

## **Emma Stamm | Syllabus concept: “Knowledge After Big Data”**

### ***About the Course***

Digital technologies generate over 2.5 quintillion bytes of data each day. As data proliferate, they present novel opportunities for inquiry, creativity, and critique. “Knowledge After Big Data” explores epistemological research and methodological approaches pertinent to Big Data and the humanities.

The first section of the course introduces theoretical perspectives on the relationship between data and knowledge. Here, we take seriously the possibility that Big Data is changing standards of evidence across multiple areas of knowledge production. In the second section, we review a suite of digital methods applicable to humanities scholars. Lectures and assigned materials promote technical literacy and present scholarly reflections on digital sensemaking tools. In the third section, we explore digital humanities projects which address the social and intellectual repercussions of Big Data. The course culminates with a final project: students have the option to submit a traditional paper or a digital humanities project proposal.

### ***Grade components/assignments***

#### **Grade Component 1: The first assignment (pass/fail)**

There are two parts to the first assignment:

- 1) Fill out the questionnaire, which I will send as a Google form on the first day of class. It should only take twenty minutes. Remember that the questionnaire is fully anonymous and that you may choose not to answer any question. This questionnaire will be used to generate the dataset for Grade Component 5.
- 2) Upload a document to Canvas under “Assignments” > “First Assignment.” This document must include a brief autobiography (minimum one paragraph) and notification of whether you choose option A or B for Grade Component 4 (see below).

Note that the first assignment is pass/fail. If you do not turn it in on time, you will lose half a point on your final grade for every day late.

#### **Grade component 2: Discussion posts (5 points each; 20 points total)**

Four times throughout the semester, you will make a three-hundred word discussion post on Canvas. These respond to a specific prompt regarding recent content.

#### **Grade component 3: Data mining exercises (2.5 points each; 5 points total)**

At the end of week six, please complete any two of the five exercises on page 15 of *Discovering Knowledge in Data*. You will upload this as a document to Canvas under “Assignments” > “Data mining exercises.”

#### **Grade component 4: Group qualitative coding (20 points)**

For this assignment, your collaboration group will come up with a qualitative coding scheme for the class dataset. You will use this schema to analyze the data. Submissions include multiple deliverables, including a written. More information will be provided early in the semester. One selected group member will upload your deliverables to Canvas under “Assignments” > “Group qualitative coding.”

#### **Grade component 5: Midterm exam or midterm paper (you choose) (25 points)**

You may choose to A) take an exam or B) submit a short paper (minimum 1,200 words) at midterms. More information will be provided early in the semester.

#### **Grade component 6: Final project (5 points for peer review; 25 points for project; total 30 points)**

You may choose to A) write a traditional paper or B) submit a comprehensive proposal for a humanities research project that deploys digital methods. For both options, the submission will be no shorter than 2,000 words.

Two weeks before the final project is due, you will send a working draft to each of your collaboration group members. You will complete a peer-review sheet for each draft you receive and provide it to the author of the project one week later. The peer-review sheets for your projects — i.e., the sheets you receive from your peers, not those you fill out — are due to the instructor along with the final project submission.

### ***Course schedule***

#### **Unit I: Theory**

##### **Week 1**

##### **Reading assignments:**

Lisa Gitelman and Virginia Jackson: “Introduction” from *Raw Data is an Oxymoron* (PDF on Canvas)

Alexander R. Galloway: [“From Data to Information”](#)

##### **Week 2**

##### **Reading assignments:**

Kate Crawford, Kate Miltner, and Mary L. Gray: [“Critiquing Big Data: Politics, Ethics,](#)

## Epistemology”

Rob Kitchin: [“Big Data, New Epistemologies and Paradigm Shifts”](#)

### **Week 3**

#### **Reading assignments:**

Chris Anderson: [“The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”](#)

Nathan Jurgenson: [“View from Nowhere”](#)

Geoffrey Bowker: [“The Theory/Data Thing”](#)

#### **Video assignment:**

Theorizing the Web 2018 Keynote (various participants): [“God View”](#)

### **Week 4**

#### **Reading assignments:**

Alan Liu: [“Theses on the Epistemology of the Digital”](#)

Peter de Bolla: [“Digital Knowledge: Format, Scale, and the Information-knowledge Parallax at the Total Knowledge Horizon — A Reply to Alan Liu”](#)

#### **Video assignment:**

Timandra Harkness: [What is Knowledge in the Age of Big Data?](#)

## **Unit II: Methods**

### **Week 5**

#### **Reading assignment:**

Boris Lubarsky: [Re-Identification of Anonymized Data](#)

#### **Video assignment:**

Katharina Rasch: [What every data scientist should know about data anonymization](#)

### **Week 6**

#### **Reading assignment:**

Qualitative Coding (PDF on Canvas)

#### **Video assignment:**

[Intro to Qualitative Coding](#)

### **Week 7**

#### **Reading assignment:**

Daniel T. Larose and Chantal D. Larose: [Discovering Knowledge in Data: An Introduction to Data Mining](#) pp. 1-15

### **Week 8**

#### **Reading assignment:**

Pedro Domingos: [A Few Useful Things to Know About Machine Learning](#)

Adrian Mackenzie: Introduction and Chapter 1, "Into the Data," from *Machine Learners: Archaeology of a Data Practice* (PDF on Canvas)

Week 9

**Reading assignment:**

Colin Allen and Jaimie Murdock: [LDA Topic Modeling: Contexts for the History and Philosophy of Science](#)

**Unit III: Critical Digital Humanities projects**

Week 10

**Reading assignment:**

Oliver L. Haimson and Anna Lauren Hoffman: "[Constructing and Enforcing 'Authentic' Identity Online: Facebook, Real Names, and Non-Normative Identities](#)"

**Project case studies:**

Ben Grosser: [Instagram, Twitter, and Facebook Demetricators](#)

Week 11

**Reading assignment:**

Finn Brunton and Helen Nissenbaum: "[Vernacular Resistance to Data Collection and Analysis: a Political Theory of Obfuscation](#)"

**Project case studies:**

Daniel C. Howe, Mushon Zer-Aviv, and Helen Nissenbaum: [Ad Nauseam](#)

Ben Grosser: [Go Rando; ScareMail](#)

Week 12

**Reading assignments:** Ingrid Burrington: "[Mapping the Amazon](#);" *ibid.*: "[Policing Is An Information Business](#)"

**Project case study:**

Ingrid Burrington: [Networks of New York](#)

Week 13

**Reading assignments:**

John Cheney Lippold, "[Who we are as data might soon become more important than who we are as people](#)"

Sun-ha Hong: "[Data's Intimacy: Machinic Sensibility and the Quantified Self](#)"

**Project case study:**

James Bridle: [Citizen Ex](#)

**Week 14**

**Reading assignment:**

Olivia Solon: ["The rise of 'pseudo-AI': How Tech Firms Quietly Use Humans to Do Bots' Work"](#)

Lilly Irani: ["Justice For Data Janitors"](#)

**Project case study:**

Lilly Irani: [Turkopticon](#)

**Weeks 15/16:**

No assigned material. Students are expected to work on class presentations and final projects.