.3 Does the project use, collect, or maintain Social Security numbers (SSNs)? *This includes truncated SSNs.*

☒ Yes

If yes, provide the purpose of collection:	The primary identifier to create a record is the Enrollment ID from US Access System. The US Access System includes the SSN. OHR uses the SSN to match employees with their Department of Interior (DOI) records. The complete SSN is used for that matching. The complete SSN is also used to complete the employee record in EHCR, and is used to match employee information from multiple inbound systems.
If yes, provide the legal authority:	The legal authority to collect the SSN is outlined in the specific System of records notice, including Executive Orders 9397, as amended by 13478, 9830, and 12107; Sections 19(c) and 20(a) of the Securities Act of 1933; Section 21(a) of the Securities Exchange Act of 1934; Section 321(a) of the Trust Indenture Act of 1939; Section 42(b) of the Investment Company Act of 1940; Section 209(b) of the Investment Advisors Act of 1940; and 17 C.F.R. §§ 200.21, 201.102(e), 202.5(a).

2.4 Do you retrieve data in the system by using a personal identifier?

☐ No

☐ Yes, a SORN is in progress

☒ Yes, there is an existing SORN

Since EHCR does not collect information directly from individuals and no new information is created through the system about individuals, the information contained in the source systems performing the original collection is covered by the individual SORNs for those systems as listed below:

1. US Access: GSA/GOVT-7, Federal Personal Identity Verification Identity Management System
2. WTTS: OPM/GOVT-1, General Personal Records; and OPM/GOVT-5, Recruiting, examining and placement records
3. FPPS: OPM/GOVT-1, General Personal Records; and OPM/GOVT-5, Recruiting, examining and placement records
4. Active Directory: SEC -56, Mailing, Contact and Other Lists
5. LEAP: SEC-37, Automated Personnel Management Information System; SEC-39, Personnel Management Employment and Staffing Files; and SEC-40, Office of Personnel Management Training Files
6. CP List: SEC-67, General Information Technology Records
7. Archibus: SEC-46, Identification and Access Control Cards, Special Credentials, Press Passes, and Building Access Control Cards.
8. HR4ME: SEC SEC-39: Personnel Management Employment and Staffing File
9. WebTA: SEC-15: Payroll, Attendance, Retirement and Leave Records

2.5 Is the information covered by the Paperwork Reduction Act of 1995 (PRA)?

☒ No

PRA does not apply to EHCR because EHCR does not collect information directly from members of the public.

2.6 Considering the purpose of the collection, what privacy risks were identified and how were those risks mitigated?

Primary privacy risks are that 1/personal information may be ingested into EHCR without a clear purpose or used from EHCR without clear legal authority; 2/information ingested into EHCR is either unnecessary or excessive; or 3/the information provided for one purpose in the source system, may be used inappropriately once ingested in EHCR. The design of the EHCR system limits the data elements that may be ingested from the primary source system to those elements required for the specific purpose of the data in EHCR. Another primary risk is inadvertent or unauthorized disclosure of personally identifiable information (PII) and other identifying information linking individuals to human resource records. An approved System Security Plan describes the access controls to limit access to staff with a need to know.

Section 3: Data Collection, Minimization, and Retention

3.1 What information is collected, maintained, used, or disseminated about individuals? *Check all that apply.*

☐ The system does not collect, maintain, use, or disseminate information about individuals.

Identifying Numbers

☒ Social Security Number	☐ Alien Registration	☐ Financial Accounts
☐ Taxpayer ID	☐ Driver's License Number	☐ Financial Transactions
☐ Employee ID	☐ Passport Information	☐ Vehicle Identifiers
☐ File/Case ID	☐ Credit Card Number	☐ Employer ID
☐ Other: Click here to enter text.		

General Personal Data

☒ Name	☒ Date of Birth	☐ Marriage Records
☐ Maiden Name	☐ Place of Birth	☐ Financial Information
☐ Alias	☒ Home Address	☐ Medical Information
☐ Gender	☒ Telephone Number	☐ Military Service

☐ Age	☒ Email Address	☐ Mother's Maiden Name
☐ Race/Ethnicity	☐ Education Records	☐ Health Plan Numbers
☐ Civil or Criminal History	☒ Zip Code	
☐ Other: Click here to enter text.		

Work-Related Data

☐ Occupation	☒ Telephone Number	☒ Salary
☒ Job Title	☒ Email Address	☒ Work History
☐ Work Address	☐ Certificate/License Number	☐ Business Associates
☐ PIV Card Information	☐ Fax Number	
☐ Other: Click here to enter text.		

Distinguishing Features/Biometrics

☐ Fingerprints	☐ Photographs	☐ Genetic Information
☐ Voice Recording	☐ Video Recordings	☐ Voice Signature
☐ Other: Click here to enter text.		

System Administration/Audit Data

☐ User ID	☐ Date/Time of Access	☐ ID Files Accessed
☐ IP Address	☐ Queries Ran	☐ Contents of Files
☐ Other: Click here to enter text.		

3.2 Why is the PII listed in Question 3.1 collected, used, shared, or maintained by the system or project?

The data is being collected to create a single "source of truth" record for all employees at the SEC to support a variety of reporting and downstream data needs. EHCR provides structured information and associated data schemata such that every data element is stored exactly once; and updates to the data elements in the primary location propagate to the EHCR system without the possibility of a duplicate value somewhere being forgotten. Any possible linkages to the data elements are by reference only. The data is used for internal OHR reporting and downstream data subscriptions (ex. providing name to source system to ensure names are used consistently across systems).

3.3 Whose information may be collected, used, shared, or maintained by the system?

☒ SEC Employees
Purpose: For truth verification of SEC workers HR information.

☒ SEC Federal Contractors
Purpose: For verifying truth about SEC workers

☒ Interns
Purpose: For truth verification of SEC workers HR information.

☐ Members of the Public
Purpose: Describe the purpose of collecting the information from this source.

☐ Employee Family Members
Purpose: Describe the purpose of collecting the information from this source.

☒ Former Employees
Purpose: For truth verification of SEC workers HR information.

☐ Job Applicants
Purpose: Describe the purpose of collecting the information from this source.

☐ Vendors
Purpose: Describe the purpose of collecting the information from this source.

☐ Other: List other sources of information.
Purpose: Describe the purpose of collecting the information from this source.

3.4 Describe the PII minimizing mechanisms and if the PII from the system is being used for testing, training, and/or research efforts.

EHCR system collects the minimally required PII to manage workforce personnel files through its lifecycle. PII is not used for testing, training, and/or research efforts.

3.5 Has a retention schedule been established by the National Archives and Records Administration (NARA)?

☒ Yes.

The data retention periods for EHCR will vary in nature based on each underlying source system. Data is retained within EHCR as long as indicated by each source system's NARA approved records retention and disposal schedule. The SEC Records Office is currently drafting a Record Retention Schedule for data specific to EHCR, i.e., reports and analyses.

3.6 What are the procedures for identification and disposition at the end of the retention period?

The disposition of records is facilitated by the source system once the records have reached the end of their retention period, as source system updates are replicated in EHCR. The SEC Records Office is currently drafting a record retention and disposition schedule for data specific to EHCR, i.e., reports and analyses.

3.7 Will the system monitor members of the public, employees, and/or contractors?

☐ N/A

☐ Members of the Public

Purpose: If the system or project monitors the members of the public, explain the purpose of the monitoring.

☒ Employees

Purpose: The system contains audit capabilities that track user access to and actions within EHCR. To the extent this functionality constitutes monitoring, the purpose is to ensure the system is accessed and used only by authorized users to further official SEC business.

☒ Contractors

Purpose: The system contains audit capabilities that track user access to and actions within EHCR. Some EHCR users may be contractors supporting OHR activities. To the extent this functionality constitutes monitoring, the purpose is to ensure the system is accessed and used only by authorized users to further official SEC business.

3.8 Considering the type of information collected, what privacy risks were identified and how were those risks mitigated?

Information pertaining to human resource data is sensitive, and access is limited to those who require access to the information. The primary risk is inadvertent or unauthorized disclosure of PII and human resource records on individuals. This risk is mitigated by ensuring that EHCR utilizes granular access controls to protect the data at all levels, and deployment of encryption of data-at-rest in accordance with NIST standards, and role-based access. Additionally, the system will employ audit capabilities to ensure system access and use is appropriate. The system will leverage AD for authentication to further mitigate the potential for unauthorized access of data. All users are required to take mandatory training on Cyber Security and Privacy Awareness, Protecting Nonpublic Information, and Records Management.

Section 4: Openness and Transparency

4.1 What forms of privacy notice were provided to the individuals prior to collection of data? *Check all that apply.*

☐ Privacy Act Statement

Where was the notice provided?

☒ System of Records Notice

The source systems and their applicable SORNs are:

- US Access: GSA/GOVT-7, Federal Personal Identity Verification Identity Management System
- WTTS: OPM/GOVT-1, General Personal Records; and OPM/GOVT-5, Recruiting, examining and placement records
- FPPS: OPM/GOVT-1, General Personal Records; and OPM/GOVT-5, Recruiting, examining and placement records
- Active Directory: SEC -56, Mailing, Contact and Other Lists
- LEAP: SEC-37, Automated Personnel Management Information System; SEC-39, Personnel Management Employment and Staffing Files; and SEC-40, Office of Personnel Management Training Files
- CP List: SEC-67, General Information Technology Records
- Archibus: SEC-46, Identification and Access Control Cards, Special Credentials, Press Passes, and Building Access Control Cards.
- HR4ME: SEC SEC-39: Personnel Management Employment and Staffing File
- WebTA: SEC-15: Payroll, Attendance, Retirement and Leave Record

☒ Privacy Impact Assessment

Date of Last Update: Current PIA

☐ Web Privacy Policy

Where was the notice provided?

☐ Other notice:

What type of notice was provided? Where was the notice provided?

☐ Notice was not provided.

If no notice was provided, please explain why not.

4.2 Considering the method(s) of notice provided, what privacy risks were identified regarding adequate notice and how were those risks mitigated?

Given that the information is covered by multiple SORNs and other notices, the primary risk is inadequate notice. Individuals may not know that information about them is being collected and maintained, who will use it, or the purpose for which the information may be used. However, when submitting their information to the various source systems for EHCR, individuals are provided privacy notices, as appropriate, and they have access to SORNs on the various source systems. Furthermore, this PIA provides additional notice.

Section 5: Limits on Uses and Sharing of Information

5.1 What methods are used to analyze the data?

Data is analyzed via search and reporting capabilities, which may present existing information in the form of reports, graphs, charts, and related management metrics. The system does not otherwise analyze or derive new information.

5.2 Will internal organizations have access to the data?

☒ Yes

Organizations: Authorized staff will have access to the data. Data is aggregated to provide a single repository for ease of use. Data is moved electronically via ETL tools.

5.3 Describe the risk to privacy from internal sharing and describe how the risks are mitigated.

The primary privacy risks associated with internal sharing are inadvertent or unauthorized disclosure of information to individuals without authorization; use of personal information collected without legal authority by receiving office; or disclosure of information for a use not directly related to the primary

purpose of the collection. These risks are mitigated by ensuring that access to the database is limited to EDW EHCR database administrators. Also, end user interface is only via security-controlled reporting access from Business Objects (BOXI). BOXI reporting will be limited to authorized users and all reports will be limited to defined data sets. End user access to EHCR data is controlled by AD groups and only authorized users will have access to that data. Ad hoc User Interface exists in Business Objects.

5.4 Will external organizations have access to the data?

☒ No

5.5 Describe the risk to privacy from external sharing and describe how the risks are mitigated.

Data is not shared with external organizations.

Section 6: Data Quality and Integrity

6.1 Is the information collected directly from the individual or from another source?

☐ Directly from the individual.

☒ Other source(s):

1. DOI Datamart: This feed provides employee data and triggers for activating the employee record, including pay position, grade, salary, clearances, and department
2. US Access: This feed provides Enrollment (ID), SSN, name, and adjudication status information
3. LEAP: This feed provides details on completed training courses
4. Active Directory: This feed provides employee e-mail, and office telephone number
5. Contractor Personnel (CP) List: This feed provides CP start date, name, email, office location, phone number.
6. Archibus: This feed provides employee and contractor workplace location information
7. HR4ME: This feed provides employee telework agreements, and CPA and Bar membership details
8. WebTA: This feed provides employee supervisor's name and organization

6.2 What methods will be used to collect the data?

The data are ingested from various source systems using SEC's enterprise ETL tools. EHCR will be built on the EDW platform and will consist of an Oracle database, a Netezza database with ETL jobs handled by DataStage running on Linux, and Gardium for access/audit control and security monitoring. The ETL jobs are executed daily and kicked off by scheduling scripts. Manual jobs can also be executed.

6.3 How will the data collected from individuals, or derived by the system, be checked for accuracy and completeness?

System authentication, audit capabilities and business rules ensure data integrity. Each inbound data source is checked for accuracy based on already defined business rules. For "data file", the details of the inbound file are verified against control details included in the inbound file. For "web services", the data in the incoming feed are verified against a predefined schema to ensure that the outputs match. The requirements for data quality (accuracy, completeness, timeliness, validity, precision) are the responsibility of the underlying source system as source system updates are replicated in EHCR. These requirements will vary based on the unique business requirements of each system. System replication occurs daily for most sources and weekly for others.

6.4 Does the project or system process, or access, PII in any other SEC system?

☒ Yes.

System(s): The data are ingested from source systems using SEC's enterprise ETL tools for the purpose of aggregating a single view of SEC workers and contractors. Source Systems include: DOI, US Access, Active Directory, LEAP, CP List, Archibus, HR4ME, and WebTA.

6.5 Consider the sources of the data and methods of collection and discuss the privacy risk for this system related to data quality and integrity? How are these risks mitigated?

The primary risk is data accuracy as the ultimate requirement for data quality (accuracy, completeness, timeliness, validity, precision) is the responsibility of the underlying source systems. This risk is mitigated by the ensuring the scheduled updates are replicated in EHCR. The requirements for replication will vary based on the unique business requirements of each system. System replication occurs daily for most sources and weekly for others.

Section 7: Individual Participation

7.1 What opportunities are available for individuals to consent to uses, decline to provide information, or opt out of the project? If no opportunities are available to consent, decline or opt out, please explain.

Any opportunities for individuals to decline to provide information, opt out, or consent to uses of their information occurs at the point of collection via the original source system. EHCR only provides reporting technology using the information ingested from the original source. No new information collection occurs via EHCR.

7.2 What procedures are in place to allow individuals to access their information?

EHCR is not a system of records. The information related to individuals ingested in EHCR may be accessed via procedures outlined in the SORN related to the original source system of records (see Question 4.1). Persons wishing to obtain information on the procedures for gaining access to the contents of records maintained in the original source may contact the FOIA/Privacy Act Officer, Securities and Exchange Commission, 100 F Street, N.E., Washington, D.C. 20549-2736.

7.3 Can individuals amend information about themselves in the system? If so, how?

EHCR is not a system of records. The information related to individuals ingested in EHCR may be accessed via procedures outlined in the SORN related to the original source system of records (see Question 4.1). Persons wishing to obtain information on the procedures for gaining access to the contents of records maintained in the original source may contact the FOIA/Privacy Act Officer, Securities and Exchange Commission, 100 F Street, N.E., Washington, D.C. 20549-2736.

7.4 Discuss the privacy risks related to individual participation and redress? How were these risks mitigated?

EHCR is not a system of records. The reports and analysis features of the system rely on source systems and limit opportunities for participation and redress. Therefore, the primary privacy risks are lack of access to personal information. This risks is mitigated by ensuring that SORNs are current and adequately cover procedures for participation and redress, and that PIA reports are published and adequately describe how personal information will be managed.

Section 8: Security

8.1 Has the system been authorized to process information?

☒ Yes
SA&A Completion Date: 7/20/2015
Date of Authority to Operate (ATO) Expected or Granted: 7/25/2015

8.2 Identify individuals who will have access to the data in the project or system and state their respective roles.

☒ Users
Roles: Access to system data

☒ Contractors
Roles: Access to system data

☒ System Administrators

Roles: Access to system data and configurations. Manage access/permissions, make configuration changes, and perform other administrative actions in support of the system.

8.3 Can the system be accessed outside of a connected SEC network?

☒ No

8.4 How will the system be secured?

Access to EHCR will be strictly governed by the business owners and all requests must be approved by OHR. EHCR will leverage AD role based access. AD will be used to authenticate users and provision the database level of access (read only). OHR will have a report of users and roles to verify for accuracy. OHR will notify the AD administrators to remove users who no longer need access to EHCR.

Underlying network protocols provide some integrity protection as do higher layer cryptographic mechanisms such as Secure Socket Layer (SSL) as implemented within EHCR. Application components are segregated and each segment of the system is located on a different server.

8.5 Does the project or system involve an online collection of personal data?

☒ No

8.6 Does the site have a posted privacy notice?

☒ No

8.7 Does the project or system use web measurement and/or customization technologies?

☒ No

8.8 Describe any privacy risks for this system that relate to the technology and security of the system and how those risks are mitigated.

Handling and managing PII was a major consideration during the design phase of EHCR. The primary privacy risk is failing to limit access to data, or to limit, monitor or enforce access controls, which can lead to misuse or unauthorized disclosure. To mitigate this risk, EHCR was designed without a user interface and access to the data is strictly controlled via security at both the database and Business Intelligence (BI) report levels. Also, this risk is mitigated by ensuring that access to the database is limited to EDW EHCR database administrators. End user interface is only via security-controlled reporting access from BOXI. BOXI reporting will be limited to authorized users and all reports will be limited to defined data sets. Access to PII is controlled through AD groups and only authorized users will have access to that data. EHCR will only be available to designated OHR staff and authorized users.

Section 9: Accountability and Auditing

9.1 Describe what privacy training is provided to users, either general or specific to the system or project.

Users take regular mandatory training on Cyber Security and Privacy Awareness, Protecting Nonpublic Information, and Records Management. EHCR-specific training has not yet been planned.

9.2 Does the system generate reports that contain information on individuals?

☒ Yes

Data is analyzed via search and reporting capabilities, which may present existing information in the form of reports, graphs, charts, and related management metrics. These outputs could potentially contain information on individuals. Any reports will be handled in accordance with applicable records regulations and policies.

9.3 Do contracts for the system include Federal Acquisition Regulation (FAR) and other applicable clauses ensuring adherence to the privacy provisions and practices?

☒ This is not a contractor operated system

9.4 Does the system employ audit logging or event logging?

☒ Yes

The Detailed Level logs exist in the IIS Prod Server. Any failure of the process is notified via email and investigated for root cause. The database side of EHCR is subject to audit processes in place for Oracle and Netezza; not specific to EHCR.

9.5 What auditing measures/controls and technical safeguards are in place to prevent misuse (e.g., unauthorized browsing) of the data? What mechanisms are in place to identify security breaches?

The records are protected from unauthorized access through password and/or PIV card authentication using AD, role-based access, firewalls, and other system-based protections. Access to the underlying data in EHCR is controlled by a DBA group and managed under the SEC procedures governing database access. EHCR data is accessed by authorized business users via a business intelligence reporting tool, BOXI. The BI tool has defined security controls to prevent unauthorized access or misuse of data.

9.6 Given the sensitivity of the PII in the system, manner of use, and established safeguards, describe the expected residual risk related to access.

Although the system maintains sensitive PII, the manner of use and the system controls and user safeguards described in this PIA largely mitigate the risks; expected residual risks related to access is minimal.