

# INF312H1: Worlds become data

Rohan Alexander

2024-01-19

## Table of contents

<b>1</b>	<b>Preamble</b>	<b>2</b>
1.1	Official course description . . . . .	2
1.2	Course Objectives . . . . .	2
1.3	Past Syllabus . . . . .	2
1.4	Pre-requisites . . . . .	2
1.5	Textbook . . . . .	2
<b>2</b>	<b>Content</b>	<b>3</b>
<b>3</b>	<b>Assessment</b>	<b>4</b>
3.1	Summary . . . . .	4
<b>4</b>	<b>Other</b>	<b>6</b>
4.1	Children in the classroom . . . . .	6
4.2	Accommodations with regard to assessment . . . . .	6
4.3	Re-grading . . . . .	7
4.4	Plagiarism and integrity . . . . .	7
4.5	Late policy . . . . .	8
4.6	Writing . . . . .	8
4.7	Minimum submission requirement . . . . .	8
4.8	Use of Generative AI in assignments . . . . .	8
4.9	Accommodations . . . . .	9
4.10	Writing Support . . . . .	9
4.11	Equity, Diversity, and Inclusion . . . . .	10
4.12	Learning Hub . . . . .	11
4.13	Academic integrity . . . . .	12
4.14	Course Learning Outcomes . . . . .	12
4.15	Relation to Program Learning Outcomes . . . . .	12

# 1 Preamble

- Course Instructor: Rohan Alexander
- E-mail: rohan.alexander@utoronto.ca

## 1.1 Official course description

This course (INF312H1 — Worlds Become Data) covers issues in the practices of translating phenomena to data and algorithmic description. What happens, what is gained, what is lost, when things that happen in the world are recorded and made into information or recorded as a document? The course explores representation, modeling, correctness, reliability, and bias in different types of data and algorithms. We will learn about diverse topics such as cultural and algorithmic bias, challenges of big data, what happens when the world is transformed into images, what are the implications of having your social status determined by data and scores on your social media profile, and what we gain or miss when we deal with geographical information systems.

## 1.2 Course Objectives

To a certain extent we are wasting our time. We have a perfect model of the world—it is the world! But it is too complicated. Because of this we must simplify the world in order for it to become data. In this course we explore how we do this, and the implications.

## 1.3 Past Syllabus

- [2023](#).
- [2022](#).

## 1.4 Pre-requisites

- INF301H1 – Introduction to Information and Power and
- INF302H1 – Integrative Approaches to Technology and Society.
- Or permission from the instructor

## 1.5 Textbook

*Telling Stories with Data*

## 2 Content

- Week 1
  - Drinking from a fire hose
- Week 2
  - Reproducible workflows
  - Guest: Samita Prabhasavat
- Week 3
  - R essentials
- Week 4
  - Writing research
  - Guest: Steven Coyne - “Who Owns This? The Ethics of Copyright”
- Week 5
  - Static communication
- Week 6
  - Farm data
- Week 7
  - Gather data
- Week 8
  - Hunt data
- Week 9
  - Clean and prepare data
- Week 10
  - Store and share data
- Week 11
  - Exploratory data analysis
- Week 12
  - Linear models

### 3 Assessment

#### 3.1 Summary

Item	Weight (%)	Due date	Notes
Quiz	7	Wednesdays, noon, Weeks 1-12	Only best seven out of twelve count.
SQL quiz	1	Wednesday, noon, Week 6	
Personal website	1	Wednesday, noon, Week 9	Create a personal website using Quarto and make it live via GitHub Pages. At a minimum, it must include a bio and a CV in PDF form.
Tutorials	6	Wednesdays, noon, Weeks 1-12	Only best three out of twelve count.
Term papers	48	<p>Wednesdays, noon, Weeks 3, 6, 9</p> <p>Term Paper I: 24 January 2024</p> <p>Term Paper II: 14 February 2024</p> <p>Term Paper III: 13 March 2024</p>	<p>You must submit Term Paper I in order to pass the course.</p> <p>Only best two of three term papers count.</p> <p>Marking starts, noon, on the Friday after submission, and you can update until then i.e. submissions made by noon, Wednesday, Week 3 can be updated until noon, Friday, Week 3 (this is to allow you to incorporate peer review comments).</p> <p>Please do not make any changes after marking starts.</p> <p>Term Paper I: <a href="#">Donaldson Paper</a></p> <p>Term Paper II: <a href="#">Mawson Paper</a>.</p> <p>Term paper III: <a href="#">Howrah Paper</a>.</p>

Item	Weight (%)	Due date	Notes
Conduct peer review of Term/Final papers	3	Thursdays, noon, Weeks 3, 6, 9, 12	Conduct peer review for six other term/final papers, by creating a GitHub Issue or Pull Request. Papers will be distributed by a spreadsheet—add a link to the Issue/PR to a term paper that does not have four other entries. You will only have 24 hours to do this.
Final paper	34	Wednesday, noon, Week 12 (3 April 2024)	You must submit this paper. Marking starts, noon, Monday 22 April and you can update until then i.e. submissions made by noon, Wednesday, Week 12 can be updated until noon, Monday, 22 April (this is to allow you to incorporate peer review comments). Please do not make any changes after marking starts.   <a href="#">Final paper.</a>

### **You must submit Term Paper 1. You must submit the Final Paper.**

Beyond that, you have scope to pick an assessment schedule that works for you. I will take your best three of the twelve tutorials for that six per cent, and your best seven of twelve quizzes for that seven per cent. I take your two best papers from the three term papers for that 48 per cent (24 per cent for each). You get up to three percentage points for conducting peer review of other student papers, (half a percentage point per review). There is 34 per cent allocated for the Final Paper.

Additional details:

- Quiz questions are drawn from those in the Quiz section that follows each chapter of *Telling Stories with Data*. Some of them are multiple choice, and you should expect to know the mark within a few days of submission. Please do them before coming to class.
- Tutorial questions are drawn from those in the Tutorial section that follows each chapter of *Telling Stories with Data*. The general expectation (although this differs from week to week) is about two pages of written content. You should expect to know the mark within a few days of the tutorial.
- In general term papers require a considerable amount of work, and are due after the material has been covered in quizzes and tutorials (i.e. you would draw on knowledge tested in the quizzes, and potentially material could be re-used from the tutorial material). In general, they require original work to some extent. Papers are taken from the Papers appendix of *Telling Stories with Data* and students have access to the grading rubrics before submission.
- If you already have a website, please communicate with me about this early in the term so that I can let you know whether it can be used for the purposes of this submission.
- While they vary, a rough rubric for tutorial is:

- 0 - Any typos, grammatical errors, other table stakes issues for this level. Submission is too short. Other basic mistakes.
  - 0.25 - Tables/graphs not properly labeled, no references, other aspects that affect credibility.
  - 0.5 - Makes some interesting and relevant points, related to course material (including required materials), but lacking in terms of structure and story/argument.
  - 0.80 - Interesting submission that is well-structured, coherent, and credible.
  - 1 - As with 0.80, but exceptional in some way.
- Only the best two of three term papers counts. This means each is worth 24 per cent.
  - Peer review will occur for Term Paper I, but students are typically not yet familiar enough with the expectations of the course so as to be able to provide valuable comments (other than noticing whether R has been cited!).

## 4 Other

### 4.1 Children in the classroom

Babies (bottle-feeding, nursing, etc) are welcome in class as often as necessary. You are welcome to take breaks to feed them or express milk as needed, either in the classroom or elsewhere including [here](#). A list of baby change stations is also available [here](#). Please communicate with me so that I can make sure that we have regular breaks to accommodate this.

For toddlers and older children, I understand that unexpected disruptions in childcare/school can happen. You are welcome to bring your child to class in order to cover unforeseen gaps.

### 4.2 Accommodations with regard to assessment

Please do **not** reveal your personal or medical information to me. I understand that illness or personal emergencies can happen from time to time. The following accommodations to assessment requirements exist to provide for those situations.

Straight-forward (will automatically apply to all students—there is no need to ask for these):

- Quiz: Only your best seven quizzes count.
- Tutorial: Only your best three tutorials count.
- Term Papers: Only your best two term papers count.

So for those, if you have a situation, then just do not submit (or in the case of Term Paper I, just submit a blank page).

Slightly more involved:

- Term Paper I: You must submit something for Term Paper I, even if it gets zero. If you have a medical emergency that makes it impossible for you to submit even a blank page for Term Paper I, then please email me. In that situation one of the remaining term papers must be done individually to ensure fairness with the rest of the class.
- Peer review: No accommodation or late submission is possible for this because it would hold up the rest of the class. That said, there are many opportunities to get the peer review marks, so if you cannot do any for a particular paper, then just do the others. If you have a medical emergency that makes this impossible, then please email me and cc your faculty/department/college advisor so that we can work out an alternative plan.
- Final paper: The final paper is a critical piece of assessment. It is also up against deadlines for submission of grades (especially for graduating students). If you have a medical emergency that makes it impossible for you to submit before marking begins, then I **may** be able to grant you an extension of up to three days. Email me and cc your faculty/department/college advisor so that we can work out a alternative plan.

### 4.3 Re-grading

Marking mistakes happen and I want to correct those. Requests to have your work re-graded will **not** be accepted within 24 hours of the release of grades. This is to give you a chance to reflect. Similarly, requests to have your work re-graded more than seven days after the release of the grades will not be accepted. This is to ensure the course runs smoothly.

Inside that 1-7 day period if you would like to request a re-grade, please email rohan.alexander@utoronto.ca and use the subject line “INF312: re-grade request”. Please specify where the marking mistake was made in relation to the marking guide. The entire assessment will be re-marked and it is possible that your grade could reduce.

Plenty of students get 0 on the first paper, but go on to get an A+ overall in the course. The nature of the work in this course requires students to adjust from what is expected in other courses, and the forgiving assessment weighting is designed to allow this.

### 4.4 Plagiarism and integrity

Please do not plagiarize. In particular, be careful to acknowledge the source of code—if it is extensive then through proper citation and if it is just a couple of lines from Stack Overflow then in a comment immediately next to the code.

You are responsible for knowing the content of the University of Toronto’s Code of Behaviour on Academic Matters.

Academic offenses include (but are not limited to) plagiarism, cheating, copying code without acknowledgement, purchasing labor for assessments (of any kind). Academic offenses will be

taken seriously and dealt with accordingly. If you have any questions about what is or is not permitted in this course, please just email me.

Please consult the University's [site](#) on Academic Integrity. Please also see the definition of plagiarism in section B.I.1.(d) of the University's Code of Behaviour on Academic Matters available [here](#). Please read the Code. Please review Cite it Right and if you require further clarification, consult the site [How Not to Plagiarize](#).

#### **4.5 Late policy**

If no extension has been granted and no accommodation applies, then late submissions will not be accepted.

#### **4.6 Writing**

Papers and reports should be well-written, well-organized, and easy to follow. They should flow easily from one point to the next. They should have proper sentence structure, spelling, vocabulary, and grammar. Each point should be articulated clearly and completely without being overly verbose. Papers should demonstrate your understanding of the topics you are studying in the course and your confidence in using the terms, techniques and issues you have learned. As always, references must be properly included and cited. If you have concerns about your ability to do any of this then please make use of the writing support provided to the faculty, colleges and the [SGS Graduate Centre for Academic Communication](#).

#### **4.7 Minimum submission requirement**

If you are going to not be able to submit at least two term papers, and/or be unable to submit the final paper then it would be unfair on the other students to allow you to pass the course. But it is not a situation that I want to get into. Please ensure you and your college registrar or faculty/department advisor get in touch with me as early as possible if this may be the case for you so that we can work out a solution.

#### **4.8 Use of Generative AI in assignments**

In general, students are encouraged to use generative AI tools as a starting point. Specific course policies are:



- **Code:** Students may wish to use generative AI tools to aid in initial development and writing of code to answer assignment questions and carry out the research project analysis. If this is the case, the use of such tools should be explicitly acknowledged in the submitted work, and the relevant prompts and responses should be included in a text file in the repo. All code, regardless of how it is generated must be thoroughly commented and explained. Failure to do so may result in penalties.
- **Written work:** Using generative AI tools to generate written answers to assignment questions of text contained in the final research project is prohibited in this course. Representing as one's own an idea, or expression of an idea, that was AI-generated may be considered an academic offense in this course. However, you are welcome to use it to generate a first draft, which you then completely edit. Again, if you use such tools, it should be explicitly acknowledged in the submitted work, and the relevant prompts and responses should be included in a text file in the repo.

This course policy is designed to promote your learning and intellectual development and to help you reach course learning outcomes.

## 4.9 Accommodations

Students with diverse learning styles and needs are welcome in this course. In particular, if you have a disability/health consideration that may require accommodations, please feel free to approach me and/or the Accessibility Services Office as soon as possible. Accessibility Services staff are available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations. The sooner you let us know your needs, the quicker we can assist you in achieving your learning goals in this course.

To book an appointment with an Accessibility Advisor, please connect with the Accessibility Services front desk via email at [accessibility.services@utoronto.ca](mailto:accessibility.services@utoronto.ca) or call (416) 978-8060. Consultation appointments are available to discuss any questions about the Accessibility Services registration process and/or potential accommodation support. The on-location Accessibility Advisor at the Faculty of Information is Michael Mercer.

## 4.10 Writing Support

As stated in the Faculty of Information's Grade Interpretation Guidelines, "work that is not well written and grammatically correct will not generally be considered eligible for a grade in the A range, regardless of its quality in other respects." With this in mind, please make use of the writing support provided to graduate students by the [SGS Graduate Centre for Academic Communication](#). The services are designed to target the needs of both native and non-native speakers and all programs are free. Please consult the current SGS Workshops Schedule for more information.

## 4.11 Equity, Diversity, and Inclusion

The University of Toronto is committed to equity, human rights and respect for diversity. All members of the learning environment in this course should strive to create an atmosphere of mutual respect where all members of our community can express themselves, engage with each other, and respect one another's differences. U of T does not condone discrimination or harassment against any persons or communities.

The Equity, Diversity, and Inclusion Unit (EDIU) at the School of Information, in collaboration with U of T community members, works to promote and encourage an equitable and inclusive work and classroom environment, free from discrimination and/or harassment based on any of the code grounds. The EDI Unit is responsible for developing and delivering EDI programs and services, works with all stakeholders, and provides confidential services. Key areas of services include:

- Training and educational opportunities
- Community building and engagement
- Systemic change initiatives
- Providing confidential advice/consultations
- Supports with resolving concerns of discrimination and/or harassment

EDIU Complaints Resolution Assistance Process:

- Meet with individuals (all stakeholders) to listen and discuss concerns or questions related to any of the human rights protected grounds
- A complaint is not necessary to approach the EDI Director with questions or to seek information
- Talk about options available for resolution assistance, including informal and formal complaint options
- Provide referrals or liaise with other departments or stakeholders where necessary
- Maintain confidentiality of queries that people bring forward as legally required. Limitations to confidentiality are discussed if safety is raised
- No steps are taken to address a complaint without consent

Please note, there will be changes to this process in the near future; however, if you do have any questions/concerns, feel free to reach out to [ediu.ischool@utoronto.ca](mailto:ediu.ischool@utoronto.ca)

FOI Equity, Diversity, and Inclusion Unit (EDIU)

Website: [Equity, Diversity & Inclusion Resources - Faculty of Information \(iSchool\) | University of Toronto \(utoronto.ca\)](#)

Tri-Campus Equity Offices

U of Toronto Tri-Campus Equity Offices

Website: [Equity Offices – The Division of People Strategy, Equity & Culture \(utoronto.ca\)](#)

- AODA Office
- Anti-Racism & Cultural Diversity Office (ARCD O)
- Family Care Office
- Office of Indigenous Initiatives
- Sexual & Gender Diversity Office (SGDO)
- Sexual Violence Prevention & Support Centre (SVPSC)

Related Offices (websites):

- Community Safety Office – University of Toronto (utoronto.ca)
- First Nations House - UofT Student Life (utoronto.ca)
- Wellness – The Division of People Strategy, Equity & Culture (utoronto.ca)
- Multi-Faith Centre for Spiritual Study & Practice - UofT Student Life (utoronto.ca)

## 4.12 Learning Hub

The Faculty of Information Learning Hub can support your learning in this course in a range of ways. We offer programs, workshops, and services to support your learning, as well as a physical place – on the 4th floor of Bissell – for gathering, seeking help, finding resources, studying, creative making, relaxing, playing and collaborating. Additionally, we provide an Virtual Learning Hub that provides resources and sign ups for services and events. Below are an abbreviated list of our services:

**Cite it Right:** All incoming students must complete the Cite it Right online workshop and quiz within the month of September. Cite it Right, with its focus on academic integrity, was designed to familiarize students with the University’s Code of Behaviour on Academic Matters and, more generally, help them build confidence as they work with sources. Both the workshop and quiz are located in the Virtual Learning Hub. Please note that the Dean’s Office monitors the completion of these modules, as well as quiz scores.

**iSkills Workshops:** The iSkills co-curricular workshop series is an expansive program that addresses scholarly, professional, and technical competencies aligned with Faculty of Information academic programs. Rosters are built every term to reflect students’ current needs along with trends in the information and heritage professional worlds. View the current roster of workshops and learn more about the program on our iSkills site.

**Tutors:** The Learning Hub offers one-on-one tutoring services to support writing, research, and technical skills. You can learn more about our tutors’ specific areas of expertise, how they can support you, and sign up for individual tutoring on our Writing, Research & Technical Skills Support page. They can help you with assignments for this course at any stage – conceptualizing and planning, drafting, refining, and even after you have received your mark, to help you understand your instructor’s comments and plan for your next assignment.

**Library Support:** The University of Toronto Libraries (UTL) provides a liaison to the Faculty of Information, who is familiar with the specific needs of our students. Yoonhee Lee can connect

you to UTL resources, services, and tools, as well as support you with research projects, citation management, and other research-related tasks.

### 4.13 Academic integrity

Please consult the University's site on [Academic Integrity](#). The Faculty of Information has a zero-tolerance policy on plagiarism as defined in section B.I.1.(d) of the University's Code of Behaviour on Academic Matters (PDF). You should acquaint yourself with the Code. Please review the material in Cite it Right and if you require further clarification, consult the resource How Not to Plagiarize (PDF). Cite it Right covers relevant parts of the U of T Code of Behaviour on Academic Matters (1995). It is expected that all Faculty of Information students complete the Cite it Right module and the online quiz prior to the second week of classes of their first term.

### 4.14 Course Learning Outcomes

1. Critique the major methodological and computational challenges related to the creation and representation of structured and unstructured data (All papers and quizzes)
2. Analyze, and synthesize the major ethical debates and recurring patterns pertaining to data stewardship and circulation (All papers)
3. Develop innovative research design to overcome the methodological, sociopolitical and ethical challenges associated with current methods of analysis of complex information practices involving structured and unstructured data. (All papers)
4. Communicate and present research proposal to overcome the methodological, sociopolitical and ethical challenges associated with current methods of analysis of complex information practices. (Papers and tutorials)

These will be measured through assignments that support understanding of the methodological and computational debates related to data creation and representation (CLO #1) and the ethical debates pertaining to data stewardship and circulation (CLO#2), reflective papers that help develop and relate insights regarding the methodological, sociopolitical and ethical challenges associated with current methods of analysis of complex information practices involving structured and unstructured data (CLO#1, CLO#2, CLO#3), and produce work that help relate their reflections in textual and non-textual forms. (CLO #4.)

### 4.15 Relation to Program Learning Outcomes

Huge amounts of data are produced everyday including different types of data such as structured quantifiable data, unstructured text, and multimedia data. Studying the social world through shaping it into one of the above-mentioned data types has its own advantages and limitations. This course will help students to understand and assess the social, political,

economic, and ethical entailments of information creation, ownership, stewardship, and circulation, especially in light of enduring and emerging ethical and political questions (PLO1r). Given the case studies applied in class and practical assignments in this course, the students will be able to critique the conceptual and philosophical foundations of representation and computation, and to recognize recurring patterns of unresolved intellectual and social tension (PLO3i and PLO11r). With the knowledge acquired in this course, the students will be able to develop, defend, and use methods of analysis of complex information practices and the political, economic, technical, and cultural contexts in which they occur (PLO10i).