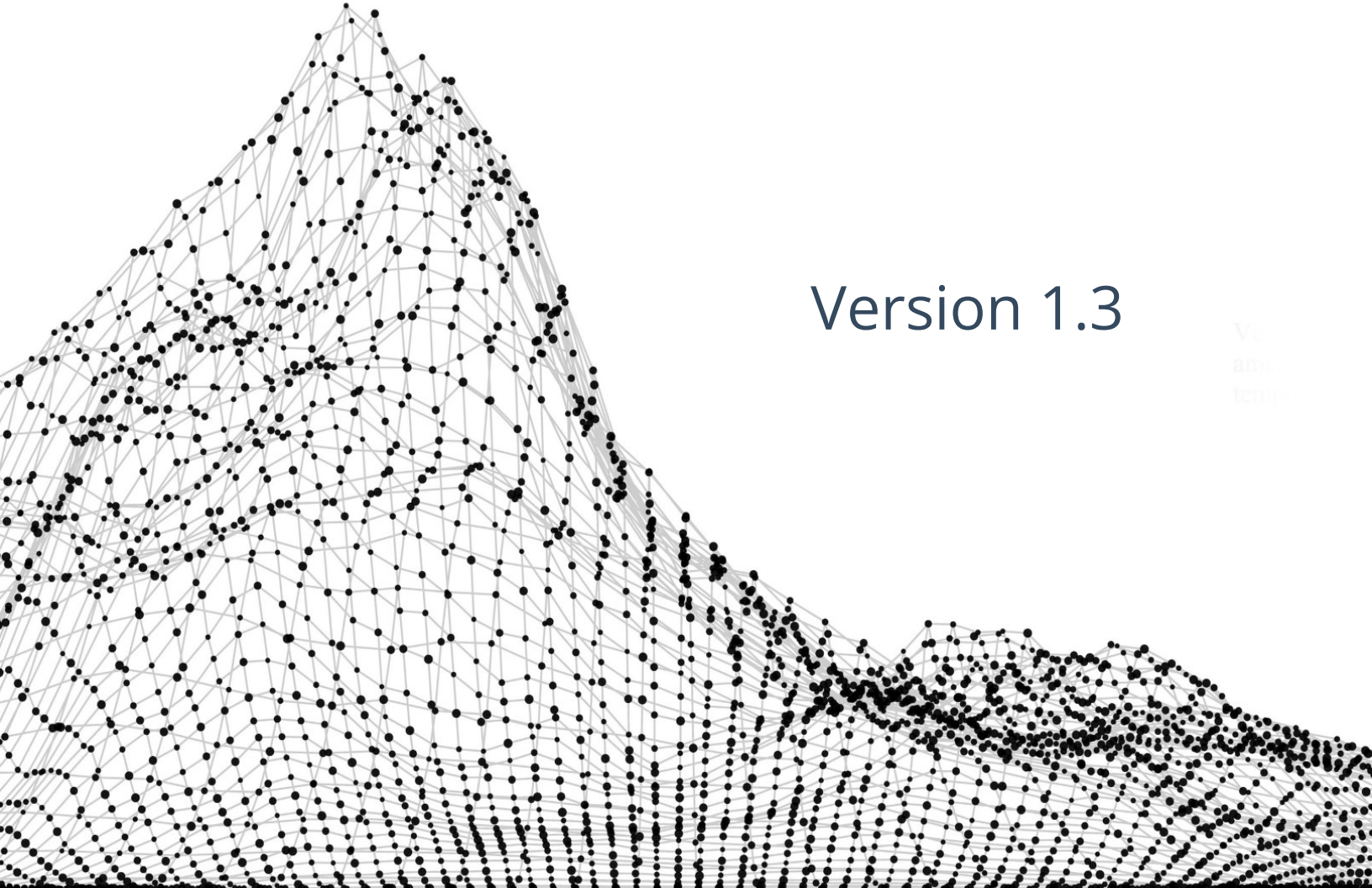


UTAH DATA
RESEARCH CENTER

General Governance

Version 1.3

Utah
Data
Research
Center



Contents

1	Introduction	3
2	Mission and Responsibilities	3
3	Structure	4
4	Research Products	5
5	Review of Research	6
6	Technology	8
7	Data Privacy	9
8	Data Security	10
9	Data Access	11
10	Data Retention and Disposition	13



1 | INTRODUCTION

The Utah Data Governance Center (UDRC) establishes and operates the State of Utah’s program for data-informed and evidence-based research on education and workforce issues and serves as Utah’s state longitudinal data system. UDRC collects longitudinal data from multiple State of Utah data partners, including the Utah System of Higher Education (USHE), the Utah State Board of Education (USBE), the Department of Commerce (UDOC), the Department of Health and Human Services (DHHS), and the Department of Workforce Services (DWS). The data provided from our partner agencies constitute the “P20W” pipeline. P20W captures the behavior of individuals starting in early childhood, continuing through primary and secondary schools and higher education, and lastly as they enter and move about the workforce. UDRC operates data systems that link, de-identify, and warehouses that data for use in research its own research as well as research by data partners, other state agencies, and the general public that addresses questions integral to public policy at all levels in Utah.

UDRC’s research provides objective insights into public policy queries to help decision makers find meaningful and impactful solutions to the issues facing the state of Utah. UDRC significantly decreases the cost to our partners and key stakeholders of data and analyses by leveraging technology resources to create efficiencies in the exchange and release of data between agencies and external requestors. UDRC data research program is committed to the use of the best statistical practices and careful peer review when releasing any research informing public policy.

UDRC’s data governance plan establishes the rules for the collection, storage, access, and use of UDRC data. The data governance plan ensures that UDRC can fulfill its mission and responsibilities, meeting statutory requirements for providing quality research and data for state agencies under Utah Code chapter 53B-33 and effectively fulfilling its role as Utah’s state longitudinal data system while protecting the rights of data subjects. The plan implements the responsibilities delegated to UDRC by the Utah Higher Education Data Privacy and Governance Act of 2022, as amended (HESDP), and is consistent with the student data protections delineated in the Utah Higher Education Data Privacy and Governance Act of 2022, as amended; the Federal Education Rights Privacy Act of 1974 (FERPA); and other applicable state and federal data privacy laws. UDRC maintains reasonable and appropriate data protection methods to ensure that UDRC adequately protects the personal data that it holds; meets its own responsibilities under applicable state, federal, and foreign laws; adheres to industry norms and emerging practices for data privacy and security, and protects its data partners from data breaches and legal violations that might result from UDRC mishandling of partner data. UDRC collects only data that it is legally authorized to collect. Security and privacy practices are complementary, with security practices protecting data both at rest and in motion and privacy practices extending through the full data lifecycle.

2 | MISSION AND RESPONSIBILITIES

2.1 | Mission and Vision

The mission statement of UDRC is:

The Utah Data Research Center conducts meaningful research to inform public policy initiatives, providing the platform for policymakers, practitioners, and the general public to make data-informed decisions.

The vision statement of UDRC is:

To create a world-class State of Utah data and research platform to empower stakeholders.

2.2 | Responsibilities

UDRC’s responsibilities are primarily established in various provisions of Utah Code Chapter 53B-33. Its key statutory responsibilities include:

- Education and Workforce Data. UDRC maintains a longitudinal database of connected, aggregated, and de-identified data from USBE, USHE, DWS, DHHS, and DoC that can support education and workforce research.
- Research Program. UDRC operates a data research program for the purpose of developing public policy for the state based on an annual research agenda using data from its partner agencies.
- Research Requests. UDRC responds to data research requests from its data partners, state government officials of the legislative and executive branches, political and private organizations, and members of the public based on UDRC’s prioritization of research requests.



- Data Visualization. UDRC maintains online data visualizations that allow the public access to aggregated, connected, and de-identified data and creates role-based dashboards that enable querying of data in real-time to allow for customizable views.
- Publications. UDRC publishes the results of its research and releases de-identified datasets for public use through its website.
- Privacy. UDRC protects the privacy of data subjects by ensuring that all data made public through research is de-identified and avoids secondary disclosure.

UDRC is authorized under Utah Code section 53B-33-302(9)(b) to charge reasonable use fees for completing data research requests to assist in meeting center costs.

The UDRC database serves as Utah's state longitudinal data system, a system supported by the United States Department of Education under the Educational Technical Assistance Act of 2002. Such systems are designed to connect individual-level statewide information from early childhood through K-12 education, postsecondary education, and the workforce. State longitudinal data systems are intended to enable cross-sector data insights that help leaders answer pressing policy questions, direct resources, and better support individuals along their education and workforce journeys.

2.3 | Accountability

Utah Code section 53B-33-304 requires UDRC to report to the Education Interim Committee and Business and Labor Interim Committee its research priorities for the year, including ongoing data research items from the previous year and new prioritized research for the starting year, completed research from the previous year, and additional activities and accomplishments of the prior year by July 1 annually. Before December 1 of each year, UDRC reports the current data research priority list to the Education Interim Committee as required in Utah Code section 53B-33-302.

3 | STRUCTURE

3.1 | Organization

UDRC began operations on July 1, 2017 under DWS. Since July 1, 2022, UDRC is part of the USHE Office of Research and Data Science (ORDS). While UDRC is a government entity, UDRC does not speak on behalf of any data partner agency, its leadership, or their governing board.

3.2 | Leadership

The UDRC leadership team is responsible for the day-to-day operations of UDRC, including identifying appropriate policy questions and prioritizing data research requests, developing and operating UDRC's technology solutions, conducting research and compiling reports, providing training and support for UDRC staff, developing and reviewing UDRC's policies and procedures, providing system utilization reports to UDRC's data partners, and ensuring data privacy through compliance with the privacy and security requirements described in sections 5.1.2 and 5.2.1 below.

UDRC's leadership team consists of three staff members:

- The Associate Commissioner for Workforce and Institutional Research is appointed as the director of the center by the Commissioner of Higher Education under Utah Code section 53B-33-203 as part of their wider responsibilities within USHE. The Associate Commissioner oversees the attainment of UDRC's mission and the pursuit of its vision, is entrusted with budgetary oversight of UDRC, and oversees the performance of the Senior Assistant Commissioner and provides guidance when required.
- The Senior Assistant Commissioner of Research and Data Systems is immediately responsible for the leadership of the center, directing UDRC staff in attaining UDRC's mission and pursuing its vision. The Senior Assistant Commissioner communicates the operational directives of the Associate Commissioner and is expected to make executive decisions for the purpose of furthering those operational directives. The Senior Assistant Commissioner establishes the operational infrastructure of UDRC, provides leadership and ultimate decision-making in the creation of UDRC's technology solutions, is entrusted with operational budget decisions, and oversees the performance of the Assistant Commissioner providing guidance when required. The Senior Assistant Commissioner and the Associate Commissioner are the primary points of contact for internal and external entities and individuals in regard to communication, with some communication responsibilities delegated to the Assistant Commissioner.
- The Assistant Commissioner of Research is tasked with the day-to-day management of UDRC and with



executing the vision and mission promulgated by the directives of the Associate Commissioner and Senior Assistant Commissioner. The Assistant Commissioner provides project management for technology solutions from the business side, leads the annual research agenda, and manages the conduct of research

3.3 | Advisory Board

UDRC is supported by the Utah Data Research Advisory Board, which is chaired by the Commissioner of Higher Education and composed of representatives of its data partners. The advisory board is established under Utah Code section 53B-33-202, and consults in the development of the UDRC research agenda and approval of research requests. The management team provides updates on UDRC to the advisory board. The advisory board offers advice and guidance when necessary.

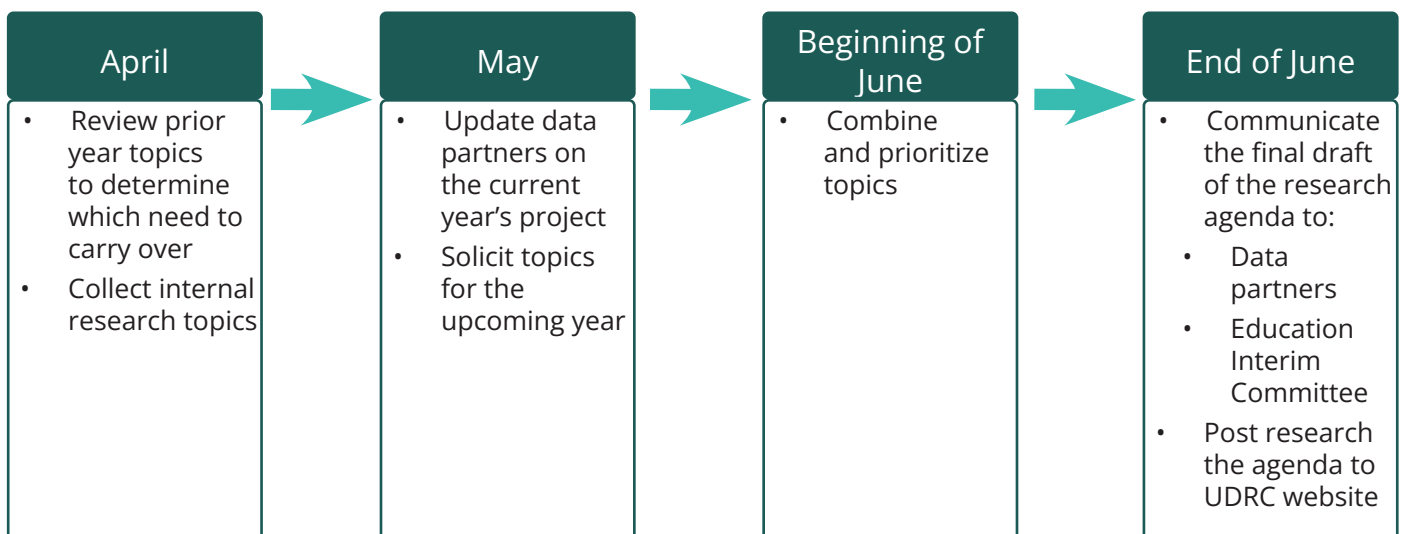
4 | RESEARCH PRODUCTS

The UDRC data research program produces “data products” and “data as a product.” UDRC data products help communicate complex ideas in easily understood formats to aide interested parties in making data-informed decisions. Data as a product is a simpler and more efficient way to access the P20W data, helping ensure access only by approved external researchers to conduct and provide answers to research within, and potentially extending beyond, the P20W domain. UDRC delivers four main products: research reports, analysis reports, visualizations, and external-use data.

4.1 | Research

The primary data research output from UDRC is research reports. The topic of each research report is an item from the yearly Primary Research Agenda (see section 4.2.1 below). The research reports are longitudinal, following outcomes across time and data agency. The results, communicated as written documents and available as interactive data visualizations available on UDRC website, are designed to contribute to data-informed, evidence-based public policy.

UDRC’s research agenda determines the primary questions driving UDRC research and prioritizes the research topics for each fiscal year. The research agenda process adheres to the requirements outlined in Utah Code section 53B-33-302. Requests to include a project in the research agenda are, by statute, limited to legislative committees or staff offices, the executive branch, data partners, or members of the advisory board. The process of formulating the research agenda ensures that formally submitted research requests are adequately formulated, appropriately prioritized, and, as possible, addressed during each program year. UDRC research reports are based on the annual research agenda. The research agenda is communicated to the Interim Education Committee before July 1 of every year.



UDRC, in consultation with the advisory board, ensures timely delivery of the research agenda through the following process:

1. In April of each year, UDRC begins collecting potential research topics. UDRC also reviews the prior year's items to identify items that will be continued to the next year's agenda if needed.
2. UDRC contacts data partners in May to communicate progress on the current year's research agenda, propose topics for the next year's agenda, and request agency needs for topics and priorities.
3. By June, UDRC combines agency requests and prioritize requests. UDRC may add at least one topic from each data agency to the research agenda. An agency may have multiple topics included in the research agenda if time and space permit.
4. Additional topics that were proposed but not included on the research agenda will be added to a research backlog. These topics may be assigned a lower priority for the requesting agencies or identified as projects for which UDRC does not currently have the data but the data may become available in the future.
5. UDRC will communicate a prioritized research agenda draft to data partners for review.
6. Before the end of June, UDRC will create a final draft of the research agenda. UDRC will communicate a final draft of the research agenda to the Education Interm Committee.
7. The final draft is publicly available on UDRC website.

4.2 | Analytic Requests

UDRC provides analysis of education and workforce data for state agencies and the public. Organizations and individuals other than those who may submit requests to the primary research agenda such as state government entities, political subdivisions of the state, private entities, or members of the public may request data research as UDRC resources become available. Requests for analysis are submitted through UDRC website. Analytic requests are prioritized in UDRC's secondary research agenda by the UDRC leadership team based on the source of the data request, the visibility of the final report, the urgency of the report, the expected impact of the final report and the availability of the data. The topics are longitudinal, addressing questions that involve data linked across time and agency. The results are relevant to developing policy throughout the state. Analysis reports are housed on UDRC website.

4.3 | Data Visualizations

UDRC maintains online data visualizations that are available to the public. These data visualizations allow public access to aggregated, connected, and de-identified data. UDRC also creates role dashboards that enable querying of data in real-time to allow for customizable views. Data visualizations are published through the online data visualization portal on the UDRC website.

4.4 | Data Requests

UDRC provides appropriately limited access to data collected from its data partners to other parties. Researchers may request linked, aggregated, and de-identified datasets developed as part of the UDRC data research program for purposes of advancing public policy or academic research subject to UDRC's prioritization process as permitted in Utah Code section 53B-33-302. Datasets may be requested by data partners, executive branch agencies, legislative staff, additional government agencies, or those with legitimate research use for the data. Datasets are transmitted to researchers securely and must, under Utah Code subsection 53B-33-302(10), be connected, aggregated, and de-identified. Datasets are also obfuscated as necessary to prevent secondary disclosure.

Dataset requests are submitted through UDRC website. For UDRC to fulfill a dataset request, the dataset must use linked data from at least two data partners. If a data request is appropriate for UDRC, a representative of UDRC will meet with the requester to further assess if UDRC is able to fulfill the dataset request. If approved, UDRC will notify data partners whose data is part of the dataset request. An IRB must approve requests for research datasets outside government entities. All dataset requesters must adhere to UDRC's Investigator Assurance Policy, Data Disclosure Agreement (NDA and data destruction), and Peer Review Process.

As authorized by Utah Code section 53B-33-302, UDRC may charge a fee to cover the costs of preparing a dataset request.



5 | REVIEW OF RESEARCH

5.1 | Peer Review

All UDRC products undergo formal peer review. The peer review process is designed to ensure:

- the protection of all PII including avoiding secondary disclosure in aggregate reporting,
- the use of appropriate methodologies and accurate statistical reporting,
- high-quality and informative writing,
- that conclusions follow from the reported evidence and avoid overly subjective interpretation of the results,
- that data partners' data are accurately represented, and
- that data partners have time to formulate any responses to findings if necessary.

UDRC peer review ensures accuracy at each step of the data research process. The peer review process is not intended or designed to ensure that researchers' findings conform with the views of any state agencies or officials. The staff of UDRC will put forth significant effort to preserve the center's role in supporting all research, regardless of the implications of any research findings.

The two primary components of UDRC peer review process are internal review and external review. An internal peer review is performed at each step of the data research process for research, analysis, and data requests. UDRC research scientists act as peer reviewers for each other's data queries and analysis. The research scientists are subject matter experts in how disparate data partners' data work together, proper statistical coding, and reviewers for secondary disclosure issues. Along with a research scientist, the members of the UDRC leadership team review draft reports and provide feedback. Researcher scientists incorporate feedback prior to communicating the draft for external review.

For research and analysis, the data partners provide external review. The data partners are subject matter experts on the submitted data used for any research or analysis. Data partners' primary peer review responsibilities are to ensure the data research product correctly uses the partners' data, including but not limited to evaluating whether the data were capable of addressing the research question the data were used for and, when appropriate, whether the data definitions in the resulting research accurately reflect those of the data agency. Of secondary importance, the data partners help to ensure that the document is broadly accurate, that the findings are statistically and methodologically sound, and that interpretation and discussion of results are supportable by the data provided.

When UDRC provides data, UDRC will act as an external peer reviewer. The external researcher will provide UDRC with a draft of the research document. UDRC will provide statistical and methodological feedback. UDRC will communicate the document to the relevant data partners, who will act as subject matter experts, ensuring accurate definitions and descriptions of their data. The external researchers will incorporate the feedback from both UDRC and the data partners. Additionally, the external research will include the following disclaimer:

Data for this research was accessible through Utah's state longitudinal data system database administered by the Utah Data Research Center which includes data supplied by UDRC member. This research, including the methods, results, and conclusions neither necessarily reflect the views of, nor are endorsed by, UDRC members. All errors are the responsibility of the author.

Data agency research will not fall under the peer review process. The data partners employ their own internal peer review. Additional information for the specific steps in the peer review process are outlined in UDRC peer review document available at udrc.ush.edu.

5.2 | Institutional Review Board Approval for Scholarly Research

To ensure the protection of human subjects, research conducted for scholarly research and activity purposes must be reviewed by a certified institutional review board or the Utah State Board of Education Institutional Review Board (USBE IRB). UDRC will not provide data to the researchers unless:

- the IRB has determined that the research does not require IRB review because it is not within the scope of the "Common Rule," 45 CFR section 46.101 (for example, if the work is not research or does not involve human subjects as defined in the Common Rule);
- the IRB has determined that the research is exempt from the Common Rule under the provisions of 45 CFR section 46.104; or
- the IRB has approved a human subjects research protocol for the research.



A determination by the researcher personally that the research is not within the scope of or is exempt from the Common Rule's requirements is not acceptable. UDRC will not make its own determinations regarding IRB applicability on behalf of a scholarly researcher.

Scholarly researchers are strongly encouraged to consult with UDRC before submitting an IRB request to any institution, paying particular attention to privacy and security controls in the protocol. An approved IRB protocol may not be amended without the approval of the IRB. IRB approval does not constitute or authorize any exception or exemption from UDRC policy or state or federal law. If UDRC cannot provide data as described in an approved IRB protocol without violating law or policy, or the approved protocol does not provide sufficient guidance for UDRC to determine whether it can comply with the protocol with law or policy, the researcher will need to amend the IRB protocol and receive IRB approval for the amended protocol before UDRC can provide the requested data. UDRC will consider an approved IRB that includes all required provisions for a written agreement to disclose data to constitute such an agreement.

Data accessed for the purposes of internal UDRC research, as part of UDRC research agenda, or by data partner or external researchers for program and policy research is intended to inform the development of public policies to improve a Utah citizen's ability to access services at an institution of higher education or graduate with a postsecondary certificate or degree and is presumed to not develop or contribute to generalizable knowledge. Such research is presumed to not constitute research within the meaning of the Common Rule, 45 CFR section 46.102(l) and thus to not require IRB approval. This presumption is subject to determination to the contrary by UDRC personnel, who may require an IRB determination as described for scholarly research.

6 | TECHNOLOGY

6.1 | Platform

The UDRC is built in the Google Cloud Platform (GCP), a premier provider of cloud based services. Specific services of the platform are employed to create the UDRC infrastructure, micro-services, processes, procedures, and associated applications; which work in concert to produce the objectives of the UDRC data system. GCP's native tools and services, like BigQuery and Cloud Storage, are intricately woven into the UDRC's architecture, providing seamless data management and analytics capabilities. The incorporation of Vertex AI into the UDRC framework empowers researchers to harness advanced machine learning and AI capabilities, facilitating innovative research methodologies and data analysis. With Vertex AI, the UDRC offers a versatile platform for researchers, accommodating various programming languages like Python and R, which are widely used in data analysis and scientific computing. The UDRC system is protected by the information system controls required to align with the Federal Risk and Authorization Management Program (FedRAMP) Moderate security controls baseline. The FedRAMP Moderate baseline, published by the U.S. Federal government, provides a comprehensive list of cybersecurity controls required to manage, process and store sensitive data in cloud environments. The comprehensive cybersecurity controls of the FedRAMP framework enable the UDRC to conduct regular security audits and risk assessments, ensuring continuous improvement in its defense mechanisms.

6.2 | Objectives

To ensure the UDRC is able to effectively address all duties assigned to the UDRC in Utah Code chapter 53B-33, the UDRC designed and, on an ongoing basis, evaluates its technology solutions to ensure the following:

- **Security.** Encryption, obfuscation, suppression, and access management are integrated. Thus, risk is mitigated to an acceptable level, given the sensitivity of the data.
- **Efficiency.** Automation of data extraction and a common extraction protocol that decreases the resource (capital and labor) intensity required to accomplish the foundational task of retrieving data. In addition, the consolidation of data increases access and reduces the effort for individuals seeking information.
- **Quality.** Data governance, automated protocols, and IT structures are combined to produce longitudinal data with a very high level of consistency, accuracy, and of credible research value.
- **Efficacy.** Automation, quality systems, and data access points contribute to the decreased time for research production and additional data products.



6.3 | Master Person Index Process

Linking individuals across disparate data sets is a core competency of UDRC and a key function of its data processes. Matching across agency and time presents challenges, namely noisy data in the form of differing variables across agencies, changing PII over time, and potential errors in data entry. To address the uncertainty created by noisy data, UDRC has developed a probabilistic matching system, the Master Person Index (MPI) system. The MPI system is a multi-step technical and statistical solution that handles probabilistic matching over data sets of substantial size. The MPI process indexes potential existing identities to newly added data using both deterministic and non-deterministic steps, compares the indexed row to the potential MPIs through a series of BigQuery queries based on multiple identifiers, and classified using a probability match for each potential MPI attached to the new individuals. The potential MPI with the highest probability match is assigned to the individual if it is above the probability threshold; if there is no MPI above the threshold the record is considered a new individual and a new MPI is created with that information. Given the uncertainty introduced by noise in the matching process, this probabilistic process reduces potential lost matches, increasing the power of research and potentially reducing biases in results.

6.4 | Metadata

The UDRC data dictionary provides high-level metadata on the various data elements available for request. The dictionary is organized by data partner agency. Details include field values, years available, limitations, and other notes. Variables available include those from kindergarten to 12th grade, postsecondary education, health, and employment in Utah. Variables with 10 or fewer records are never disclosed to protect personal information. The data are updated on a quarterly basis. The dictionary is published on the UDRC website.

7 | DATA PRIVACY

7.1 | Data Collection Authority

UDRC collects data only as authorized by law. Utah Code sections 53B-1-109(2) and 53E-4-308(3) through (5) require USBE and USHE to coordinate to allow individual student academic achievement to be tracked through both the public and higher education systems, including coordinating access to the unique student identifier of a public education student who later attends an institution within USHE. UDRC is designated as USBE and USHE's agent for executing these requirements in their MOU of June 29, 2022. Under Utah Code section 53B-33-301, UDRC is authorized to collect data from its data partners for purposes permitted by statute. Data managed by UDRC must be aggregated, connected, and de-identified upon collection.

While state law authorizes UDRC to collect data from its data partners, data partners who are educational institutions or agencies nonetheless may only disclose personally identifiable data from a student's education records to third-parties with the consent of the parent or eligible student (a student who is over the age of 18 or attending a postsecondary institution) or under one of the exceptions to consent described in FERPA, 34 C.F.R. section 99.31(a). UDRC's data research program serves as Utah's state longitudinal data system, which operates under the authority of USDE to enable state educational agencies to design, develop, and implement statewide, longitudinal data systems to efficiently and accurately manage, analyze, disaggregate, and use individual student data, which may be used to facilitate research to improve student academic achievement and close achievement gaps. UDRC receives personally identifiable information from its education data partners under the audit and evaluation exception, 34 C.F.R. section 99.31(a)(4), in order to facilitate future evaluations of federal-or state-supported education programs by establishing linked data, thereby building capacity to regularly evaluate federal-and state-supported education programs. USBE designated UDRC as their authorized representative for the purpose of auditing or evaluating educational programs on their behalf in a memorandum of understanding of June 29, 2022.

Unemployment information may, under 20 C.F.R section 603.5(e)(2)(iii), be disclosed to UDRC as a contractor of public officials as required for education or workforce training program performance accountability and reporting under federal or state law.

7.2 | PII Protection Requirements

UDRC is not directly governed by FERPA, but it is obligated to uphold its educational data partners' FERPA obligations in order for the educational data partners to disclose data to UDRC. FERPA section 99.33(a)(1) thus prohibits UDRC from disclosing of personally identifiable information from education records without the consent of the student except as allowed under section 99.31(a) just as it does for educational agencies and



institutions. In some instances, data may also be protected by the Parts B and C of the federal Individuals with Disabilities Education Act (IDEA), 34 C.F.R. sections 300, 302. IDEA incorporates all the provisions of FERPA and adds additional requirements to safeguard privacy.

Data provided to UDRC is further protected by the Utah Student Data Protections Act, Utah Code Title 53E, Chapter 9, for primary and secondary education students and the Utah Higher Education Student Data Protections Act, Utah Code Title 53B, Chapter 28, Part 5, for students at postsecondary institutions. UDRC is a third-party contractor of USBE and UBHE, and is thus required to use personally identifiable information from education records strictly for the purpose of providing the products as described in Utah Code 53B-28-505(1) and data sharing agreements with its data partners.

Workforce-related data are protected and secured by federal and state law. States are required to provide for maintaining the confidentiality of any unemployment compensation information which reveals the name or any identifying particular about any individual or any past or present employer or employing unit, or which could foreseeably be combined with other publicly available information to reveal any such particulars, and must bar the disclosure of any such information, except as provided in 20 C.F.R Part 603. These requirements are put into effect under Utah Administrative Code section 994-312-103. The Workforce Innovation and Opportunity Act of 2014 exempts from disclosure information collected under the auspices of the workforce development system that would “constitute a clearly unwarranted invasion of personal privacy,” 29 U.S.C. section 3245(a) (4)(B)(i). UDRC thus cannot release or share workforce information about individuals that would constitute an unwarranted invasion of privacy, even to advance its workforce development system.

7.3 | De-Identification Protocol

UDRC applies a de-identification protocol to substitute a UDRC-assigned individual identifier for personally identifiable information provided by the data partners. Data tables from the agencies are copied into UDRC Research-Use database. A matching algorithm matches student identities in the input files to existing records in the Master Person Index (MPI) table and either associates an existing MPI with the input records or assigns a new MPI to students who are not matched. Primary PII received from the agencies are removed or transformed sufficiently to remove the identifiable component of the data while maintaining the research value of the variable. The primary unique identifier used in working data files is UDRCMPI. Only authorized UDRC staff are allowed access to the PII for the purpose of matching and linking records together and validating the accuracy of the input and matching process.

The details of these procedures are not included in this plan to prevent the disclosure of information that could be used to compromise the security or integrity of UDRC’s data systems.

7.4 | Secondary Disclosure Mitigation and Disclosure Avoidance Protocol

UDRC adopts robust safeguards to reduce the risk of secondary disclosure (such as applying statistical methods to protect PII in aggregate data tables. These safeguards, often referred to as disclosure avoidance methods, can take many forms (e.g., data suppression, rounding, recoding, obfuscating, etc.). UDRC adopts best practices strategies of disclosure avoidance as recommended by the Privacy Technical Assistance Center (PTAC). All data must be reported in a way that individuals cannot be identified based on the information available, meaning that a reasonable person in the community, who does not have personal knowledge of relevant circumstances, cannot identify an individual person with reasonable certainty. To accomplish this, UDRC requires that, at a minimum, results of less than 10 individuals ($n < 10$) in a category or subgroup be suppressed. Other appropriate techniques will be applied based on the specific needs of the project to further protect student privacy.

The details of these procedures are not included in this plan to prevent the disclosure of information that could be used to compromise the security or integrity of UDRC’s data systems.

8 | DATA SECURITY

8.1 | Security Requirements

FERPA does not specify technical data security requirements. Under FERPA, UDRC’s education data partners are required to use reasonable methods to ensure to the greatest extent practicable that UDRC, as their authorized representative, uses personally identifiable information only for the purposes identified in Utah Code section 53B-33-301, destroys personally identifiable information when no longer needed for those purposes, protects the education partners’ personally identifiable information from further disclosure, and protected in a manner



that does not permit personal identification of individuals.

Data protection standards for contractors under HESDP section 501(4)(a) require that an education entity “incorporates reasonable data industry best practices to maintain and protect student data and other education-related data” and “describes the data breach response processsection .”

The data partner memoranda of understanding (MOUs) require that UDRC use appropriate security controls in accordance with the most recent NIST 800-53 guidelines and State of Utah Enterprise Information Security Policy and security standards. USHE Policy R345 implements the Center for Internet Security (CIS) Critical Security Controls as the security standard for USHE data systems. The CIS controls largely map to the NIST standards and use methods comparable to the state standards, and are understood by USHE and USBE to comply with the MOUs.

The data partner MOUs require that data is physically secure from unauthorized access and not accessible from the internet or other remote means. UDRC must notify USBE of actual or suspected breaches. UDRC must provide access for inspections and maintain systems to support audit requirements.

8.2 | Security Infrastructure and Conrols

UDRC data is stored through a secure cloud storage provider, details of which are not included in this document for security reasons. Administrative access to cloud storage is limited to authorized UDRC staff. UDRC defers to the policies of DWS and the Department of Technology Services (DTS) regarding data security, and complies with FedRAMP standards for medium impact level systems. The data partner MOUs require that UDRC provide privacy training regarding the data shared under it and potential consequences of unauthorized sharing.

8.3 | Incident Response

UDRC incident response plan manages potential cybersecurity incidents. This plan is compliant with UDRC’s fundamental responsibilities as the State of Utah’s longitudinal data system. It was created in coordination, and consultation, with the University of Utah’s Information Technology (UIT) group. UDRC is covered by the umbrella cybersecurity insurance plan held by the University of Utah and subject to its requirements. UDRC is supported by the UIT’s Incident Response Team (IRT), which provides liaison services between UDRC and the cybersecurity insurance personnel.

The details of these procedures are not included in this plan to prevent the disclosure of information that could be used to compromise the security or integrity of UDRC’s data systems.

9 | DATA ACCESS

9.1 | Data Control and Access Requirements

UDRC is charged with transactional provision of de-identified data beyond its data partners subject to the provisions of Utah Code Title 53B, Chapter 33. While UDRC has wide scope to conduct research for government and the public and is obligated to make completed research available to the public, it is obligated to ensure that any data contained in a completed data research request is de-identified. Research requests are governed by section 302. Section 302(8) requires data released in a completed data research request is de-identified. Section 302(10) allows, but does not require, release of datasets to a participating entity, executive branch agency, or legislative staff office if the datasets are connected, aggregated, and de-identified. Section 303(1) requires UDRC to make available connected, aggregate, and de-identified data to the public in a data visualization portal.

Release of student data by UDRC is governed by FERPA based on its origin with educational institutions. As an authorized representative of a state or local education authority under section 99.31(a)(4), UDRC may provide information from students’ education records to users other than the education agency or institution that provided them only as permitted under the FERPA audit and evaluation exception described in section 99.35. UDRC must protect the data in a manner that does not permit personal identification of individuals by anyone other than its data partners. UDRC may provide data on behalf of USBE or USHE institutions, in which case it must record further those disclosures itself if USBE or a USHE institution does not do so. UDRC may also provide information where its collection is authorized by Federal law. In accordance with FERPA, UDRC does not redisclose personally identifiable information.

UDRC is not authorized to release personally identifiable information in any form to any parties other than its data partners. The USHE-USBE and USHE-DWS MOUs provide that UDRC will not release personally identifiable information to the public or publish such information, and will not release information if there is a reasonable



possibility of identification whether directly or by inference. UDRC must ensure data provided to third-party researchers is de-identified and not linkable to personal identifying information, and will use a written agreement to ensure third-party researchers comply with the confidentiality terms of the MOUs. The MOUs permit UDRC to release de-identified data with a matching code in accordance with FERPA section 99.31(b)(2).

UDRC is required to release data to the legislative auditor under Utah Code section 53B-33-301(9).

Implementation of this provision, however, is constrained to aggregate data, de-identified data, and limited disclosure of personally identifiable information under FERPA as described above. Under 34 CFR section 99.35(a) (1), personally identifiable information may only be disclosed to the legislative auditor in connection with an audit or evaluation of Federal or State supported education programs, or for the enforcement of or compliance with Federal legal requirements that relate to those programs. The legislative auditor must be designated as an authorized representative of the Commissioner of Higher Education or the State Superintendent of Public Instruction by written agreement meeting the requirements of 34 CFR section 99.35(a)(3). UDRC data is not subject to release under GRAMA as described in Utah Code section 53B-33-301(8).

9.2 | User Access

UDRC user access procedures are intended to limit disclosure of and access to partner data in accordance with industry-standard data protections, the requirements of all federal and state laws applicable to UDRC data partners, and DWS data privacy and security policies including training on the handling of sensitive data apply to UDRC. UDRC provides access to data that integrates the resources of its data partners. If a project requires data from only one data contributor, the researcher must contact the individual state agency and will not be granted access to the data through UDRC.

UDRC data warehouse is created and used in accordance with state and federal legal authorities that govern information exchange, access, and confidentiality of individual records. Storing data in UDRC data warehouse does not create new rights of access to the data loaded in UDRC data warehouse for UDRC members or any other entity. UDRC data warehouse is a tool for storing data that is intended to enable research for the purpose of informing public policy. Violations of this policy could result in loss of access to UDRC data warehouse and criminal prosecution pursuant to applicable state law.

Data access procedures are based on the type of user requesting access to the data and the purposes of the data request. UDRC recognizes the following types of users:

- Authorized UDRC staff: UDRC employees who are specifically designated as having access to personally identifiable data.
- UDRC researchers: UDRC employees assigned to work on projects led by UDRC.
- Data partner researchers: Individuals who are requesting data for program and policy research or scholarly research and activities and who are affiliated with a partner that is currently providing data to UDRC, including individuals or groups conducting research under a contract with or as a consultant for one of the data partners and ancillary research groups associated with a data partner (for example, the Kem C. Gardner institute when conducting research for USHE).
- External researchers: Individuals and groups who are requesting data for program and policy research or scholarly research and activities and who are affiliated with the State of Utah (including agencies and political subdivisions) but are not affiliated with UDRC or its data partners, or who are affiliated with an organization with an interest in formal policy research.
- Public: All other individuals or groups.

Users may request data for purposes of:

- Program and policy research: Data that will be used to draw conclusions about government operations such as program needs evaluation, audit and evaluation of programs or organizations, the development and evaluation of public policy, or required reporting.
- Scholarly research and activities: Data that will support systematic investigation designed to develop or contribute to generalizable knowledge or the development of academic literatures, including student projects that are part of academic programs.
- Informational requests: Data intended to generally inform the public without a specific research purpose.

9.3 | Access Controls



UDRC uses a wide range of data access procedures to ensure effective user access within appropriate limitations required by law. Data protection procedures include:

- Archived when loaded: Files from data partners are used to load records into UDRC data system but are they stored in a separate data archive with limited access after loading. Archives are used only to verify system integrity.
- GCP IAMS access controls: Data are kept secure at all times and authorization for access is controlled through the operational and standard GCP IAMS controls.
- Separation of Master Person Index: Direct personal identifiers are stored in a Master Person Index table separately from the working data.
- De-identification: Primary PII is replaced with a UDRC unique ID and generalized using identity-masking demographic information.
- Data request required: External researchers must request data through UDRC data request process and are not permitted to access the database directly.
- Project-specific dataset design: External researchers work with the data engineers to understand the metadata of a given dataset, be informed of the data available, and for the creation of custom datasets for their particular research question and are not permitted to access the database directly.
- NDA required: External researchers must sign a non-disclosure agreement to promote the security and privacy of data to which they have been granted access.
- Linkable PII mitigation: Processes using USDE-promoted disclosure avoidance methodologies to reduce the possibility of linking de-identified data to specific students or employers.
- Aggregation: Data are reported as statistics about groups of students or employers unless necessary for the purpose of context or mandated by legislative directives.

The details of these procedures are not included in this plan to prevent the disclosure of information that could be used to compromise the security or integrity of UDRC’s data systems.

9.4 | Data Categorization and Control

UDRC applies access, security, and privacy controls according to a four-category schema: (1) Restricted-Use Data, (2) UDRC Research-Use Data, (3) External Research-Use Data, and (4) Public-Use Data. In addition, level (0) Partner Submission Data is used to create UDRC level (1) data files but is not stored in the active UDRC database.

Level	Type	Description	Security Procedures	Access and Purposes	FERPA Status
0	Partner Submission Data (Raw Data)	Source data files submitted by data partners for UDRC processing	<ul style="list-style-type: none"> • Archived when data loaded to active UDRC database • GCP IAMS access controls 	Authorized UDRC staff for data system management	Protected
1	Restricted-Use Data	Data that includes information about the identity of individuals and employers	<ul style="list-style-type: none"> • GCP IAMS access controls • Separation of Master Person Index 	Authorized UDRC staff for data system management	Protected



Level	Type	Description	Security Procedures	Access and Purposes	FERPA Status
2	UDRC Research-Use Data	Unit-level (individual students or employers) records that contain no primary individually identifiable information but include limited indirect identifiers	<ul style="list-style-type: none"> De-identification Data request required for external researchers GCP IAMS access controls 	UDRC and data partner researchers for program and policy research	Presumed to be protected unless determined otherwise in specific cases
3	External Research-Use Data	Unit-level (individual students or employers) records that contain no primary individually identifiable information and are evaluated for secondary disclosure	<ul style="list-style-type: none"> Project-specific dataset design NDA required Linkable PII mitigation De-identification Data request required for external researchers 	External researchers for program and policy research or scholarly research and activities	Presumed to be unprotected unless determined otherwise in specific cases
4	Public-Use Data	Aggregated from level 2 data and contain no unit-level record data.	Aggregation above level of individual agencies	Public for information requests	Not protected

10 | DATA RETENTION AND DISPOSITION

10.1 | Data Disposition Requirements

UDRC is required under FERPA to destroy personally identifiable information from education records disclosed by its educational data partners when the information is no longer needed for the purposes for which it was disclosed. The data partner MOUs require UDRC to destroy personally identifiable information within 180 days after the purpose for which the information is disclosed is served, but states that the purpose will continue as long as UDRC research program continues. This is consistent with UDRC's role as Utah's state longitudinal data system as articulated by USDE to enable state educational agencies to design, develop, and implement statewide, longitudinal data systems to efficiently and accurately manage, analyze, disaggregate, and use individual student data.

10.2 | Data Retention Schedule

UDRC retains data based on the following schedule.

Level	Data Type/Event	Archive Timeframe	Disposition Timeframe	NIST Disposition Standard
1	Partner Submission Data (Raw Data)	When added to UDRC database	Never, as long as UDRC program is in operation	Purge
2	Restricted-Use Data	Variable	Variable	Purge
3	Deidentified Working Database Files	Never	Never	None



Level	Data Type/Event	Archive Timeframe	Disposition Timeframe	NIST Disposition Standard
4	UDRC Research Project Files (Research-Use Data)	On completion of project	Never	None
5	External Research Datasets	TBA	As required by UDRC's NDA associated with the external data request process	Clear (Researcher) TBA (UDRC)
6	Public-Use Datasets	When deleted	Variable Dependent on the purpose of the data and the utilization of said data	Delete (UDRC) None (Public)
7	Dissolution of UDRC	Upon dissolution	Upon dissolution	Destroy

Data should be disposed of to the specified standards of NIST SP 800-88 Rev. 1, "Guidelines for Media Sanitization."

- "Delete" uses the operating system's Delete commands to designate a file as logically deleted within the operating system. Note that "Delete" is not considered a sanitization standard under NIST SP 800-88 as deleted files are easily restored using simple non-invasive data recovery techniques.
- "Clear" applies logical techniques to sanitize data in all user-addressable storage locations for protection against simple non-invasive data recovery techniques; typically applied through the standard read and write commands to the storage device, such as by rewriting with a new value or using a menu option to reset the device to the factory state (where rewriting is not supported).
- "Purge" applies physical or logical techniques that render target data recovery infeasible using state of the art laboratory techniques.
- "Destroy" renders target data recovery infeasible using state of the art laboratory techniques and results in the subsequent inability to use the media for storage of data.

UDRC's cloud storage provider uses deletion protocols including cryptographic erase that generally meet the NIST standard for purged data within 180 days of deletion and destroyed data within two years of deletion.

