



University at Buffalo

Department of
Information Science

Graduate School of Education

We acknowledge the land on which the University at Buffalo operates, is the territory of the Seneca Nation, a member of the Haudenosaunee/Six Nations Confederacy.

LIS 501 / 612: Information Retrieval (Special Topics in Information Management)



Thomas Rowlandson, *Inside View of the Public Library, Cambridge*, 1809.

Table of Contents

Instructor Information	2
Course Information	2
Course Description	3
Course Topics	3
Learning Outcomes	3
Mode of Instruction	4
Course Technologies & Required Technology Skills	4
University Libraries	5
Course Assignments	5
Grading Policy	5
Weekly Course Schedule	6
Required Course Materials	14
Course Communication	14
Course Evaluation	15
Academic Integrity	15
Accessibility Policies and Services	15
Student Success and Wellbeing Policies and Services	16

Instructor Information

Instructor: Samuel Dodson, Ph.D. ([he/him/his](#))
 Office Location: 524 Baldy Hall (North Campus)
 E-mail: smdodson@buffalo.edu
 Phone: (716) 645-1488
 Office Hours: By appointment

Course Information

Program: Information and Library Science
 Semester: Spring 2024
 Dates: 1/24/2024-5/7/2024
 Credits: 3
 Type of Instruction: Lecture
 Delivery Mode: Online
 Location: Remote
 Learning Management System: [UB Learns](#)
 Course Prerequisites: None

This course satisfies the *Information Management Electives* category of the [Plan of Study](#).

As per the [Handbook](#), successful completion of this course is defined as completion with a grade of C or better or with a grade of S.

Course Description

This course provides an introduction to information retrieval. Students will learn about theories and techniques for automatically processing, storing, and retrieving documents. Topics include indexing data structures and algorithms, retrieval models (geometric, logic-based, and probabilistic), query languages, search user interfaces, methods of system evaluation, and ethical issues, such as bias, censorship, and privacy. Students will apply these teachings to a number of domains, such as conversational search, digital libraries, multimedia search, recommender systems, and search engines.

Course Topics

- Current information retrieval research issues and trends,
- Indexing data structures, with emphasis on storage and retrieval of information on the Web,
- Use of search vocabularies (e.g., controlled and free-text),
- Models of information retrieval, including Boolean, vector space model, and probabilistic,
- Presentation of results to users,
- Evaluation of efficiency and effectiveness, and
- Applications of information retrieval.

Learning Outcomes

Upon completion of this course, students will be able to:

- Understand the foundations of information retrieval concepts, models, and theories,
- Analyze and discuss current issues and research in information retrieval,
- Understand the capabilities and limitations of information retrieval systems, and
- Identify and discuss challenges and future developments in information retrieval.

This course addresses the following Master of Science in Information and Library Science (MS ILS) [program goals](#):

1. Graduates demonstrate theoretical and conceptual understanding of information science, including the creation, representation, organization, retrieval, dissemination, use and curation of information.
2. Graduates are prepared to apply disciplinary knowledge and skills in a variety of information contexts.

3. Graduates demonstrate professional competences, including leadership, critical and analytical thinking, research, communication, collaboration, cultural competence, reflective practice, and adherence to professional ethics.
4. Graduates demonstrate values, attitudes and behaviors that are essential for information and library professionals, including diversity, equity and inclusion.

Mode of Instruction

This is a remote and asynchronous course. All course materials will be available through UB Learns, [UB Libraries](#), or the Web.

The objectives for this course will be accomplished through the following:

- Instructional video lectures,
- Readings from the textbook, book chapters, journal articles, and conference papers, and
- Individual assignments.

Each week will begin on Monday at 9:00 AM EST / EDT, except for [holidays officially observed by the University and Fall / Spring Break](#). Weekly course materials will be available on UB Learns at the beginning of each week. Weekly activities should be completed six days later, on Sunday at 11:59 PM.

Course Technologies & Required Technology Skills

You must have a UBITName to log in to UB Learns and access the course materials. If you do not have a UBITName, please contact the [UB Information Technology \(UBIT\) Help Center](#) at (716) 645-3542 or ubithelp@buffalo.edu.

Reliable access to the Web on a device capable of accessing UB Learns is required. A detailed description of course technology requirements may be found in the [UB Student Computer Standards](#).

For general questions about course technologies, please contact the UBIT Help Center. For questions about UB Learns, please see [UB Learns for Students](#). You can contact UB Learns at (716) 645-6188 or ublearns@buffalo.edu.

You are expected to have [Prerequisite Technology Skills](#). It is your responsibility to gain proficiency to successfully complete the course objectives and assignments.

The University at Buffalo provides students with free licenses for a number of software programs, including Microsoft 365. Please visit <http://www.buffalo.edu/ubit/service->

[guides/software.html](#) for more information.

University Libraries

You have full access to the [University Libraries](#) resources and services. Through the University Libraries, you have online access to many full-text databases, including: [Library Literature & Information Science](#) and [Library, Information Science & Technology Abstracts](#). You can also request physical and electronic materials using the [Delivery+](#) service.

[Molly Dahl Poremski](#) is the Information Science subject librarian, and is available by phone: (716) 645-7750 and e-mail: poremski@buffalo.edu.

If you have a general reference question, you can also use the [Instant Librarian](#) service.

Course Assignments

<i>Assignment</i>	<i>Due Date</i>	<i>Weight</i>	<i>MS ILS Goals</i>
A1: Search System Profile	February 11 (Week 3)	20%	1, 2, 3
A3, Part 1: Final Project Proposal	February 25 (Week 5)	5%	1, 2, 3
A2: Wiki Page	March 10 (Week 7)	20%	1, 2, 3
A3, Part 2: Final Project Presentation	April 14 (Week 11)	10%	1, 2, 3, 4
A3, Part 3: Final Project Paper	May 5 (Week 14)	40%	1, 2, 3, 4
Participation	Throughout the semester	5%	1, 2, 3, 4

Grading Policy

This course uses a letter grading system. Grades for this course will be similar to

assessments of performance in a typical workplace:

- A = Outstanding,
- B = Meets expectations,
- C = Needs improvement,
- D = Unacceptable quality, and
- F = No effort to perform in the course.

For each letter grade, a plus (+) or minus (-) may be used.

You are expected to submit assignments by the due date. Late or missing work will be handled at my discretion.

Assignment feedback and grades will be posted throughout the semester in UB Learns. Final grades will be available through the [HUB Student Center](#).

The assignment of an incomplete grade (I) is at my discretion, and is only given in exceptional circumstances. Furthermore, you must have a passing average in coursework already completed. You must make a request for an incomplete grade in writing with me before the last day of final examinations. A detailed description of an interim grade of incomplete may be found in the [Graduate Incomplete Policy](#).

Weekly Course Schedule

You are responsible for completing all readings by the dates noted in the course schedule:

<i>Date</i>	<i>Topic</i>	<i>Required & Recommended Materials</i>
Week 1 Jan. 24-28	Introduction	<p style="text-align: center;"><i>Required</i></p> <hr/> <p>Croft, W. B., Metzler, D., & Strohman, T. (2010). Search engines and information retrieval. In <i>Search engines: Information retrieval in practice</i> (ch. 1, pp. 1-12). Pearson. https://ciir.cs.umass.edu/downloads/SEIRiP.pdf</p> <p>Croft, W. B., Metzler, D., & Strohman, T. (2010). Architecture of a search engine. In <i>Search engines: Information retrieval in practice</i> (ch. 2, pp. 13-29). Pearson.</p>

Recommended

Harman, D. (2019). Information retrieval: The early years. *Foundations and Trends in Information Retrieval*, 13(5), 425–577. <http://dx.doi.org/10.1561/15000000065>

Week 2
Jan. 29–Feb. 4

Document
Representation
& Processing

Required

Croft, W. B., Metzler, D., & Strohman, T. (2010). Processing text. In *Search engines: Information retrieval in practice* (ch. 4, pp. 73–124). Pearson.

Recommended

Manning, C. & Schütze, H. (1999). Corpus-based work. In *Foundations of statistical natural language processing* (ch. 4, pp. 117–147). MIT Press.

→PDF available on UB Learns.

Week 3
Feb. 5–11

Indexing

Nota bene: Assignment 1 is due by the end of the week.

Required

Manning, C., Raghavan, P., & Schütze, H. (2008). Boolean retrieval. *Introduction to information retrieval* (ch. 1, pp. 1–18). Cambridge University Press. <https://nlp.stanford.edu/IR-book/information-retrieval-book.html>

Manning, C., Raghavan, P., & Schütze, H. (2008). The term vocabulary and postings lists. *Introduction to information retrieval* (ch. 2, pp. 19–47). Cambridge University Press. <https://nlp.stanford.edu/IR-book/information-retrieval-book.html>

Recommended

Zobel, J., & Moffat, A. (2006). Inverted files for text search engines. *ACM Computing Surveys*, 38(2), 1–56. <https://doi.org/10.1145/1132956.1132959>

Week 4
Feb. 12–18

Query
Representation
& Processing

Required

Baeza-Yates, R. & Navarro, G. (2011). Queries: Language & properties. In R. Baeza-Yates & B. Ribeiro-Neto (Eds.), *Modern information retrieval: The concepts and technology behind search* (2nd ed., ch. 7, pp. 255–280). Addison Wesley.

→PDF available on UB Learns.

Week 5
Feb. 19–25

Access Models

Nota bene: Part 1 of Assignment 3 is due by the end of the week.

Required

Rasmussen, E. (2011). Access models. In I. Ruthven & D. Kelly (Eds.), *Interactive information seeking, behaviour and retrieval* (ch. 6, pp. 95–112). Facet Publishing.

Recommended

Roelleke, T. (2013). Information retrieval models: Foundations and relationships. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 5(3), 1–163. <https://doi.org/10.2200/S00494ED1V01Y201304ICR027>

Week 6
Feb. 26–Mar. 3

Search User
Interfaces

Required

Hearst, M. (2009). The design of search user interfaces. In *Search user interfaces* (ch. 1, pp.

1-28). Cambridge University Press.
https://searchuserinterfaces.com/book/sui_ch1_design.html

Hearst, M. (2009). Presentation of search results. In *Search user interfaces* (ch. 5, pp. 120-140). Cambridge University Press.
https://searchuserinterfaces.com/book/sui_ch5_retrieval_results.html

Recommended

Hearst, M. (2009). Integrating navigation with search. In *Search user interfaces* (ch. 8, pp. 174-210). Cambridge University Press.
https://searchuserinterfaces.com/book/sui_ch8_navigation_and_search.html

Morville, P., & Callender, J. (2010). *Search patterns: Design for discovery*. O'Reilly Media.

Week 7
 Mar. 4-10

Evaluation

Nota bene: Assignment 2 is due by the end of the week.

Required

Manning, C. D., Raghavan, P., & Schütze, H. (2008). Evaluation in information retrieval. In *Introduction to information retrieval* (ch. 8, pp. 139-161). Cambridge University Press.
<https://nlp.stanford.edu/IR-book/information-retrieval-book.html>

Recommended

Harman, D. (2011). Information retrieval evaluation. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 3(2), 1-119.
<https://doi.org/10.2200/S00368ED1V01Y201105ICR019>

Week 8
Mar. 11–17

Ethics

Required

Herman, T. (2020). Search engines and ethics. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2020 ed.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2020/entries/ethics-search/>

Noble, S. (2013). Google Search: Hyper-visibility as a means of rendering Black women and girls invisible. *In Visible Culture*, (19). <https://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/>

Recommended

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (ch. 9, pp. 167–193). MIT Press.

→PDF available on UB Learns.

Granka, L. A. (2010). The Politics of Search: A Decade Retrospective. *The Information Society*, 26(5), 364–374. <https://doi.org/10.1080/01972243.2010.511560>

Mar. 18–24

No Class
(Spring Break)

No readings.

Week 9
Mar. 25–31

Search
Engines

Required

Croft, W. B., Metzler, D., & Strohman, T. (2010). Crawls and feeds. In *Search engines: Information retrieval in practice* (ch. 3, pp. 31–72). Pearson.

Zuboff, S. (2019, September 5). How Google

discovered the value of surveillance. *Longreads*.
<https://longreads.com/2019/09/05/how-google-discovered-the-value-of-surveillance/>

Recommended

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the dangers of stochastic parrots: Can language models be too big?* In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT '21) (pp. 610–623). ACM. <https://doi.org/10.1145/3442188.3445922>

Cambazoglu, B. B., & Baeza-Yates, R. (2015). Scalability challenges in Web search engines. *Synthesis Lectures on Information Concept, Retrieval, and Services*, 7(6), 1–138. <https://doi.org/10.2200/S00662ED1V01Y201508ICR045>

Taylor, G. (2012, June). *Behind the white curtain: Search engine economics* [Video]. Retrieved from https://youtu.be/boPj_tCqZ_M

Week 10 Multimedia
 Apr. 1–7 Search

Required

Little, S., Brown, E., & R ger, S. (2011). Multimedia: Information representation and access. In I. Ruthven & D. Kelly (Eds.), *Interactive information seeking, behaviour and retrieval* (ch. 13, pp. 235–254). Facet Publishing.

Walker, R. (2009, October 14). The song decoders. *The New York Times Magazine*. <https://www.nytimes.com/2009/10/18/magazine/18Pandora-t.html>

Recommended

- Downie, J. S. (2003), Music information retrieval. *Annual Review of Information Science and Technology*, 37, 295–340. <https://doi.org/10.1002/aris.1440370108>
- Enser, P. G. B. (2008), Visual image retrieval. *Annual Review of Information Science and Technology*, 42, 1–42. <https://doi.org/10.1002/aris.2008.1440420108>
- Hill, K. (2022, May 26). A face search engine anyone can use is alarmingly accurate. *The New York Times*. <https://www.nytimes.com/2022/05/26/technology/pimeyes-facial-recognition-search.html>
- Smeaton, A. F. (2004), Indexing, browsing, and searching of digital video. *Annual Review of Information Science and Technology*, 38, 371–407. <https://doi.org/10.1002/aris.1440380109>

Week 11
Apr. 8–14

Digital
Libraries

Nota bene: Part 2 of Assignment 3 is due by the end of the week.

Required

Rasmussen, E. (2011). Library systems. In R. Baeza-Yates & B. Ribeiro-Neto (Eds.), *Modern information retrieval: The concepts and technology behind search* (2nd ed., ch. 16, pp. 685–710). Addison Wesley.

→PDF available on UB Learns.

Somers, J. (2017, April 20). Torching the modern-day Library of Alexandria. *The Atlantic*. <https://www.theatlantic.com/technology/>

[archive/2017/04/the-tragedy-of-google-books/523320/](https://archive.org/details/2017/04/the-tragedy-of-google-books/523320/)

Recommended

Borgman, C. L. (2000). Is it digital or is it a library? Digital libraries and information infrastructure. In *From Gutenberg to the global information infrastructure: Access to information in the networked world* (ch. 2, pp. 33–52). MIT Press.

→PDF available on UB Learns.

Labbe, C. (Host). (2020, February, 11). Jake Goldenfein on Google Scholar [Audio podcast episode]. In *Good Code*. Cornell Tech.
<https://www.dli.tech.cornell.edu/goodcode/episode/lea86721/jake-goldenfein-on-google-scholar>

Week 12 Social Search
 Apr. 15–21

Required

Nichols, D., & Twidale, M. (2011). Recommendation, collaboration and social search. In I. Ruthven & D. Kelly (Eds.), *Interactive information seeking, behaviour and retrieval* (ch. 11, pp. 205–220). Facet Publishing.

Thompson, C. (2008, November 21). If you liked this, you're sure to love that. *The New York Times Magazine*. <https://www.nytimes.com/2008/11/23/magazine/23Netflix-t.html>

Recommended

Konstan, J. A. (2008). *Introduction to recommender systems*. In Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '08) (pp. 1373–1374). ACM.

<https://doi.org/10.1145/1376616.1376776>

→See https://dl.acm.org/action/downloadSupplement?doi=10.1145%2F1376616.1376776&file=p1373-konstan-complete_768k.mov.

Roose, K., & Childs, M. (Hosts). (2021, May 26). Runaway recommendation engine [Audio podcast episode]. In *Planet Money*. National Public Radio. <https://www.npr.org/2021/05/26/1000529597/runaway-recommendation-engine>

Schrage, M. (2020). *Recommendation engines*. MIT Press.

Week 13 Apr. 22–28	Final Project Presentations	No readings.
Week 14 Apr. 29–May. 5	No Class (Final Project Paper Preparation)	<i>Nota bene: Part 3 of Assignment 3 is due by the end of the week.</i> No readings.

Required Course Materials

All course materials will be freely available through UB Learns, UB Libraries, or the Web. The required and recommended readings are listed in the course schedule.

Course Communication

I endeavor to create a learning environment in which people of all identities are welcomed and supported. I ask you to, optionally, share your pronouns and preferred chosen name with me at the beginning of the semester. Similarly, you can update your [pronouns and gender identity](#) and [preferred chosen name](#) in the HUB Student Center.

I encourage you to actively participate in the UB Learns discussion forum throughout the semester. You will find that reading your peers' posts and writing your own will be beneficial to your understanding of the course materials. I read all discussion forum posts,

but am not always able to respond. In this class, you are expected to engage in respectful communication. I will not tolerate any form of disrespect directed towards anyone in the class.

If you have any questions or comments about the course, please do not hesitate to contact me. Because this is an asynchronous online course, I prefer to communicate by e-mail. All messages sent to me should come from your @buffalo.edu e-mail address. When contacting me, please include the course number in the subject of your e-mail. I will do my best to respond to you within 48 hours between Monday at 9:00 AM and Friday at 5:00 PM.

Course Evaluation

Course evaluations are an important part of the University's commitment to academic excellence. Once you complete the course, you will receive an e-mail from [UB Course Evaluations](#) notifying you that course evaluations are open. When you complete a course evaluation, you give me the tools to strengthen the course offerings and improve student learning. I cannot see individual responses; I receive a summary report that compiles responses across all students in the class. I do not receive a copy of the report until after grades have been submitted.

Academic Integrity

The University is founded on civility, honesty, and integrity. As a member of the community, you are expected to understand and follow the codes of conduct regarding academic integrity. [Academic integrity](#) is critical to the learning process. It is your responsibility to work in an honest fashion, upholding my expectations. The goal is to ensure that students learn in accordance with the University's academic integrity principles, regardless of whether instruction is in-person or remote. The use of any tools, including artificial intelligence-based large language models (e.g., ChatGPT), to produce whole or parts of assignments in this course is prohibited. Thank you for upholding your integrity and ensuring the University's tradition of academic excellence. A detailed description of academic integrity, including the University's policies and procedures, may be found in the [Graduate Academic Integrity Policy](#).

Accessibility Policies and Services

The University is committed to providing equal access to individuals with disabilities, in part, through [Accessibility Resources](#). If you have any disability which requires reasonable accommodations to enable you to participate in this course, please contact the Office of Accessibility Resources and also me during the first week of class. It is your responsibility to make a request for academic accommodation with Accessibility

Resources:

Accessibility Resources

University at Buffalo

60 Capen Hall (North Campus)

Buffalo, NY 14260

Phone: (716) 645-2608

Web: <https://www.buffalo.edu/studentlife/who-we-are/departments/accessibility.html>

Accessibility Resources will review appropriate arrangements for reasonable accommodations. If Accessibility Resources determines a request is reasonable, a memorandum describing the recommended academic adjustments or auxiliary aids will be issued. It is your responsibility to provide the memorandum to me and to arrange to meet with me regarding implementation of the recommendations. I may contact Accessibility Resources for consultation if there are any questions or concerns about a recommendation.

Student Success and Wellbeing Policies and Services

The University provides resources to support student learning and wellbeing. You can learn more about these programs and services by contacting:

Counseling Services

University at Buffalo

120 Richmond Quadrangle (North Campus)

Buffalo, NY 14261

Phone: (716) 645-2720

Web: <https://www.buffalo.edu/studentlife/who-we-are/departments/counseling.html>

Health Promotion

University at Buffalo

114 Student Union (North Campus)

Buffalo, NY 14260

Phone: (716) 645-2837

Web: <https://www.buffalo.edu/studentlife/who-we-are/departments/health-promotion.html>

Health Services

University at Buffalo

Michael Hall, 3435 Main Street (South Campus)

Buffalo, NY 14214

Phone: (716) 829-3316

Web: <https://www.buffalo.edu/studentlife/who-we-are/departments/health.html>

The University values and respects all members of the community. Harassment and discrimination are not tolerated. A detailed description of discrimination and harassment may be found in the [Discrimination and Harassment Policy](#).

The University provides appropriate accommodation for religious and cultural observances. Students who require a religious accommodation should make the request directly to me. A detailed description of religious accommodation may be found in the [Religious Accommodation and Expression Policy](#).

If you have experienced violence or harassment on the basis of sex or gender, the University has resources to help. For more information, please contact the Title IX Coordinator:

Office of Equity, Diversity and Inclusion

University at Buffalo

406 Capen Hall (North Campus)

Buffalo, New York 14260

Phone: (716) 645-2266

Web: <https://www.buffalo.edu/equity/obtaining-assistance/sex-discrimination-and-sexual-harassment/title-ix.html>

For confidential assistance, you may also contact a Crisis Services Campus Advocate at (716) 796-4399.

University at Buffalo faculty are mandated to report violence or harassment on the basis of sex or gender. This means that if you tell me about a situation involving violence or harassment, I will need to report it to the Office of Equity, Diversity and Inclusion. If you do not wish to have the University proceed with an investigation, your request will be honored unless the University's failure to act does not adequately mitigate the risk of harm to you or other members of the community. You also have the option of speaking with trained counselors who can maintain confidentiality. A full explanation of the resources and services available may be found in the [Options for Confidentially Disclosing Sexual Violence](#).