NROC/EdReady Data-Collection, Operations, Access, and Security Response Protocols

Updated: June 2021

The NROC College Readiness System, called EdReady, is a cloud-based application that personalizes the pathways to college readiness via diagnostic assessments and referral to appropriate digital resources. The application requires users to create and maintain unique identities across multiple usage periods; therefore, EdReady must store and utilize student data – including personally identifiable data (PII) – as part of its regular operations.

The NROC Project takes data protection and access seriously. EdReady employs best practices in data security, and only the minimum possible amount of PII data is acquired by default. This document details the specific data collected under different implementation scenarios as well as related security, operations, and performance protocols. These protocols are subject to further review at any time. Consult with NROC staff with any questions or comments.

Data Collection

Data collected by EdReady

For Students – Account Information

For initial creation of all student EdReady accts, we collect:

- First name
- Last name
- Unique ID (email or other)

The first name and last name are not validated and can be edited at any time by the user.

For students of our public site (edready.org) and any partner sites that allow students to sign up from the home page, the unique ID must be a valid, unique email address. This address is used to verify the account. The address must be accessible during the sign-up process, but once the EdReady account has been created, it is not technically necessary for that email account to remain active; however, students with inactive (or temporary) email accounts will not be able to reset their passwords, nor will they receive any progress communications from EdReady.

Many partner sites provide access to EdReady via a single-sign-on (SSO) bridge. The specific information passed to EdReady via the SSO bridge varies depending on the SIS, LMS, or directory service in use. In all cases, a unique ID is supplied that is reserved in EdReady for that user (and cannot be changed) and is visible to authorized administrators. EdReady can also store other information (e.g., email addresses, zip codes, other custom fields) passed through by the SSO bridge, but these are not required.

EdReady also stores a default location (as zip code or city/state) for certain location-based recommendations. This information can be changed at any time.

Finally, EdReady provides an option for partner sites to request or require students to provide additional information as part of the sign-up process. This information could include student IDs, age or grade level, school affiliations, reasons for using EdReady, etc. These fields are activated and managed by authorized partner-site administrators, not NROC staff. If these

fields are activated, EdReady will store these data in the same manner as the other protected account information, subject to override by the partner-site administrator.

For Administrators – Account Information

Administrative accounts (of any role) require the same information as for students, except that the supplied email address (as the unique ID) need not be valid, only unique. In addition, EdReady does not collect any location-based information for administrators by default.

Administrative accounts can also be created via SSO, and the same protocols and conditions apply as for students.

There is no administrative access to the public EdReady site (edready.org) except for maintenance and support purposes by NROC staff. Each partner's EdReady site should have an authorized site-version administrator who is then responsible for creating and managing any additional administrative accounts on that site.

For Students – Activity Data

EdReady gathers and stores substantial activity data for each student in the course of using EdReady. These data include: login and usage times and durations, progress data, resource-access data, etc. These data are all securely stored (see below) and provided to authorized administrative users via on-screen or downloaded reports, as well as via a suite of secure APIs.

For Administrators – Activity Data

EdReady gathers and stores basic activity data for all administrators, including session-level login and usage data.

Data collected by 3rd parties (e.g., via cookies, plug-ins, ad networks, web beacons etc.)

There are no data collected by third parties in EdReady, other than anonymous use data for web-based protocols such as data collected by Google Analytics (for high-level views of site activity metrics) and by New Relic (server-management software).

Network Operations Center Management and Security

NROC content, web applications (such as EdReady) and web services are hosted with Amazon Web Services (AWS).

EdReady is a collection of Java/Tomcat/PostgreSQL web applications that makes use of AWS Elastic Beanstalk architecture to provide redundant load-balanced web servers that auto-scale based on server demand. EdReady uses Amazon RDS for database storage and management. EdReady also employs AWS DynamoDB and AWS ElastiCache for log and queue processing.

The NROC Content Server, which includes interventions provided as part of the EdReady resource library, consists of an Amazon Simple Storage Service (S3) data store delivered via AWS Cloudfront CDN for improved durability and reduced network latency.

DNS service for each of the domains we use to host NROC content and services are provided through Amazon Route 53, which also provides integrated monitoring and failover services.

A more detailed description of the service infrastructure and capabilities for each of the services we use can be found here:

AWS EC2:	http://aws.amazon.com/ec2/
AWS RDS:	http://aws.amazon.com/rds/
AWS S3:	http://aws.amazon.com/s3/
AWS Elastic Beanstalk:	http://aws.amazon.com/elasticbeanstalk/
AWS Route 53:	http://aws.amazon.com/route53/
AWS Cloudfront:	http://aws.amazon.com/cloudfront/
AWS ElastiCache	http://aws.amazon.com/elasticache/
AWS DynamoDB	http://aws.amazon.com/dynamodb/

You can find a description of the physical facilities, primary systems, backup systems, and physical security provided by AWS here.

NROC web servers are implemented in a single Region and distributed across multiple Availability Zones (see AWS Documentation for explanation of Regions and Availability Zones). However, we make use of the redundancy provided by AWS by keeping backup instances of our webservers and databases in different Regions and Availability Zones.

Penetration testing, vulnerability management, and intrusion prevention

NROC examines the security architecture and high-level administrative access points to its various platforms and services once or twice a year. All of our online platforms continue to be developed and updated, and security testing is a regular, ongoing process as part of development. Furthermore, our host provides some managed services that protect us from vulnerabilities and intrusions like automated critical patches, scale to prevent DDoS, etc.

Backup schedule, performance, and storage

The EdReady database operates across multiple availability zones. The live database operates in one availability zone while a replicate, which is synched with the live database, is operated in another availability zone. If the live database becomes inoperable, the replica is immediately brought into service.

The EdReady database is backed up on a daily and quarterly basis using the database services provided by Amazon RDS. We rotate the daily backups, keeping the most recent 20 days. Quarterly backups are copied and the duplicate is encrypted and stored off-site (downloaded). In addition to the automated backups provided by AWS, we perform manual (scripted) backups four times a day and copy those backups to an alternate AWS region.

Authorized NROC staff and our contract software developers are the only personnel with secured-access permissions to backups and other service-management tools.

Data Protection and Access

Protecting data “at rest”

Personally identifiable student information (PII), including passwords and sensitive data (i.e., any custom data obtained on account creation), in EdReady is securely stored at rest by default using AES 256-bit encryption. All other student and administrator data are stored unencrypted at rest. We have found that encryption/decryption of all non-PII data results in significant performance degradation and is not justified given the other security measures in place.

As stated above, all data are stored with Amazon RDS. Amazon RDS security is described on page 28 of the AWS Security Whitepaper. In order to take advantage of the security systems and safeguards provided by Amazon Web Services, we employ all of the best practices described on the AWS site.

Protecting data in transit

All EdReady instances run on an HTTPS protocol and are transmitted using the TLS 1.2 standard.

We employ additional encryption (via JWT Tokens) for all API endpoints, including for authentication tokens (for users querying the API endpoints). Returned JSON objects are not encrypted.

Data access permissions

Login access

All EdReady users must create a strong password prior to performing any activities in the platform. This password is BCRYPT encrypted and cannot be retrieved or recreated by any means. If a student forgets the password, the only recourse is to reset the password and create a new one, which can only be done by the user if there is a valid email address on file associated with the account.

For users accessing EdReady via single-sign-on (SSO), EdReady accepts any properly configured authentication token from the student-information system or other account-holding system. The security of this connection is largely dependent on whatever security protocols are in force for users logging into the external application. However, setting up the SSO bridge requires some additional security measures (a unique consumer key and a key secret) in addition to the inherent security of modern, supported SSO protocols (including OAuth and SAML), and all data are encrypted in transit through HTTPS connections.

Authorized personnel

Only a limited number of authorized NROC staff have access to EdReady data, and then only for site management and troubleshooting purposes. All personnel with such access are trained and monitored for proper use of their access privileges.

Each NROC partner is entitled to one or more EdReady instances where access permissions are completely segregated for each site. Access permissions to any one site are granted to one or more authorized personnel designated by the institutional representatives, at least one of which retains permission to create additional administrative accounts and grant appropriate access as required. These decisions are made by the partner institutions who are then responsible for managing those permissions and maintaining user security protocols.

Sub-contracted access

There are no subcontractors outside of our code-development lead who are authorized to access EdReady data on an unlimited basis. The NROC Project partners with certain research organizations, such as SRI, to analyze EdReady activity patterns and outcomes. These research activities require access to certain levels of data for certain select EdReady sites. These research activities are approved via data-sharing agreements between the NROC partner institutions and SRI personnel. Certain non-PII data may be shared after further aggregation and anonymization for research and outreach purposes.

Data and metadata retention

EdReady data are retained as long as there is no specific request for their deletion. If institutional partners require other terms to be met in this regard, we will work with them individually to establish appropriate protocols for data preservation, management, and disposal.

If a partner institution wants to delete all data on their EdReady site, we will execute that order and confirm completion within 72 hours of the request. Data that are deleted in this manner are not retrievable and cannot be restored for any reason.

Incidents and Response

Security notifications and updates

If any security-related updates become necessary, NROC will take immediate steps to secure the existing information and alert all users via existing onboard and external communications channels. Any required code updates will be deployed as soon as practicable, generally within 24 hours. Any further updates that might be required for any integrated third-party systems (e.g., member SISs or LMSs) will be handled as quickly as possible via direct support with the necessary institutional personnel.

We audit current access permissions, site integrity, and external threats on a regular basis. We institute a review of active administrative accounts at least annually, and we evaluate our active security protocols against known updates from the field annually as well.

Availability and recovery

To date, EdReady availability (up-time) has been greater than 99% for all users and regions.

The AWS Elastic Beanstalk architecture provides redundant web servers that are deployed across multiple availability zones. The servers are monitored and replaced or augmented as faults are discovered or demand increases. The AWS RDS service provides a redundant database architecture that provides for a live database in one availability zone, and a synched replica in another availability zone, along with a health monitoring service that can fail over to the replica when a fault is discovered.

Additionally, redundant monitoring systems at AWS and Uptime Robot provide instant alerts about servers that are down or performing poorly. If there is an outage in a particular hosting zone or data center, or there is a server that does not respond to a restart, we can quickly bring up another server in a different location and modify our DNS records to point to the new server. This process typically takes less than five minutes.

Database recovery procedures would typically involve:

1) Diagnosis of the database issue that would be remedied by restoring from a backup.
2) Identification of the best backup file to use.
3) Pausing the application or turning off user logins (to prevent database writes during the restore process).
4) Copying the backup file to replace the corrupt database.
5) Restarting the application or restoring user logins.