

1. THÔNG TIN CHUNG - GENERAL INFORMATION

Tên học phần	Tìm kiếm thông tin
Course name:	Information retrieval
Mã học phần	IT4863
Course ID:	IT4863
Khối lượng	3(3-1-0-6)
Credit:	<ul style="list-style-type: none">- Lý thuyết - Lecture: 45 hours- Bài tập/BTL -Exercises/Course project: 15 hours- Thí nghiệm - Experiments: 0 hours
Học phần tiên quyết	- MI1141: Đại số
Pre-requisite courses:	<i>MI1141: Algebra</i> <ul style="list-style-type: none">- MI2020: Xác suất thống kê <i>MI2020: Probability and Statistics</i>- IT3011:Cấu trúc dữ liệu và thuật toán hoặc IT3010: Cấu trúc dữ liệu và giải thuật <i>IT3011: Data Structures and Algorithms</i> or <i>IT3010: Data Structures and Algorithms</i>
Học phần học trước	Không có
Prior courses:	None
Học phần song hành	Không có
Co-requisite courses:	None

2. MÔ TẢ HỌC PHẦN - COURSE DESCRIPTION

Học phần này cung cấp một cái nhìn khái quát về lĩnh vực nghiên cứu tìm kiếm thông tin, kiến thức cơ bản quan trọng về máy tìm kiếm, mô hình tìm kiếm thông tin, cấu trúc chỉ mục, thuật toán thực hiện truy vấn. Đồng thời học phần cũng rèn luyện các kỹ năng vận dụng kiến thức toán học và lập trình để phân tích vấn đề và xây dựng mô hình, phát triển giải pháp công nghệ, triển khai dịch vụ, kiểm nghiệm giả thuyết và đánh giá giải pháp tìm kiếm thông tin.

Bên cạnh đó học phần còn phân tích mối liên hệ giữa giải pháp tìm kiếm thông tin và các yếu tố kinh tế, xã hội, rèn luyện các kỹ năng thuyết trình và làm việc nhóm.

This course provides an overview on the research field of information retrieval, briefly and precisely covers fundamental and important topics about search engines, information retrieval models, index structures and algorithms for query processing. During this course students will use prior knowledge and experiences in mathematics and programming to analyze problems and design models, develop technological solutions and deploy services, check hypotheses and evaluate information retrieval systems.

At the same time, in this course some economical and social aspects of search engines will be discussed. Students will also practice presentation and teamwork skills, and do the course project in teams.

3. MỤC TIÊU VÀ CHUẨN ĐẦU RA CỦA HỌC PHẦN - LEARNING OUTCOMES

Sinh viên hoàn thành học phần này có khả năng

After this course the student will obtain the followings

Mục tiêu /Course learning outcomes	Mô tả mục tiêu/Chuẩn đầu ra của học phần <i>Description of course learning outcomes</i>	CĐR được phân bổ cho HP/ Mức độ (I/T/U) Mapping to Program learning outcomes (I/T/U)
[1]	[2]	[3]
M1	Vận dụng kiến thức toán học để phân tích vấn đề và xây dựng mô hình giải pháp <i>Apply prior knowledge in mathematics to analyze practical problems and develop solutions.</i>	1.1.1; 1.1.2; 1.1.4
M1.1	Vận dụng kiến thức đại số để phân tích và xây dựng mô hình tìm kiếm thông tin. <i>Apply algebra to analyze and develop information retrieval models.</i>	[1.1.1] (UT)
M1.2	Vận dụng kiến thức xác suất thống kê để phân tích và xây dựng mô hình tìm kiếm thông tin. <i>Apply probability and mathematical statistics to analyze and develop information retrieval models.</i>	[1.1.2] (UT)
M1.3	Áp dụng kiến thức biểu diễn dữ liệu để phân tích các phương pháp nén và tối ưu hóa lưu trữ. <i>Exploit the knowledge of data encoding in development of compression methods and data storage optimizations.</i>	[1.1.4] (UT)
M2	Vận dụng kiến thức cốt lõi ngành để triển khai các mô hình toán học, phát triển các giải pháp công nghệ <i>Apply fundamental knowledge to implement mathematical models, and develop technological products.</i>	1.2.1; 1.2.2
M2.1	Vận dụng kiến thức cấu trúc dữ liệu và giải thuật để xây dựng các cấu trúc chỉ mục. <i>Exploit data structures and algorithms knowledge in index structures developments.</i>	[1.2.1] (UT)

Mục tiêu /Course learning outcomes	Mô tả mục tiêu/Chuẩn đầu ra của học phần <i>Description of course learning outcomes</i>	CĐR được phân bổ cho HP/ Mức độ (I/T/U) Mapping to Program learning outcomes (I/T/U)
M2.2	<p>Áp dụng kiến thức về hệ thống máy tính, các đặc điểm phần cứng để phát triển giải pháp tính toán tối ưu.</p> <p><i>Apply knowledge about computing systems to optimize computational solutions.</i></p>	[1.2.2] (U)
M3	<p>Vận dụng kiến thức chuyên ngành để triển khai dịch vụ tìm kiếm</p> <p><i>Apply specialized knowledge to deploy information retrieval services</i></p>	1.3.2; 1.3.3; 1.3.4; 1.3.5
M3.1	<p>Nắm vững nguyên lý vận hành và kiến trúc máy tìm kiếm và có khả năng vận dụng kiến thức về hệ thống thông tin để thiết kế và triển khai dịch vụ tìm kiếm đơn giản ở quy mô nhỏ.</p> <p><i>Mastering the working principles and architectures of search engines, and being able to apply specialized knowledge in information systems to design and deploy simple information retrieval services on a small scale.</i></p>	[1.3.2] (T)
M3.2	<p>Nắm vững nguyên lý và có khả năng phân tích vấn đề tìm kiếm thông tin ở quy mô lớn.</p> <p><i>Mastering the working principles and being able to analyze information retrieval problems on a large scale.</i></p>	[1.3.3] (T)
M3.3	<p>Có khả năng sử dụng các nền tảng mã nguồn mở chuyên dụng và áp dụng các kỹ thuật gợi ý trên dữ liệu văn bản trong xây dựng hệ thống tìm kiếm thông tin.</p> <p><i>Being able to use specialized open sources and apply recommendation methods on textual data for information retrieval system developments.</i></p>	[1.3.4] (T) [1.3.5] (T)
M4	<p>Áp dụng các kỹ năng làm việc nhóm trong phát triển hệ thống theo nhóm</p> <p><i>Apply teamwork skills in system development by team.</i></p>	3.1.1; 3.1.2; 3.1.3, 3.1.4
M4.1	<p>Có khả năng thành lập nhóm và tham gia nhóm phù hợp với kỹ năng để phát triển các giải pháp công nghệ.</p> <p><i>Being able to organize a team or join a suitable team by skills to develop technological solutions.</i></p>	[3.1.1] (U)
M4.2	<p>Có khả năng tổ chức hoạt động nhóm và phân chia công việc trong nhóm.</p> <p><i>Being able organize teamworks and breakdown a big task to smaller tasks for team members.</i></p>	[3.1.2] (U)

Mục tiêu /Course learning outcomes	Mô tả mục tiêu/Chuẩn đầu ra của học phần <i>Description of course learning outcomes</i>	CĐR được phân bổ cho HP/ Mức độ (I/T/U) Mapping to Program learning outcomes (I/T/U)
M4.3	Có khả năng quản lý tiến độ dự án. <i>Being able to manage project progress.</i>	[3.1.3] (U)
M4.4	Có khả năng phối hợp hiệu quả với các thành viên khác để cùng giải quyết công việc. <i>Can effectively cooperate with other members to solve tasks together.</i>	[3.1.4] (U)
M5	Nhận thức về tác động của công nghệ đối với xã hội và các quy định đối với công nghệ <i>Awareness of the impacts of information technology solutions on society and regulations of technologies.</i>	4.1.2; 4.1.3
M5.1	Nhận thức về những tác động của công nghệ tìm kiếm thông tin với kinh tế, xã hội. <i>Awareness of the social and economical impacts of information retrieval technologies.</i>	[4.1.2] (I)
M5.2	Nhận thức về những quy định đối với công nghệ tìm kiếm thông tin. <i>Awareness of regulations of information retrieval technologies.</i>	[4.1.3] (I)

4. TÀI LIỆU HỌC TẬP - COURSE MATERIALS

Giáo trình – Textbook

Sách tham khảo - Reference book

- [1] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- [2] W. Bruce Croft, Donald Metzler and Trevor Strohman (2010). *Search Engines: Information retrieval in practice*. Addison-Wesley.

Online open courseware

...

5. CÁCH ĐÁNH GIÁ HỌC PHẦN - EVALUATION

Điểm thành phần Module	Phương pháp đánh giá cụ thể Evaluation method	Mô tả Detail	CĐR được đánh giá Output	Tỷ trọng Percent
[1]	[2]	[3]	[4]	[5]
A1. Điểm quá trình Mid-term (*)	Đánh giá quá trình Progress evaluation	Bài tập lớn	M3.1÷M3.3; M4.1÷M4.4	30%

		Course project		
A2. Điểm cuối kỳ Final term	Thi cuối kỳ Final exam	Thi viết Written exam	M1.1÷M1.3, M2.1, M2.2, M3.1÷M3.3	70%

* Điểm quá trình sẽ được điều chỉnh bằng cách cộng thêm điểm chuyên cần. Điểm chuyên cần có giá trị từ -2 đến +1, theo Quy chế Đào tạo đại học hệ chính quy của Trường ĐH Bách khoa Hà Nội.

The progress score can be adjusted by diligence evaluation. The adjusted value should be in the range [-2, +1], according to the policy of Hanoi University of Science and Technology.

6. KẾ HOẠCH GIẢNG DẠY – SCHEDULE

Tuần <i>Week</i>	Nội dung <i>Content</i>	CDR học phần <i>Learning outcomes</i>	Hoạt động dạy và học <i>Activities</i>	Bài đánh giá <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
1	<p>Tổng quan học phần</p> <p>Chương 1. Vấn đề tìm kiếm thông tin và giải pháp Boolean</p> <p>1.1. Vấn đề tìm kiếm thông tin</p> <p>1.2. Mô hình Boolean</p> <p>1.3. Một cách đánh chỉ mục trong RAM</p> <p>1.4. Xử lý truy vấn Boolean</p> <p>1.5. Mô hình Boolean mở rộng và xếp hạng</p> <p>Thông tin bài tập lớn</p> <p><i>Course overview</i></p> <p><i>Chapter 1. The Information Retrieval Problem and the Boolean solution</i></p> <p><i>1.1. An information retrieval problem</i></p> <p><i>1.2. Boolean model</i></p> <p><i>1.3. One indexing algorithm in RAM</i></p> <p><i>1.4. Boolean queries processing</i></p> <p><i>1.5. Extended Boolean models and ranked retrieval</i></p> <p><i>Course project introduction</i></p>	M1.1, M2.1	<p>Giảng bài, thảo luận.</p> <p><i>Teaching, discussion.</i></p> <p>Sinh viên đăng ký nhóm bài tập lớn.</p> <p><i>Students sign up for the course project teams.</i></p>	A2
2	<p>Chương 2. Bộ từ vựng, danh sách định vị và phương pháp mềm hóa tìm kiếm</p> <p>2.1. Tách nội dung văn bản từ tài liệu</p>	M2.1, M3.3	Giảng bài, thảo luận.	A1, A2

Tuần <i>Week</i>	Nội dung <i>Content</i>	CĐR học phần <i>Learning outcomes</i>	Hoạt động dạy và học <i>Activities</i>	Bài đánh giá <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
	<p>2.2. Tổng hợp bộ từ vựng</p> <p>2.3. Tăng tốc xử lý danh sách định vị bằng con trỏ ngắt quãng</p> <p>2.4. Thẻ định vị có vị trí và truy vấn nguyên câu</p> <p>2.5. Cấu trúc tra cứu cho bộ từ vựng</p> <p>2.6. Truy vấn với ký tự đại diện</p> <p>2.7. So khớp mềm</p> <p>2.8. Gợi ý từ khóa</p> <p><i>Chapter 2. Vocabulary, posting lists and keywords soft matching</i></p> <p><i>2.1. Extract text from documents</i></p> <p><i>2.2. Vocabulary creation</i></p> <p><i>2.3. Speed up posting lists processing with skip pointers</i></p> <p><i>2.4. Positional index and phrase query</i></p> <p><i>2.5. Search structure for vocabulary</i></p> <p><i>2.6. Query with wildcard characters</i></p> <p><i>2.7. Soft matching</i></p> <p><i>2.8. Keyword suggestion</i></p>		Teaching, discussion.	
3	<p>Chương 3. Xây dựng và nén chỉ mục ngược</p> <p>3.1. Căn bản phần cứng</p> <p>3.2. Giải thuật BSBI</p> <p>3.3. Giải thuật SPIMI</p> <p>3.4. Giải thuật dựa trên MapReduce</p> <p>3.5. Chỉ mục ngược thay đổi động</p> <p>3.6. Các loại chỉ mục khác</p> <p>3.7. Các đặc điểm thống kê từ vựng trong tìm kiếm thông tin</p> <p>3.8. Nén chỉ mục ngược</p> <p><i>Chapter 3. Index construction and compression algorithms</i></p> <p><i>3.1. Hardware basics</i></p> <p><i>3.2. Blocked Sort-based Indexing (BSBI)</i></p> <p><i>3.3. Single-pass In-memory Indexing (SPIMI)</i></p>	M1.3, M2.1, M2.2, M3.2	Giảng bài, thảo luận. <i>Teaching, discussion.</i>	A2

Tuần <i>Week</i>	Nội dung <i>Content</i>	CĐR học phần <i>Learning outcomes</i>	Hoạt động dạy và học <i>Activities</i>	Bài đánh giá <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
	3.4. <i>MapReduce indexing</i> 3.5. <i>Dynamic index</i> 3.6. <i>Other indexes</i> 3.7. <i>Statistical properties of terms</i> 3.8. <i>Index compression</i>			
4	Chương 4. Tính điểm, trọng số từ và giải pháp không gian vec-tơ 4.1. Chỉ mục tham số và chỉ mục miền 4.2. Tần suất và đánh giá trọng số từ 4.3. Thiết lập không gian vec-tơ và tính điểm 4.4. Các biến thể của trọng số tf-idf 4.5. Xử lý truy vấn trong mô hình không gian vec-tơ 4.6. Các kỹ thuật cắt tỉa nhánh tìm kiếm 4.7. Các thành phần của máy tìm kiếm 4.8. Tính điểm trong không gian vec-tơ và mối liên hệ với các dạng truy vấn <i>Chapter 4. Scoring, term weighting and vector space model for information retrieval</i> 4.1. <i>Parametric index and zone indexes</i> 4.2. <i>Term frequency and term weighting</i> 4.3. <i>Vector space construction and scoring</i> 4.4. <i>Variants of tf-idf weighting</i> 4.5. <i>Query processing in vector space model</i> 4.6. <i>Search branches pruning</i> 4.7. <i>Components of a search engine</i> 4.8. <i>Scoring in the vector space and relationships with other query types</i>	M1.1, M2.1	Giảng bài, thảo luận. <i>Teaching, discussion.</i>	A2
5	Chương 5. Đánh giá trong tìm kiếm thông tin 5.1. Vấn đề đánh giá hệ thống tìm kiếm 5.2. Một số bộ dữ liệu kiểm thử phổ biến 5.3. Đánh giá kết quả dạng tập hợp 5.4. Đánh giá kết quả dạng danh sách	M2.1, M3.1	Giảng bài, thảo luận. <i>Teaching, discussion.</i>	A1, A2

Tuần <i>Week</i>	Nội dung <i>Content</i>	CĐR học phần <i>Learning outcomes</i>	Hoạt động dạy và học <i>Activities</i>	Bài đánh giá <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
	<p>5.5. Đánh giá tính phù hợp và vấn đề xây dựng bộ dữ liệu kiểm thử</p> <p>5.6. Mối liên hệ giữa chất lượng hệ thống và tính hữu ích từ góc nhìn người dùng</p> <p>5.7. Trích đoạn kết quả tìm kiếm</p> <p><i>Chapter 5. Evaluations in information retrieval</i></p> <p><i>5.1. The problem of evaluation</i></p> <p><i>5.2. Well-known test collections</i></p> <p><i>5.3. Evaluation of result sets</i></p> <p><i>5.4. Evaluation of ranked results lists</i></p> <p><i>5.5. Relevance assignment and test collection construction</i></p> <p><i>5.6. System quality and user utility</i></p> <p><i>5.7. Search results snippets</i></p>			
6	<p>Thảo luận và kiểm tra tiến độ bài tập lớn</p> <p>Máy tìm kiếm</p> <p>Các nền tảng mã nguồn mở</p> <p>Xây dựng phân hệ tìm kiếm</p> <p><i>Discussion and coursework checkpoint</i></p> <p><i>Search engine</i></p> <p><i>Popular open source projects</i></p> <p><i>Searcher development</i></p>	<p>M3.3</p> <p>M4.1÷M4.4</p>	<p>Làm việc nhóm, thuyết trình, thảo luận.</p> <p><i>Teamwork, presentation, discussion.</i></p>	A1
7	<p>Chương 6. Các kỹ thuật nâng cao độ chính xác và độ đầy đủ</p> <p>6.1. Phản hồi phù hợp và giải thuật Rocchio</p> <p>6.2. Các hình thức phản hồi</p> <p>6.3. Phản hồi phù hợp trong môi trường Web</p> <p>6.4. Các kỹ thuật viết lại câu truy vấn</p> <p>6.5. Từ điển đồng nghĩa và các kỹ thuật tự động tổng hợp từ điển đồng nghĩa</p> <p>6.6. Các đánh giá</p>	<p>M1.1, M3.3</p>	<p>Giảng bài, thảo luận.</p> <p><i>Teaching, discussion</i></p>	A1, A2

Tuần <i>Week</i>	Nội dung <i>Content</i>	CĐR học phần <i>Learning outcomes</i>	Hoạt động dạy và học <i>Activities</i>	Bài đánh giá <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
	<p><i>Chapter 6. Techniques to improve precision and recall</i></p> <p><i>6.1. Relevance feedback and the Rocchio algorithm</i></p> <p><i>6.2. Feedback types</i></p> <p><i>6.3. Relevance feedback in the Web</i></p> <p><i>6.4. Query reformulation</i></p> <p><i>6.5. Thesaurus and thesaurus construction</i></p> <p><i>6.6. Evaluation</i></p>			
8	<p>Chương 7. Tìm kiếm thông tin theo xác suất và các mô hình ngôn ngữ</p> <p>7.1. Căn bản lý thuyết xác suất</p> <p>7.2. Nguyên lý xếp hạng xác suất</p> <p>7.3. Mô hình nhị phân độc lập</p> <p>7.4. Mô hình Okapi BM25 và các mô hình xác suất khác</p> <p>7.5. Các mô hình ngôn ngữ</p> <p>7.6. Ước lượng xác suất dựa trên mô hình ngôn ngữ</p> <p>7.7. So sánh mô hình hóa ngôn ngữ với các cách tiếp cận khác</p> <p>7.8. Các phương pháp mô hình hóa ngôn ngữ mở rộng</p> <p><i>Chapter 7. Probabilistic information retrieval and language models</i></p> <p><i>7.1. A brief review of probability theory</i></p> <p><i>7.2. Probability Ranking Principle</i></p> <p><i>7.3. Binary independent model (BIM)</i></p> <p><i>7.4. Okapi BM25 and other probabilistic models</i></p> <p><i>7.5. Language models</i></p> <p><i>7.6. Probability estimation based on language modeling</i></p> <p><i>7.7. A comparison of language modeling approach with other approaches</i></p> <p><i>7.8. Extended language modeling approaches</i></p>	M1.2	Giảng bài, thảo luận <i>Teaching, Discussion</i>	A2

Tuần <i>Week</i>	Nội dung <i>Content</i>	CĐR học phần <i>Learning</i> <i>outcomes</i>	Hoạt động dạy và học <i>Activities</i>	Bài đánh giá <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
9	<p>Chương 8. Phân lớp và phân cụm văn bản</p> <p>8.1. Bài toán phân lớp văn bản và các ứng dụng trong tìm kiếm thông tin</p> <p>8.2. Các phương pháp phân lớp văn bản</p> <p>8.3. Các phương pháp trích chọn đặc trưng</p> <p>8.4. Bài toán phân cụm văn bản trong tìm kiếm thông tin</p> <p>8.5. Các phương pháp phân cụm văn bản</p> <p>8.6. Các đánh giá</p> <p><i>Chapter 8. Text classification and clustering</i></p> <p><i>8.1. The text classification problem and applications in information retrieval</i></p> <p><i>8.2. Text classification methods</i></p> <p><i>8.3. Features selection methods</i></p> <p><i>8.4. The text clustering problem</i></p> <p><i>8.5. Text clustering methods</i></p> <p><i>8.6. Evaluations</i></p>	M1.2, M3.3	Giảng bài, thảo luận <i>Teaching, Discussion</i>	A1, A2
10	<p>Thảo luận và kiểm tra tiến độ bài tập lớn</p> <p>Các nền tảng mã nguồn mở</p> <p>Các định dạng tài liệu</p> <p>Xây dựng phân hệ thu thập</p> <p><i>Discussion and coursework checkpoint</i></p> <p><i>Popular open source projects</i></p> <p><i>Document formats and parsing</i></p> <p><i>Crawler development</i></p>	M3.3 M4.1÷M4.4	Làm việc nhóm, thuyết trình, thảo luận. <i>Teamwork, presentation, discussion.</i>	A1
11	<p>Chương 9. Căn bản tìm kiếm thông tin trên Web</p> <p>9.1. Kiến thức cơ bản về Web</p> <p>9.2. Biểu diễn đồ thị của Web</p> <p>9.3. Nội dung rác trong môi trường Web</p> <p>9.4. Quảng cáo như mô hình kinh tế</p> <p>9.5. Kinh nghiệm người dùng</p>	M3.1, M3.2, M5.1, M5.2	Giảng bài, thảo luận. <i>Teaching, discussion.</i>	A1, A2

Tuần <i>Week</i>	Nội dung <i>Content</i>	CĐR học phần <i>Learning outcomes</i>	Hoạt động dạy và học <i>Activities</i>	Bài đánh giá <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
	9.6. Đặc điểm nhu cầu thông tin 9.7. Kích thước chỉ mục và ước lượng 9.8. Trùng lặp gần và giải pháp <i>Chapter 9. Web search basics</i> <i>9.1. Web basic characteristics</i> <i>9.2. Graph representation for Web documents</i> <i>9.3. Spam in the Web</i> <i>9.4. Advertising as the economic model</i> <i>9.5. User experiences</i> <i>9.6. Characteristics of information needs</i> <i>9.7. Index size and estimations</i> <i>9.8. Near duplicates and solution</i>			
12	Chương 10. Thu thập dữ liệu Web và chỉ mục ngược quy mô lớn 10.1. Các yêu cầu hệ thống 10.2. Tổng quan hệ thống thu thập dữ liệu 10.3. Phân giải DNS 10.4. Hàng đợi URL 10.5. Máy chủ liên kết 10.6. Chia nhỏ và phân tán chỉ mục ngược 10.7. Lưu trữ tài liệu quy mô lớn <i>Chapter 10. Web crawling and large scale inverted index</i> <i>10.1. System requirements</i> <i>10.2. Crawling system architecture</i> <i>10.3. DNS resolution</i> <i>10.4. URL queue</i> <i>10.5. Link server</i> <i>10.6. Split and distribute inverted index</i> <i>10.7. Document storage at large scale</i>	M3.1, M3.2	Giảng bài, thảo luận <i>Teaching, discussion.</i>	A1, A2
13	Chương 11. Phân tích liên kết 11.1. Văn bản liên kết và biểu diễn tài liệu Web 11.2. PageRank 11.3. Hub và Authorities	M1.1, M1.2, M2.1	Giảng bài, thảo luận. <i>Teaching, discussion.</i>	A2

Tuần <i>Week</i>	Nội dung <i>Content</i>	CĐR học <i>phần</i> <i>Learning</i> <i>outcomes</i>	Hoạt động <i>dạy và học</i> <i>Activities</i>	Bài đánh <i>giá</i> <i>Evaluation</i>
[1]	[2]	[3]	[4]	[5]
	Chương 12. Vượt qua giới hạn tập từ 12.1. Tìm kiếm thông tin theo nội dung 12.2. Hình ảnh 12.3. Âm thanh 12.4. Các xu hướng tiêu biểu khác <i>Chapter 11. Link analysis</i> <i>11.1. Anchor text and Web documents representation</i> <i>11.2. PageRank</i> <i>11.3. Hub and Authorities</i> <i>Chapter 12. Beyond the Bag of Word</i> <i>12.1. Semantic search</i> <i>12.2. Image</i> <i>12.3. Sound</i> <i>12.4. Other potential directions</i>			
14	Bảo vệ bài tập lớn <i>Coursework defense</i>	M4.1÷ M4.4	Làm việc nhóm, thuyết trình. <i>Teamwork, presentation</i>	A1
15	Tổng kết và ôn tập. <i>Course summary and review.</i>	*	Discussion, exercises	

7. QUY ĐỊNH CỦA HỌC PHẦN - COURSE REQUIREMENT

8. NGÀY PHÊ DUYỆT - DATE:

Chủ tịch hội đồng
Committee chair

Nhóm xây dựng đề cương
Syllabus development team

Nguyễn Bá Ngọc,
Nguyễn Thị Kim Anh,
Lê Thanh Hương,
Nguyễn Kiên Hiếu,
Nguyễn Thị Oanh,
Đỗ Bá Lâm

9. QUÁ TRÌNH CẬP NHẬT - DOCUMENT VERSION INFORMATION

STT No	Nội dung điều chỉnh Content of the update	Ngày tháng được phê duyet Date accepted	Áp dụng từ kỳ/ khóa Effective from	Ghi chú Note
1	- Cập nhật định dạng tài liệu, bổ xung nội dung tiếng Anh.			
2			