

## 1 Introduction

### 1.1 Rationale

1.1.1 The North Dakota University System Institutional Research (NDUS-IR) department serves the North Dakota University System Office (NDUSO), the State Board of Higher Education (SBHE), and various other stakeholders by providing relevant aggregated reporting of North Dakota University System (NDUS) student and employee data.

1.1.2 Additionally, NDUS-IR conducts analyses of individual level data to respond to stakeholder questions. Individual level data contains personally identifiable information (PII).

1.1.3 In some cases, questions arise that ask NDUS-IR to use individual level data to determine past outcomes, predict future outcomes or identify intervention targets. These are respectively referred to as descriptive analytics, predictive analytics, and prescriptive analytics.

1.1.4 There is a need to establish guiding principles to provide a clear framework for the ethical application of analytics within the NDUS-IR department.

### 1.2 **Problem statement: Access to both individual level data and analytics tools provide new opportunities for data analyses in NDUS-IR. As a reporting body for the university system, NDUS-IR must have a clear code defining expectations and limitations for the use of analytics for student and/or employee data analyses.**

1.2.1 All data captured as a result of NDUS interaction with students and employees has the potential to be used for analytics. Data should, however, only be used for analytics where there is an expectation of benefit and with a reasonable expectation that there will be no potential harm to NDUS students, employees, or system administration.

1.2.2 Artificial intelligence (AI) is a general term used to encompass a variety of applications where computers are trained to do things which humans do. AI includes many domains i.e., machine learning, data mining, automated reasoning.

1.2.3 The techniques used in analytics are based on standard statistical methods, but can involve the development of complex models, including the use of AI, which can learn from error, make automated decisions, and mimic human thought processes, albeit much more quickly than the human researcher.

1.2.4 Due to the complex nature of analytics research, a potential conflict exists between creating models which provide the most reliable outcomes and those which work in ways that can be made transparent to users and

subjects.

- 1.2.5 Analytics can be applied to individual students or employees, to defined groups of individuals, and to whole cohorts of individuals. This code and inclusive principles shall apply in all cases.
- 1.2.6 As NDUS-IR is a department whose main function is reporting rather than academic research, the most accurate and complete reporting results from the inclusion of data from all relevant students or employees, thus precluding the ability for individuals to opt out.
- 1.2.7 NDUS-IR recognizes that the use of individual level data in new ways will necessarily impact current practice. As such, several existing policies relating and referring to potential uses of student data may ultimately require review.

### 1.3 Definitions

The following definitions are intended to provide clarity about terms used throughout this code.

- 1.3.1 **Academic research** is a systematic inquiry into a subject in order to discover and/or develop theories, and generally done as part of university coursework, or by academicians furthering subject area research. Academic research is subject to the federal regulatory guidelines of the Office for Human Research Protections, and in turn often requires approval by an Institutional Review Board (IRB).
- 1.3.2 **Bias** is a prejudice in favor of or against a certain demographic when compared with another, in a way considered to be unfair.
- 1.3.3 **Cohort** refers to a specific group of students or employees established for tracking purposes.
- 1.3.4 **Descriptive Analytics** use data aggregation and data mining to provide insight into the past and answer: "What has happened?"
- 1.3.5 **Employee** refers to individuals past and currently employed at any organization within the NDUS, regardless of classification.
- 1.3.6 **In scope** is a term used to encompass data elements that are available and acceptable to use in analytic research.
- 1.3.7 **Intervention** derives from a prescriptive analytics approach and refers to information, advice and guidance directed from institutional staff to one or more students or employees.
- 1.3.8 **Personally identifiable information (PII)** is any individual level data that leads to identification of an individual in a data set. This includes data elements

such as EMPLID, name, address, etc.

- 1.3.9 **Predictive Analytics** uses statistical models, machine learning, and forecasts techniques to understand the future and answer: “What could happen?”
- 1.3.10 **Prescriptive Analytics** uses optimization and simulation algorithms and/or machine learning to advise on possible outcomes and answer: “What should we do?”
- 1.3.11 **Out of scope** is a term used to encompass data elements that are available, but not acceptable to be used in analytic research due to their sensitive nature.
- 1.3.12 **Sensitive data** defines personal data that is protected from unwarranted disclosure and/or to which access should be safeguarded. It may or may not be collected by NDUS and the NDUS institutions, and would include demographics such as gender, religion, and political affiliation. Sensitive data are defined more clearly in the **Out of Scope** section of this document, and their use is subject to limitations in analytics.
- 1.3.13 **Students** are individuals registered into credit bearing coursework at any institution within the NDUS. This does not include enquirers or informal learners.

## 2 Scope

### 2.1 In scope

- 2.1.1 In the adoption of this analytics code, the NDUS-IR establishes that certain data elements are within the scope of NDUS System Office analytic research.
- 2.1.2 In scope analytic research will be conducted only at the direction of the NDUS Director of Institutional Research, and be that which has been:
  - requested by a system office staff, NDUS institution administrator or program director, or CTS Office of the CIO,
  - requested by another state agency, legislative council or member, or governor’s office and approved by NDUS senior staff.
- 2.1.3 Data deemed to be in scope of analytic research include categories of data captured by the university system as part of its interaction with students and employees and potentially available as individual or combined data sets for use in analytics:
  - student demographic information provided by the student at registration or after their admission into the university system,

- student academic records including course enrollments and final course grades,
- student assessment data provided to the university system, including ACT scores, program placement and entry exam scores, and licensure exam scores,
- employee demographic data provided to the university system as a result of their employment with an NDUS institution, Core Technology Services, or System Office.
- employee job band, contract length, salary, and any other job-related data that falls under public record as a state employee,
- student and employee membership in any system or campus-based organizations, unless said membership is part of a group that falls under a category deemed out of scope in section 2.2.3.

## **2.2 Out of scope**

2.2.1 In the adoption of this analytics policy, the NDUS-IR establishes that certain data elements are out of scope of System Office analytic research.

2.2.2 Exceptions to the use of the following categories of data for analytic research may be considered by the NDUS-IR department under the conditions that the party requesting the research:

- owns or has possession of the data and demonstrates explicit proof of legal authority to share the out of scope data element(s), and
- signs a memorandum of understanding (MOU) that clearly indicates that the NDUS-IR is conducting the analytic research at the explicit request and with explicit permission of the requesting party.

2.2.3 Data deemed to be out of scope of analytic research include:

- sensitive data on gender, racial or ethnic origin, religious beliefs, sexual orientation, disability or other health matters, political opinions, offences (including alleged offences),
- individual level data from external sites not controlled by the university system or system institutions (e.g. social networking sites), with the exception of third-party vendors who are specifically contracted to provided services to the system or its institutions,
- data on student or employee complaints,
- data relating to enquirers and informal learners other than

matriculated/enrolled students, alumni, and current or past employees,

- data related to NDUS and institutional job applicants who have not yet gained employment in the system.

## **2.3 Ethical issues relating to the use of student data for academic research**

- 2.3.1 NDUS-IR will not conduct academic research for the purpose of publication or other public presentation without explicit guidance and permission of the institution(s), department, or agency with legal authority over the data elements.
- 2.3.2 Data requests to NDUS-IR for the purposes of academic research must fall within the boundaries of this policy and, when necessary, have prior approval by the appropriate Institutional Research Board (IRB).
- 2.3.3 The NDUS-IR is bound by the Association for Intuition Research *Code of Ethics and Professional Practice* in all analysis and reporting responsibilities.

## **3 Ownership and use of analytic research**

### **3.1 Statement regarding the ownership of analytic research**

- 3.1.1 Any reporting and/or research conducted as an employee of the NDUS-IR department is the express intellectual property of the North Dakota University System and/or institution, agency, or department who owns the data.
- 3.1.2 Reporting and/or research conducted within the NDUS-IR cannot be used for academic research, conference presentation, or any other public exhibition unless expressly approved by the Director of Institutional Research and other relevant stakeholders.

## **4 Guiding principles on the ethical use of data for analytics**

These principles aim to set out how the NDUS-IR will conduct analytic research and subsequent reporting. It is based around key principles that seek to provide legal, policy, and ethical guidance to NDUS-IR staff conducting research, as well as to stakeholders with research requests:

### **4.1 Principle 1: Adherence to federal and state laws regarding data privacy and protection.**

- 4.1.1 NDUS-IR will adhere to all federal and state laws regarding data privacy and protection, both current and those enacted in the future.
- The Federal Education Rights and Protection Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99) provides primary guidance regarding the control and use of student data in the North Dakota University System.

## **4.2 Principle 2: Humans are decision makers of ethical use of analytics**

4.2.1 NDUS-IR recognizes that technology in and of itself is amoral, and incapable of making data decisions based on perceptions of positive and negative outcomes.

- The foremost guiding principle must be to first do no harm to students and employees
- It is ultimately the responsibility of the institutional researcher to decide whether the use of analytics on a particular study is ethical
- Decisions regarding the ethical use of analytics is a decision best shared among a group of knowledgeable, informed professionals
- Ethical use of analytics includes not only whether it is used, but also which data elements are included in the analysis, and which inferences can be ethically drawn from research findings
- Examples of unethical uses of data may include using results to disincentivize or unjustly discourage students or employees from participating in certain activities or prohibiting students from enrolling in coursework for which they have the pre-requisites.

## **4.3 Principle 3: Data models based on analytics should be free from bias**

4.3.1 With awareness that data models can become established that bias individuals on a particular demographic, the researcher is responsible for implementing best practices to eliminate bias and/or to abandon any model that is found to discriminate

- The inclusion of out of scope data elements in this policy seeks to promote non-bias by removing sensitive data field from the data set
- Although out of scope data elements have been removed from analysis, the researcher has the responsibility to conduct additional testing to ensure unintended bias has not been an influence in a decision derived from prescriptive analytics

## **4.4 Principle 4: Individuals should not be wholly defined by their visible data or by NDUS-IR interpretation of that data**

4.4.1 Individuals are unique, and analytics cannot predict an individual's outcome with 100% accuracy. Rather analytics predict the chances of particular outcomes based on characteristics held in common with the studied cohort.

- Analytics generate insight that may allow researchers to determine possibilities of alternate outcomes
- Descriptive and Predictive Analytics reflect on what happened in the past, and not what will happen in the future, only what has the probability to happen
- Individuals do not always follow predicted patterns of behavior
- Individuals should not be stereotyped by any form of analytics

#### **4.5 Principle 5: NDUS-IR will be transparent in their conduct of research**

- NDUS-IR will communicate to relevant stakeholders what analytic research is being conducted and for what purpose
- NDUS-IR will maintain an updated policy for the use of data analytics for research
- NDUS-IR will respond to inquiries regarding analytic research currently being conducted, conducted in the past, and planned for the future

### **5 Implementation & Review**

**5.1** This code is implemented as of date indicated below and remains in effect until future revision by NDUS-IR or superseded by adopted NDUS SBHE policy.

**5.2** This code is modeled after **The Open University** document *Policy of Ethical use of Student Data for Learning Analytics*, which is released under the following Creative Commons license:  
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Effective date: 04.05.2019

North Dakota University System  
 Office of Institutional Research