Professional Inquiry

Welcome to the module on **Professional Inquiry** (PI), a 5-ECTS course on the Information Studies degree. The course is taught by:

- Birger Larsen (course responsible)
- Florian Meier
- Lone Dirckinck-Holmfeld

The PI module is an intensive course run over a few weeks at the beginning of the term. The course comprises the development and phrasing of empirical inquiry for the purpose of enabling students to formulate research questions and scientific problems within the field of Information Studies. This will form the basis of the problem based project work and inquiries to be carried out during the course of the Information Studies study programme.

Most of the **readings** are either freely available on the web or can be accessed through AUB - The University Library. To find literature in AUB, go to their homepage (http://www.en.aub.aau.dk), paste the title of the publication in the search box and press 'search'. Once you find the publication on the list, click 'Online adgang' and follow the instructions. If you are on-campus, then usually no password is needed. If you are off-campus, use your AAU log-in credentials.

Announcements

Messages

Active Participation

Hidden from students

Group Formation Spreadsheet

Hidden from students

Use this spreadsheet to note the members of each group and the case you wish to work with.

Students profiles

Hidden from students

Lecture 1: Doing research [BL, 4 hrs]

In this lecture we will take a closer look at:

▶ Writing Academic English

Readings

Reference	*Mandatory*	*Additional*	*Upload*
Hämäläinen, W. (2006): *Scientific Writing for Computer Science Students*. University of Joensuu, 123 p. **Chapters 1 and 4.** (http://www.cs.joensuu.fi/pages/whamalai/sciwri/sciwri.pdf)	64		
Zobel, J. (2014): *Writing for Computer Science*. 3rd edition. Springer Science Verlag, 284 p. (Full book can be downloaded via AUB). **Chapters 1 and 8.**	17		

Upload your BA/BSc paper (in English) here - for the peer review exercise

Hidden from students

7sem-PI-E2017 L1 Intro+AcademicEnglish

Hidden from students

Upload your checklists here

Hidden from students

In case you didn't get to hand me your checklists on Monday, make sure to upload them here in this forum no later than Tuesday Sept. 5th @ 23:59.

Also, remember to meet up with the other 1-2 people that also read your paper and collaboratively go through the checklists (will be scanned in Moodle) and identify the main issues with spelling, grammar, and style. Have some conclusions ready for **10.15 on Wednesday**.

Lecture 2: Writing reports [BL, 4 hrs]

In this lecture we will take a closer look at:

▶ Report writing

▶ How to write (and read) research papers

Readings

Reference	*Mandatory*	*Additional*	*Upload*
Zobel, J. (2014): *Writing for Computer Science*. 3rd edition. Springer Science Verlag, 284 p. (Full book can be downloaded via AUB). **Chapters 5, 6, 7, 11 and 13.**	100		

h

7sem-PI-E2017 ReportWriting

Hidden from students

Lecture 3: Information Literacy [CW, 2 hrs]

Charlotte Wind (Librarian @ AUB)

In this lecture we will focus on Information Literacy, the most important skill to learn at a higher education no matter what field you are in. Information Literacy is the adoption of appropriate information behavior identify, through whatever channel or medium, information well fitted to information needs, leading to wise and ethical use of information in society. IL is also a transformational process in which the learner needs to find, understand, evaluate and use information in various forms. Strong analytical, critical thinking and problem solving skills are key attributes of valuable employees in the HCI sectors.

Readings

Reference	*Mandatory*	*Additional*	*Upload*
Zobel, J. (2014): *Writing for Computer Science*. 3rd edition. Springer Science Verlag, 284 p. (Full book can be downloaded via AUB). **Chapter 3.**	25		
Forsyth Tech (2015): *Information Literacy in Today's Workplace: The Employer's Perspective Day 1*. Available at https://www.youtube.com/watch?v=PgIZ0fNrkn4		-	
Forsyth Tech (2015): *Information Literacy in Today's Workplace: The Employer's Perspective Day 2*. Available at https://www.youtube.com/watch?v=7jdMw74-rBU		-	

Information literacy by Charlotte Wind (AUB)

Hidden from students

Lecture 4: Finding a research topic and planning research [BL, 2 hrs]

Birger Larsen

In this lecture we will focus on some of the very early steps that you go through in order to find a research topic, a.k.a. solving problems. The process of defining a constrained research project leading to the solution of a complex problem related to humans and information systems is important to understand and will be used throughout the degree.

Reference	*Mandatory*	*Additional*	*Upload*
Bordens , K. S. & Abbott, B. B. "Research Design and Methods – A Process Approach". McGraw Hill. 8th edition or newer. **Chapter 3 (Getting ideas for Research) and chapter 4 (Choosing a Research Design).**	70		
Zobel, J. (2014): *Writing for Computer Science*. 3rd edition. Springer Science Verlag, 284 p. (Full book can be downloaded via AUB). **Chapters 2 and 4.**	25		

7sem-PI-E2017 L3a FindingAResearchTopic

Hidden from students

Lecture 5: Literature reviews [FM, 2 hrs]

An important part of any project report is writing a literature review—an objective, thorough summary and critical analysis of the relevant available research and non-research literature on your project topic. In this lecture you will learn the basics of how to conduct a literature review: (1) searching the literature, (2) reading & analyzing the literature, and (3) writing the review. In addition, you will be introduced to useful websites, services and tools for searching, collecting, and managing the literature used in your project.

The links to the required and supplementary reading should work when you're connected to the AAU wifi network, but otherwise you should be able to find PDF files of all of them by searching for the article titles in Google Scholar.

Readings

Reference	*Mandatory*	*Additional*	*Upload*
Cronin, P., Ryan, F., & Coughlan, M. (2008): Undertaking A Literature Review: A Step-By-Step Approach. *British Journal of Nursing*, 17(1), 38–43 [WWW link]	6		
Rowley, J., & Slack, F. (2013). Conducting a Literature Review. *Management Research News*, 27(6), 31–39 [WWW link]	8		
Randolph, J.J. (2009). A Guide to Writing the Dissertation Literature Review. *Practical Assessment, Research & Evaluation*, 14(13), 1–13 [WWW link]		13	
Torraco, R.J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, 4(3), 356–367 [WWW link]		12	
Webster, J., & Watson, R.T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), 13–23 [link]		11	

Slides ('Literature search & review')

Hidden from students

Examples

- Example 1.pdf
- Example 2.pdf
- Example 3.pdf

ht

Download folder

Hidden from students

Lecture 6: Introduction to PBL [LDH, 4 hrs]

Readings

Reference	*Mandatory (# pages)*	*Additional (# pages)*	*Upload*

Lecture 7: Cognitive human modeling and Basic HCI theory [BL, 2 hrs]

Introduction to basic HCI theory including humans, computers and their interaction, as well as how this connects with and forms the basis for empirical research.

Readings

Reference	*Mandatory (# pages)*	*Additional (# pages)*	*Upload*
Dix, A., Finlay, J, Abowd, G. D. & Beale, R. (2004): Human Computer Interaction. 3rd ed. Harlow: Pearson Education, xxv, 834 p. **Introduction + chapters 1 (The Human), 2 (The computer) and 3 (The interaction).**	163		

7sem-PI-E2017 L4 HCItheory

Hidden from students

Lecture 8: Theory in ICT [BL, 4 hrs]

In this lecture we will discuss what constitutes theory and epistemology in ICT research. The field as such is quite practice-driven and often it is hard to find any explicit or implicit use of theory. How can we then work academically in our research and papers?

Group formation will also be finalised during this day.

Readings

Reference	*Mandatory (# pages)*	*Additional (# pages)*	*Upload*
Carroll, J. M. (2014): Human Computer Interaction - brief intro. In: Soegaard, M. and Dam, R. F. (eds.). *The Encyclopedia of Human-Computer Interaction, 2nd Ed.*. Aarhus, Denmark: The Interaction Design Foundation. (WWW link)	27		
Charters (2003): The Use of Think-aloud Methods in Qualitative Research - An Introduction to Think-aloud Methods. *Brock Education*, 12(2). [AUB link]	15		

7sem-PI-E2017-L5 TheoryEpistemologyICT

Hidden from students

Slide upload

User Practice, User Analysis and Pilot Studies

Announcements

Messages

Tricider feedback page

Administrivia

This section provides you with different documents that should make successfully completing the 7th semester easier for both you and your supervisors.

- Overview of all deliverables — Link to PI overview of deliverables with an overview of all the deliverables you are expected to complete throughout the semester.
- Deliverable details — Link to PI detailed descriptions of the deliverables along with indepth instructions.
- Deliverable completion spreadsheet — This read-only spreadsheet will allow you to keep track of which semester deliverables you have already completed and which ones are still open. If you believe you have spotted an error in the spreadsheet, please contact Toine.
- Deliverable 1: Research design (template) — This Word document is a template for your first deliverable: presenting your research design. If you have any questions about how to use the template in your project, please contact your supervisor.
- HCI report template — This Word document is a template for how to write a good report at Informationsvidenskab. We expect you to use this an example to structure your Bachelor thesis. While you are free to alter the layout of the document, you should consult with your supervisor before you deviate from the structure and contents it proposes. In general, if you have any questions about how to use the template in your project, please contact your supervisor.
- Reference list bookkeeping template — According to the Studieordning, your Bachelor thesis must have at least 1,000 pages of academic literature, to be approved by your supervisor. This list of supervisor-approved literature must be handed in at least 2 weeks before the project deadline. If you fail to do so, you risk having your Bachelor thesis rejected.
  To make this cumbersome process as easy as possible on students and supervisors alike, we propose using a Google Spreadsheet to keep track of your literature and page counts. Literature is entered on the first worksheet; the second worksheet automatically calculates how many pages of approved academic literature you already have. The link below is a template spreadsheet; please do not edit it directly! Instead, select *File* > *Make a copy...* and copy it to your one of your own Google Drive accounts. Then share it with your supervisor so (s)he) can approve literature throughout the semester. Please discuss this with your supervisor if you have any questions.
- Client communication guide — This document contains some helpful tips on how to set up and maintain a productive working relationship with your case client.

Overview of deliverables

Deliverable details and instructions

Deliverable Completion Spreadsheet

Deliverable 3a - Research design (template)

IS report template

Client communication guide

Reference list bookkeeping template

The 5 cases - slideshow

Case descriptions

- Case 1 - Learning Platforms in Schools.pdf
- Case 2 - Google Explore.pdf
- Case 3 - Evaluating Siri.pdf
- Case 4 - Capturing Research Impact.pdf
- Case 5 - Exploring Citation Contexts.pdf

Download folder

Hidden from students

Lecture 0: Semester introduction [BL, 2 hrs]

The first lecture of the module will focus on the requirements of the degree and the course, and what it means to conduct and write a PBL project.

Required reading

- Pickard, A. J. (2013). "Chapter 1" in: Research Methods in Information. 2nd Ed. London: Facet Publishing.
  [35 pages / link / short description]
- Rogers, Y. (2004) New Theoretical approaches for Human-Computer Interaction. Annual Review of Information, Science and Technology, 38, 87-143.
  [56 pages / link / short description]
- Ioannidis, J. P. A. “Why Most Published Research Findings Are False”, PLoS Medicine, 2005, 8(2)
  [6 pages / link / short description]

h

7sem-UP-E2017 SemesterIntro

Hidden from students

7sem-IS-E2017 IntroAtStudyStart casesFinal

Hidden from students

Lecture 1: Research design Part I [FM, 2 hrs]

In this two-part lecture we will focus on the basics of research design: how do we design an experiment in such a way that we can draw reliable, valid, and meaningful conclusions from this. Topics include research questions and hypotheses, types of variables, levels of measurement, and experimental structure, and recruitment of participants. Teaching will consist of a combination of traditional lecturing and (group) exercises.

The first part will focus on introducing/repeating some basic statistical concepts that are needed for talking about research and research design:

- Hypotheses
- Types of variables
- Levels of measurement

Reading:

- Lazar, J., Feng, J.H., and Hochheiser, H. (2010). "Chapter 2: Experimental research". In: *Research Methods in Human-Computer Interaction*, Hoboken, NJ: Wiley (pp. 19-37). [19 pages / Good introductory chapter describing the basics of doing experimental research.]

Lecture 2: Research design Part II [FM, 4 hrs]

In this two-part lecture we will focus on the basics of research design: how do we design an experiment in such a way that we can draw reliable, valid, and meaningful conclusions from this. Topics include research questions and hypotheses, types of variables, levels of measurement, and experimental structure, and recruitment of participants. Teaching will consist of a combination of traditional lecturing and (group) exercises.

The second part will introduce the intersection of research questions, research design and methods for data collection.

Required reading

- Bryman, A. (2012). "Chapter 3: Research designs". In: *Social Research Methods* (4th ed.), New York, NY: Oxford University Press (pp. 44-78). [35 pages / Good overview chapter about the different types of research designs we have at our disposal.]
- Lazar, J., Feng, J.H., and Hochheiser, H. (2010). "Chapter 3: Experimental design". In: *Research Methods in Human-Computer Interaction*, Hoboken, NJ: Wiley (pp. 41-65). [25 pages / Good introductory chapter describing the basics of designing experiments.]

Additional reading

- more to come ...

Slides ('Research designs')

Hidden from students

Literature

Bryman (2012), chapter 3.pdf

Lazar (2010), chapter 2.pdf

Lazar (2010), chapter 3.pdf

Download folder

Lecture 3a: Data collection - Stakeholder analysis [LDH, 1 hrs]

This lecture (along with exercises) introduces stakeholder analysis as a tool for gathering requirements for software. Please note that the additional reading for this lecture is required reading for the students coming from the IV specialisation on the 6th semester, as they have had most of the material listed above.

Required reading

- Schmeer, K. "Stakeholder Analysis Guidelines". In: *Policy Toolkit for Strengthening Health Sector Reform,* LAC-HSR.
  [48 pages / link / These guidelines gives practical advise on how to carry out stakeholder analysis.]
- Dix et al. "Socio-Organizational Issues and Stakeholder Requirements", chap. 13 in "*Human Computer Interaction*"
  [24 pages / **uploaded below** / This text book chapter presents the basic of stakeholder analysis.]
- Pan, G. S. C. "Information systems project abandonment: a stakeholder analysis", *International Journal of Information Management*, 25, 2005, 173-184.
  [11 pages / AUB link / This article gives an example of a stakeholder analysis in an actual project.]

Additional reading

- Majchrzak A.; Cherbakov, L. and Ives, B. "Harnessing the Power of the Crowds with Corporate Social Networking Tools: How IBM Does It. MIS Quarterly Executive, 8(2), 2009.
  [7 pages / link / short description]

ht

- McAfee, A. P. “Enterprise 2.0: The Dawn of Emergent Collaboration”. MIT Sloan Management Review, 2006, 47(3).
  [10 pages / AUB link / short description]
- Puloudi, A. “Aspects of the Stakeholder Concept and their Implications for Information Systems Development”. Proceedings of the 32nd Hawaii International Conference on System Sciences, 1999.
  [17 pages / AUB link / short description]
- Sharp, H, Finkelstein, A. and Galal, G. 1999. “Stakeholder Identification in the Requirements Engineering Process”. In: *Proceedings of the Tenth International Workshop on Database and Expert Systems Applications (DEXA 99)*. IEEE Publishers.
  [5 pages / AUB link / short description]

E2017-7sem-UP-L4 StakeholderAnalysis

Dix+2004chapter13 SocioOrg+Stakeholders

stakeholder-analysis-worksheet

Lecture 3b: Data collection - Contextual inquiry [LDH, 1 hrs]

A common method for investigating people’s behavior & interaction with information, information systems and other people is contextual inquiry. This lecture will introduce this method to the students, followed by a practical exercise with the method.

Required reading

- Holtzblatt, K., & Jones, S. (1993). Contextual Inquiry: A Participatory Technique for System Design (chapter 9). In D. Schuler & A. Namioka (Eds.), *Participatory Design: Principles and Practices* (pp. 178-188). Lawrence Erlbaum Associates, Inc. [11 pages]
- Raven, M. E., & Flanders, A. (1996). Using Contextual Inquiry to Learn about your Audiences. *ACM SIGDOC Asterisk Journal of Computer Documentation*, 20(1), 1-13. [13 pages / Article describing the method of contextual inquiry, its history and context, and the different subtypes of contextual inquiry.]

Additional reading

- Holtzblatt, K., & Beyer, H. (2014). Contextual Design: Evolved. *Synthesis Lectures on Human-Centered Informatics*, 7(4), 1-91. [91 pages]

Literature

- Holtzblatt et al. (1993).pdf
- Raven et al. (1996).pdf

Download folder

Slides Contextual inquiry

Lecture 4: Data collection - Interviews & focus groups [LDH, 2 hrs]

Interviews and focus groups are common used methods for qualitative enquiry of people’s meaning making of practice and requirements in relation to design of IT systems. This lecture will introduce basic principles of these methods to the students, followed by a practical exercise with the methods.

Required reading

- Kvale, S., & Brinkmann, S. (2008). *InterViews: Learning the Craft of Qualitative Research Interviewing* (2nd edition). Los Angeles: SAGE Publications, Inc. p. 1 - 61
- Holtzblatt, K., & Beyer, Hugh. (2016). *Contextual design: design for life* (2nd edn). Cambridge, MA: Elsevier. p. 70 - 81 Chapter on Contextual Interview structure
- Bryman, A. (2016). *Social research methods* (Fifth Edition). Oxford ; New York: Oxford University Press. p. 500 - 525 Chapter 21 on Focus Groups

Additional reading

- Booth, A., Hannes, K., Harden, A., Noyes, J., Harris, J., and Tong, A. (2014). COREQ (Consolidated Criteria for Reporting Qualitative Studies) (chapter 21). In Moher, D., Altman, D. G., Schulz, K. F., Simera, I., and Wager, E. (Eds.), *Guidelines for Reporting Health Research: A User's Manual* (pp. 214-226). Oxford: John Wiley & Sons, Ltd.
- Bryman, A. (2016). *Social research methods* (Fifth Edition). Oxford ; New York: Oxford University Press. p. 374 - 407 (Use as a basic reader)
- Kvale, S., & Brinkmann, S. (2008). *InterViews: Learning the Craft of Qualitative Research Interviewing* (2nd edition). Los Angeles: SAGE Publications, Inc. p. 1 - 350 (Use as a basic reader)

YouTube qualitative research

YouTube: Qualitative analysis of interview data - a basic step-by-step guide

Literature

- Booth et al. (2014).pdf
- Bryman_focus groups.pdf
- Bryman_qualitative research_interviews.pdf

Download folder

Slides - Qualitative interviewing & focus groups

Lecture 5: Data collection - Surveys [CP, 2 hrs]

In this lecture we will focus on how to design and deploy (online) questionnaires. Teaching will consist of a combination of flipped classroom teaching (with class time used for Q&A and discussion) and a group exercise on questionnaire design.

When reading these texts, the focus is on (1) reflecting on the kinds of research situations where questionnaires are a relevant and beneficial method to use, and (2) formulating clear and focused questions with complete and sensible answer options. Both texts contain plenty of tips on which problems and pitfalls to avoid.

Required reading

- Bryman, A. (2016). Self-completion Questionnaires (chapter 10). In *Social Research Methods* (5th ed.), New York, NY: Oxford University Press.
- Bryman, A. (2016). Asking Questions (chapter 11). In *Social Research Methods* (5th ed.), New York, NY: Oxford University Press.
- Goodman, E., Kuniavsky, M., & Moed, A. (2012). Surveys. In *Observing the User Experience: A Practitioner's Guide to User Research* (2nd ed., pp. 327–384).

Additional reading

- Barratt, M. J., Ferris, J. A., & Lenton, S. (2015). Hidden Populations, Online Purposive Sampling, and External Validity Taking off the Blindfold. *Field Methods*, 27(1), 3–21 [link]
- Bhutta, C. B. (2012). Not by the Book: Facebook as a Sampling Frame. *Sociological Methods & Research*, *41*(1), 57–88. [link]
- Couper, M. P. (2000). Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly*, 64(4), 464–494. [link]
- Fan, W., & Yan, Z. (2010). Factors Affecting Response Rates of the Web Survey: A Systematic Review. *Computers in Human Behavior*, 26(2), 132–139. [link]
- Fricker, R. D., & Schonlau, M. (2002). Advantages and Disadvantages of Internet Research Surveys: Evidence from the Literature. *Field Methods*, 14(4), 347–367. [link]
- Galesic, M., & Bosnjak, M. (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*, 73(2), 349–360. [link]
- Grimshaw, J. (2014). SURGE (The SUrvey Reporting GuidelinE) (chapter 20). In Moher, D., Altman, D. G., Schulz, K. F., Simera, I., and Wager, E. (Eds.), *Guidelines for Reporting Health Research: A User's Manual* (pp. 206-213). Oxford: John Wiley & Sons, Ltd.
- Wright, K. B. (2005). Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. *Journal of Computer-Mediated Communication*, 10(3), 00–00. [link]

Slides ('Surveys')

Literature

- Bryman (2012), chapter 10.pdf
- Bryman (2012), chapter 11.pdf
- Goodman et al. (2012), chapter 12.pdf
- Grimshaw (2014).pdf

Download folder

Lecture 6: Data collection - Usability testing [BL, 2 hrs]

Tuesday October 31 @ 14:30-16:15

In this lecture we will focus on basic theory on usability and evaluation methods - that is, what is usability and user experience and which methods do we have test the usability of new and existing products.

Required reading

- Goodman, E., Kuniavsky, M. & Moed, A. (2012). Do a Usability Test Now! (chapter 2). In Observing the User Experience: A Practitioner's Guide to User Research (2nd edition, pp. 11-19). San Francisco, CA: Morgan Kauffman.
  [9 pages / link / Good introductory chapter on doing quick & light usability tests; read this before reading chapter 11]
- Goodman, E., Kuniavsky, M. & Moed, A. (2012). Usability Testing (chapter 11). In Observing the User Experience: A Practitioner's Guide to User Research (2nd edition, pp. 273-326). San Francisco, CA: Morgan Kauffman.
  [54 pages / link / Detailed overview of how to conduct usability tests from start to finish]

Additional reading

- Sauro, J., & Lewis, J. R. (2012) Standardized Usability Questionnaires (chapter 8). In Quantifying the User Experience: Practical Statistics for User Research (pp. 185-240). Waltham, MA: Morgan Kaufmann.
  [56 pages / link / Excellent overview chapter of the different usability questionnaires for different domains, how to interpret them, and their relative strengths & weaknesses]

Goodman+2012-Ch2+Ch11

E2017-7sem-UP-L7 UsabilityTesting

Lecture 7: Data analysis - Descriptive statistics [TB, 4 hrs]

Tuesday October 31 @ 12:30-16:15

In this lecture we will focus on descriptive statistics, a branch of statistics that deals with describing patterns in the data. Descriptive statistics can help us identify outliers, anomalies, and human errors in the data, as well as understand which statistical tests would be appropriate to use. Topics include graphical and non-graphical representations, measures of central tendency and dispersion, and frequency & probability distributions. Teaching will consist of a combination of traditional lecturing and (group) exercises.

Preparation

To prepare for the this lecture (in addition to doing the assigned reading), I'd like you to do the following things:

1. One of the topics of this lecture is levels of measurement, i.e., the different ways we have of measuring variables. I will not discuss this in class. Instead I have prepared a lecture video for you to watch and a quiz to test your knowledge. You can find both the video, the slides they're based on, and the quiz under this lecture. Make sure you watch this before class, because I will start using concepts like 'ratio' and 'ordinal' levels of measurement in the lectures and the lab sessions with explaining them again.

2. We will be using SPSS in the coming lab sessions, which is a program for statistical analysis. Please install it on your laptops and make sure it works **before** you come to class! You can download SPSS from the AAU ITS website here. Make sure you use the right license key to activate it (i.e., if you install version 24, use that version's license key for the 'Base/Advanced' version). While I can try to help you in class if you had problems installing it, your best bet is to fix such problems before class by stopping by the IT department.

Required reading

- Bryman, A. (2016). “Chapter 15: Quantitative data analysis”. In: *Social Research Methods* (4th ed., pp. 329-352), New York, NY: Oxford University Press [24 sider / Dense but comprehensive overview of the different statistical procedures and methods that could be used to analyze different types of data. Read up to and including page 341 for today's lecture, after you read Lazar.]
- Lazar, J., Feng, J.H., and Hochheiser, H. (2010). “Chapter 4: Statistical analysis“. In: *Research Methods in Human-Computer Interaction* (1st ed., pp. 71-104), Hoboken, NJ: John Wiley & Sons Ltd. [34 pages / This chapter provides a practical overview of the different statistical procedures and tests you may end up using to analyze quantitative data. For this lecture, you only need to read pages 71-76 about descriptive analysis. This is a high-level overview, so this is probably the best place to start reading.]

Supplementary reading

- Bryman, A. (2016). “Chapter 16: Using IBM SPSS for Windows”. In: *Social Research Methods* (5th ed., pp. 354-375), New York, NY: Oxford University Press [22 sider / Use this as a how-to guide for using SPSS to perform statistical analysis; for this lecture it's enough to skim it. Don't worry about the "for Windows" part; SPSS works exactly the same on a Mac.]
- Hinton, P. R. (2014). “Chapter 2: Descriptive Statistics”. In: *Statistics Explained* (4th ed.), New York, NY: Routledge [18 sider / A more traditional statistics text about the different methods for doing descriptive analysis. Useful if you want more background information about this.]

Slides ('Levels of measurement')

UAM lecture video: 'Levels of measurement'

Slides ('Descriptive statistics')

Literature

- Bryman (2012), chapter 15.pdf
- Bryman (2012), chapter 16.pdf
- Hinton (2014), chapter 2.pdf
- Lazar et al. (2017), chapter 4.pdf

Download folder

Chart selection guide

Lab session

- Lab session 1 - Descriptive statistics.pdf
- spss-questionnaire-data.sav

Download folder

Quiz ('Levels of measurement')

Quiz: 'Levels of measurement' (answers)

htLecture 8: Data analysis - Inferential statistics, part 1 [TB, 4 hrs]

Monday November 6 @ 12:30-16:15

In this lecture we will focus on inferential statistics, the branch of statistics that allows us to make inferences about the population using statistics from a sample of the population. Inferential statistics also help us reject or accept research hypotheses based on our experimental results. In this first part we will focus mainly on probability distributions, what it means to do statistical testing, and on performing Student's t-tests. Teaching will consist of a combination of traditional lecturing and lab exercises.

Required reading

- Bryman, A. (2016). "Chapter 15: Quantitative data analysis". In: *Social Research Methods* (4th ed., pp. 329-352), New York, NY: Oxford University Press [24 sider / Dense but comprehensive overview of the different statistical procedures and methods that could be used to analyze different types of data. Read the rest of the chapter starting from page 341, after you read Lazar.]
- Lazar, J., Feng, J.H., and Hochheiser, H. (2010). "Chapter 4: Statistical analysis". In: *Research Methods in Human-Computer Interaction* (1st ed., pp. 71-104), Hoboken, NJ: John Wiley & Sons Ltd. [34 pages / This chapter provides a practical overview of the different statistical procedures and tests you may end up using to analyze quantitative data. Read the rest of the chapter for this lecture (starting at page 76). Again, this is a high-level overview, so this is probably the best place to start reading.]

Supplementary reading

- Bryman, A. (2016). "Chapter 16: Using IBM SPSS for Windows". In: *Social Research Methods* (5th ed., pp. 354-375), New York, NY: Oxford University Press [22 sider / Use this as a how-to guide for using SPSS to perform statistical analysis; for this lecture it's enough to skim it. Don't worry about the "for Windows" part; SPSS works exactly the same on a Mac.]

Slides ('Inferential statistics, part 1')

Lab session

Lab session 2 (Inferential statistics, part 1).pdf

spss-ratings-netherlands.sav

spss-reading-speed.sav

Download folder

Guide to statistical testing

Additional exercise on experimental comparisons

Additional exercise (Answers)

Lecture 9: Data analysis - Inferential statistics, part 2 [TB, 2 hrs]

Wednesday November 8 @ 10:15-12:00

In this second part we will focus on correlation analysis and Chi-square testing. Teaching will consist of a combination of traditional lecturing and lab exercises.

Required reading

- Bryman, A. (2016). "Chapter 15: Quantitative data analysis". In: *Social Research Methods* (4th ed., pp. 329-352), New York, NY: Oxford University Press [24 sider / Dense but comprehensive overview of the different statistical procedures and methods that could be used to analyze different types of data. Read the rest of the chapter starting from page 341, after you read Lazar.]
- Lazar, J., Feng, J.H., and Hochheiser, H. (2010). "Chapter 4: Statistical analysis". In: *Research Methods in Human-Computer Interaction* (1st ed., pp. 71-104), Hoboken, NJ: John Wiley & Sons Ltd. [34 pages / This chapter provides a practical overview of the different statistical procedures and tests you may end up using to analyze quantitative data. Read the rest of the chapter for this lecture (starting at page 76). Again, this is a high-level overview, so this is probably the best place to start reading.]

Supplementary reading

- Bryman, A. (2016). "Chapter 16: Using IBM SPSS for Windows". In: *Social Research Methods* (5th ed., pp. 354-375), New York, NY: Oxford University Press [22 sider / Use this as a how-to guide for using SPSS to perform statistical analysis; for this lecture it's enough to skim it. Don't worry about the "for Windows" part; SPSS works exactly the same on a Mac.]

Slides ('Inferential statistics, part 2')

Lab session

Lab session 3 (Inferential statistics, part 2).pdf

spss-navigation-app-data.sav

spss-survey-data.sav

Download folder

ICT-based Data Collection & Analysis

Welcome to the Moodle course page for **ICT-based Data Collection & Analysis** (DCA), a 5-ECTS study module on the 7th semester of IS, taught by Toine Bogers (toine@hum.aau.dk, course responsible), Florian Meier (fmeier@hum.aau.dk) and Birger Larsen (birger@hum.aau.dk).

It is centered around methods and tools for ICT-based data collection and analysis, aimed at supporting students in their future ICT-oriented user analyses and projects. In addition, students will learn what the paradigm of Big Data entails and how it can relate to their own projects and practices. Finally, we will discuss the ethical concerns that data collection and analysis methods may raise.

Teaching will be a combination of lectures and lab sessions, where the students will gain practical experience with some of these data collection and analysis methods. Your exam will take place in week TBD.

All literature will either be uploaded to Moodle or is linked to directly. Please remember that these links only work if you are connected to our campus wifi or through a VPN connection with the university. If you think there is a reference missing or a link that does not work, please send me an e-mail about it. If you have any questions about this course in general, please send us an email to toine@hum.aau.dkor post a message in the discussion forum.

Announcements

1. Introduction + Big data [TB, 2 hrs]

DATE @ TIME

The first lecture will introduce you to the formalities of the DCA course and provide you with a general introduction to different paradigms of data collection. We will also discuss the different interpretations of 'Big Data' and the promises and pitfalls it brings.

Required reading

- boyd, d. & Crawford, K. (2012). Critical Questions for Big Data, *Information, Communication & Society*, 15(5), pp. 662-679 (link) [18 pages / Comprehensive article that offers a critical discussion of Big Data, its benefits, and its pitfalls]
- Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big Data and Its Technical Challenges. *Communications of the ACM*, 57(7), 86-94. (link) [9 pages / Overview article of Big Data and the life cycle of Big Data projects]
- Lohr, S. (2012). The Age of Big Data, *New York Times*, p. SR1, Retrieved September 10, from http://www.nytimes.com (link) [7 pages / Introductory newspaper article about Big Data; read this first]

Supplementary reading

- Asur, S., & Huberman, B. A. (2010). Predicting the Future with Social Media. In *WI-IAT '10: Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. (pp. 492-499). (link) [8 pages / Paper that describes experiments with predicting box office success using movie mentions on Twitter.]
- Becker, R., Volinsky, C., Cáceres, R., Hanson, K., Isaacman, S., Loh, J. M., ... Varshavsky, A. (2013). Human Mobility Characterization from Cellular Network Data. *Communications of the ACM*, 56(1), 74-82. (link) [8 pages / Article that describes a set of analyses using cellular network data focusing on different aspects of human mobility.]

Slides ('Course introduction')

Hidden from students

Slides ('Big Data')

Hidden from students

Literature

- Asur et al. (2010).pdf
- Becker et al. (2013).pdf
- boyd et al. (2012).pdf
- Jagadish et al. (2014).pdf

Download folder

2. Data collection: Observation (w/ eyetracking) [BL, 4 hrs]

DATE @ TIME

This lecture will provide you with an overview and examples of different IT-based observation techniques, including screen recording, remote usability testing and eye-tracking.

Required reading

- Webb, N., & Renshaw, T. (2008). Eyetracking in HCI. In P. Cains & A. L. Cox (Eds.),*Research Methods for Human-Computer Interaction* (pp. 35-69). Cambridge, UK: Cambridge University Press. [35 pages]

iMotions 2017 GSR-PocketGuide

Hidden from students

iMotions 2017 EyeTracking-TheCompletePocketGuide

Hidden from students

Webb+2008eyetrackingInHCI

Hidden from students

7sem-DCA-L2 IT-based-observation

Hidden from students

iMotionsTutorial

Hidden from students

Eyetracking

Hidden from students

SW-Usability-Testing-with-Morae

Hidden from students

Morae-trial_Windows.zip

Hidden from students

Morae-observer.zip

Hidden from students

Upload slides here

Hidden from students

Slides-template

Hidden from students

3. Data collection: Crawling & analysing social media data [FM, 4 hrs]

DATE @ TIME

The increased popularity of social media, such as Facebook, Twitter, Instagram, and Snapchat, has resulted in these systems influencing our work- and leisure-related information behavior. In this lab session, you will gain practical experience with collecting data from different social media websites.

A big part of the popularity of social media is the social aspect: social media users connect and interact with each other in many ways, forming implicit and explicit social networks. In this lecture, we focus on the pros and cons of collecting network data and introduce practical tools for collecting data from social media.

Required reading

- Fan, W., & Gordon, M. D. (2014). The Power of Social Media Analytics. *Communications of the ACM, 57*(6), 74-81. (link) [8 pages / Clear overview article about the different stages and applications of analytics social media.]
- Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... & Jebara, T. (2009). Life in the Network: The Coming Age of Computational Social Science. *Science*, 323(5915), 721-723. (link) [3 pages / Short overview article that places the study of social media behavior in the wider context of computational social science.]
- Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing Social Networks*. SAGE Publications Limited. (read chapters: 'Mathematical foundations', 'Visualization', 'Characterizing whole networks', 'centrality', 'subgroups') [98 pages / Introductory chapters to social network analysis; read these first before the other required reading.]
- Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PloS One*, 9(6), e98679. [22 pages]
- Jensen, P., Morini, M., Karsai, M., Venturini, T., Vespignani, A., Jacomy, M., ... & Fleury, E. (2015). Detecting Global Bridges in Networks. *Journal of Complex Networks*, 4(3), 319-329. [14 pages]

Supplementary reading

- boyd, d.m. and Ellison, N.B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230. (link) [21 pages / Good introduction to social media and social network sites.]

Tutorial ('Crawling social media using Netlytic')

Hidden from students

Network Slides

Hidden from students

Festivals

Hidden from students

Literature

- boyd et al. (2012).pdf
- Fan et al. (2014).pdf
- Lazer et al. (2009).pdf

Download folder

4. Data analysis: Log analysis [FM, 2 hrs]

DATE @ TIME

This lecture will introduce students to the theory and practice behind the large-scale analysis of interaction and transaction logs. Navigating through the web we constantly leave footprints providing information about us. User and usage data is a valuable source not only for commercial organizations like Google but also for academic research. In this session we want to explore how usage data can be leveraged as well as the challenges and limitations of log file analysis.

Required reading

- Agosti, M., Crivellari, F., Di Nunzio, G.M. (2012). Web Log Analysis: A Review of a Decade of Studies about Information Acquisition, Inspection and Interpretation of User Interaction. *Data Mining and Knowledge Discovery*, 24(3), 663-696. [34 pages]
- Jansen, B. J., Taksa, I., & Spink, A. (2009). Chapter 1: Research and Methodological Foundations of Transaction Log Analysis. In B. J. Jansen, A. Spink, & I. Taksa (Eds.), *Handbook of Research on Web Log Analysis* (pp. 1-16). Hershey, PA: Information Science Reference. [16 pages]

Slides ('Log analysis')

Hidden from students

Literature

- Agosti et al. (2011).pdf
- Jansen et al. (2009).pdf

Download folder

Data sets

- Europeana_Queries_Session_Click_Stats.xlsx
- Europeana_top_filters.xls

Download folder

Hidden from students

5. Data analysis: Content analysis [FM, 2 hrs]

DATE @ TIME

This lecture will highlight some commonly used techniques for analyzing and enriching data. Content analysis is a commonly used technique for coding textual data, such as interviews, focus groups transcripts and usability logs. Students will gain practical experience in coding qualitative data. We also discuss how to evaluate the quality of such annotations using metrics like Cohen's kappa and inter-annotator agreement. In addition, we will discuss automatic sentiment analysis, which allows us to detect the sentiment and opinions expressed in text.

Required reading

- Lazar, J., Feng, J.H., and Hochheiser, H. (2010). "Chapter 11: Analyzing qualitative data". In *Research Methods in Human-Computer Interaction* (pp. 281-306). Wiley [26 pages / Great overview chapter on qualitative data analysis, covering appropriate coding procedures as well as reliability and validity of the coding process.]

Supplementary reading

- Bryman, A. (2016). "Chapter 13: Content analysis". In: *Social Research Methods*(4th ed.), New York, NY: Oxford University Press (pp. 288-309). Wiley [22 pages / Good overview chapter about content analysis (as it refers to performing a more systematic, quantitative analysis of documents and texts).]
- Bryman, A. (2016). "Chapter 24: Qualitative data analysis". In: *Social Research Methods* (4th ed.), New York, NY: Oxford University Press (pp. 564-589). Wiley [26 pages / Good overview chapter about qualitative analysis, covering aspects like grounded theory & narrative analysis.]

Slides ('Content analysis')

Hidden from students

Literature

- Bryman (2012), chapter 13.pdf
- Bryman (2012), chapter 24.pdf
- Lazar et al. (2010), chapter 11.pdf
- Wright (2009).pdf

ht

Download folder

Data sets

- Cohens kappa.xlsx
- IMDB posts.xlsx
- iPhoneX tweets.xlsx

Download folder

Hidden from students

Example coding scheme for movie search requests

Hidden from students

6. Data analysis: Data visualization (w/ Tableau) [TB, 4 hrs]

DATE @ TIME

One of the most important steps in both analyzing and communicating research data is data visualization using statistical graphics, plots and information graphics. Effective visualization helps users analyze and reason about data and evidence, and it makes complex data more accessible, understandable and usable. Data visualization is both an art and a science. It is viewed as a branch of descriptive statistics by some, but also as a grounded theory development tool by others. Through the required reading, the basics of data visualization are introduced, followed by a lab session where students learn to use Tableau, a popular type of interactive data visualization software.

Required reading

- Fekete, J. D., Van Wijk, J., Stasko, J., & North, C. (2008). The Value of Information Visualization. *Information Visualization*, 1-18. [18 pages]
- Kirk, A. (2012). Chapter 4: Conceiving and Reasoning Visualization Design Options. In *Data Visualization: A Successful Design Process* (pp. 79-117). Birmingham: Packt Publishing. [39 pages]
- Kirk, A. (2012). Chapter 5: Taxonomy of Data Visualization Methods. In *Data Visualization: A Successful Design Process* (pp. 119-158). Birmingham: Packt Publishing. [40 pages]
- Shneiderman, B. (1996). The Eyes have it: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the IEEE Symposium on Visual Languages* (pp. 336-343). [8 pages / Classic article describing a taxonomy of different visualization techniques organized by the type of data being visualized; also a very useful guide for interface design.]

Supplementary reading

- Kirk, A. (2016). *Data Visualisation: A Handbook for Data-driven Design*. Sage. [368 pages]
- Murray, D. G. (2013). *Tableau your Data! Fast and Easy Visual Analysis with Tableau Software*. John Wiley & Sons. [528 pages]
- Iliinsky, N., & Steele, J. (2011). *Designing Data Visualizations: Representing Informational Relationships.* O'Reilly Media, Inc. [114 pages]
- Lupi, G., & Posavec, S. (2016). *Dear Data*. New York: Princeton Architectural Press. [288 pages / Insightful and entertaining experiment into personal data visualization; worth borrowing from the library if you're passionate about visualization.]
- Tegarden, D. P. (1999). Business Information Visualization. *Communications of the AIS*, 1(4), 1-38. [38 pages]
- Tufte, E. R. (2001). *The Visual Display of Quantitative Information* (2nd ed.), Cheshire, CT: Graphics Press. [200 pages]

Tutorial ('Data visualization using Tableau')

Literature

- Fekete et al. (2008).pdf
- Kirk (2012), chapter 4.pdf
- Kirk (2012), chapter 5.pdf
- Shneiderman (1996).pdf
- Tegarden (1999).pdf

Download folder

Data sets

- facebook-survey.xlsx
- geotagged-tweets.xlsx
- tweet-sentiment.xlsx
- uk-tweets.xlsx

Download folder

7. Legal & ethical aspects of DCA [TB/FM/BL, 4 hrs]

DATE @ TIME

This lecture will focus on the legal and ethical aspects of online data collection. After some introductory remarks, the majority of the time will be taken up by debates htbetween the four groups of students (that were created in the first lecture).

Preparation

Prepare for the debate by meeting at least once and going through the arguments you want to make in favor or against your debate topic. In addition to the required reading, try to read at least one of the supplementary articles to get a better feel for ethical and legal aspects of online data collection.

Required reading

- boyd, d. & Crawford, K. (2012). Critical Questions for Big Data, *Information, Communication & Society*, 15(5), pp. 662-679 (link) [10 pages.]

Supplementary reading

- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to Ideologically Diverse News and Opinion on Facebook. *Science*, 348(6239), 1130-1132 (link) [3 pages.]
- Barnes, S. B. (2006). A Privacy Paradox: Social Networking in the United States. *First Monday*, 11(9) (link) [11 pages.]
- Epstein, R., & Robertson, R. E. (2015). The Search Engine Manipulation Effect (SEME) and its Possible Impact on the Outcomes of Elections. *Proceedings of the National Academy of Sciences,* 112(33), E4512-E4521 (link) [10 pages.]
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), 1203-1205. (link) [10 pages / Article discussing how Google's Big Data project for flu tracking called Google Flu was way off the mark in its predictions.]
- Zimmer, M. (2010). "But the Data is Already Public": On the Ethics of Research in Facebook. *Ethics and Information Technology*, 12(4), 313–325 (link) [13 pages.]
- Zwitter, A. (2014). Big Data Ethics. *Big Data & Society*, 1(2) (link) [6 pages.]

Slides ('Ethical aspects of DCA')

Hidden from students

Module description

Placement: 7th semester

Module coordinator: Toine Bogers

Type and language: Study module / English

Objectives

At the end of the module, students are expected to be able to:

Knowledge

- theories and methods at the highest international level as regards qualitative and quantitative-oriented data collection and analysis in relation to user analyses and pilot studies
- ICT systems for data collection and analysis in relation to user analyses and pilot studies
- principles, including ethical principles, for managing ICT systems for data collection and analysis in relation to user analyses and pilot studies

Skills

- assessing and selecting a method for qualitative and quantitative oriented data collection and analysis in relation to user analyses and pilot studies
- selecting, configuring and adapting ICT systems for qualitative and quantitative oriented data collection and analysis in relation to user analyses and pilot studies
- communicating methods & results for ICT-based data collection and analysis to peers and laymen

Competences

- taking an analytical, reflective and critical approach to qualitative and quantitative oriented data collection and analysis in relation to user analyses and pilot studies
- engaging in interdisciplinary collaboration on ICT based data collection and analysis in relation to user anal-yses and pilot studies
- identifying own learning needs and structuring own learning in relation to the subject area of ICT-based data collection and analysis in relation to user analyses and pilot studies

Academic content and basis

The module will introduce students to ICT-based data collection and analysis offering a number of opportunities to obtain vast amounts of data on the use of for example Web based ICT solutions with relative ease. These opportunities call for fundamental consideration of options and problems, including ethical concerns on the significance of the potentially extensive mappings of individual user behaviour. During the course of the module, students will engage in ICT-based data collection and analysis for the support of ICT user analyses and pilot projects.

Scope and expectation

The module comprises 5 ECTS, which corresponds to 137.5 working hours, consisting of a mixture of lectures and lab sessions.

Literature

	Required (pages)	Supplementary (pages)	Digital upload
1. Introduction + Big data			
boyd, d. & Crawford, K. (2012). Critical Questions for Big Data, Information, Communication & Society, 15(5), pp. 662-679	18		
Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big Data and Its Technical Challenges. Communications of the ACM, 57(7), 86-94.	9		

	Required (pages)	Supplementary (pages)	Digital upload
Lohr, S. (2012). The Age of Big Data, New York Times, p. SR1, Retrieved September 10, from http://www.nytimes.com		8	
Asur, S., & Huberman, B. A. (2010). Predicting the Future with Social Media. In WI-IAT '10: Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. (pp. 492-499).		8	
Becker, R., Volinsky, C., Cáceres, R., Hanson, K., Isaacman, S., Loh, J. M., ... Varshavsky, A. (2013). Human Mobility Characterization from Cellular Network Data. Communications of the ACM, 56(1), 74-82.		8	
2. Data collection: Observation (w/ eyetracking)			
Webb, N., & Renshaw, T. (2008). Eyetracking in HCI. In P. Cains & A. L. Cox (Eds.), Research Methods for Human-Computer Interaction (pp. 35-69). Cambridge, UK: Cambridge University Press.	35		
3. Data collection: Crawling & analysing social media data			
Fan, W., & Gordon, M. D. (2014). The Power of Social Media Analytics. Communications of the ACM, 57(6), 74-81.	8		
Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... & Jebara, T. (2009). Life in the Network: The Coming Age of Computational Social Science. Science, 323(5915), 721-723.	3		
Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). Analyzing Social Networks. SAGE Publications Limited.	98		
Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PloS One, 9(6), e98679.	22		
Jensen, P., Morini, M., Karsai, M., Venturini, T., Vespignani, A., Jacomy, M., ... & Fleury, E. (2015). Detecting Global Bridges in Networks. Journal of Complex Networks, 4(3), 319-329.	14		
boyd, d.m. and Ellison, N.B. (2007). Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.		21	
4. Data analysis: Log analysis			
Agosti, M., Crivellari, F., Di Nunzio, G.M. (2012). Web Log Analysis: A Review of a Decade of Studies about Information Acquisition, Inspection and Interpretation of User Interaction. Data Mining and Knowledge Discovery, 24(3), 663-696.	34		
Jansen, B. J., Taksa, I., & Spink, A. (2009). Chapter 1: Research and Methodological Foundations of Transaction Log Analysis. In B. J. Jansen, A. Spink, & I. Taksa (Eds.), Handbook of Research on Web Log Analysis (pp. 1-16). Hershey, PA: Information Science Reference.	16		
5. Data analysis: Content analysis			
Lazar, J., Feng, J.H., and Hochheiser, H. (2010). ”Chapter 11: Analyzing qualitative data”. In Research Methods in Human-Computer Interaction (pp. 281-306). Wiley.	26		
Bryman, A. (2016). “Chapter 13: Content analysis”. In: Social Research Methods (4th ed.), New York, NY: Oxford University Press (pp. 288-309). Wiley.		22	
Bryman, A. (2016). “Chapter 24: Qualitative data analysis”. In: Social Research Methods (4th ed.), New York, NY: Oxford University Press (pp. 564-589). Wiley.		26	
6. Data analysis: Data visualization (w/ Tableau)			
Fekete, J. D., Van Wijk, J., Stasko, J., & North, C. (2008). The Value of Information Visualization. Information Visualization, 1-18.	18		
Kirk, A. (2012). Chapter 4: Conceiving and Reasoning Visualization Design Options. In Data Visualization: A Successful Design Process (pp. 79-117). Birmingham: Packt Publishing.	39		
Kirk, A. (2012). Chapter 5: Taxonomy of Data Visualization Methods. In Data Visualization: A Successful Design Process (pp. 119-158). Birmingham: Packt Publishing.	40		
Shneiderman, B. (1996). The Eyes have it: A Task by Data Type Taxonomy for Information Visualizations. In Proceedings of the IEEE Symposium on Visual Languages (pp. 336-343).	8		
Kirk, A. (2016). Data Visualisation: A Handbook for Data-driven Design. Sage.		368	
Murray, D. G. (2013). Tableau your Data! Fast and Easy Visual Analysis with Tableau Software. John Wiley & Sons.		528	
Iliinsky, N., & Steele, J. (2011). Designing Data Visualizations: Representing Informational Relationships. O'Reilly Media, Inc.		114	
Lupi, G., & Posavec, S. (2016). Dear Data. New York: Princeton Architectural Press.		288	
Tegarden, D. P. (1999). Business Information Visualization. Communications of the AIS, 1(4), 1-38.		38	
Tufte, E. R. (2001). The Visual Display of Quantitative Information (2nd ed.), Cheshire, CT: Graphics Press.		200	
7. Legal & ethical aspects of DCA			
boyd, d.m. and Ellison, N.B. (2007). Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 13(1), 210-230.	21		
Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to Ideologically Diverse News and Opinion on Facebook. Science, 348(6239), 1130-1132.		3	
Barnes, S. B. (2006). A Privacy Paradox: Social Networking in the United States. First Monday, 11(9).		11	

	Required (pages)	Supplementary (pages)	Digital upload
Epstein, R., & Robertson, R. E. (2015). The Search Engine Manipulation Effect (SEME) and its Possible Impact on the Outcomes of Elections. Proceedings of the National Academy of Sciences, 112(33), E4512-E4521		10	
Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. Science, 343(6176), 1203-1205.		10	
Zimmer, M. (2010). “But the Data is Already Public”: On the Ethics of Research in Facebook. Ethics and Information Technology, 12(4), 313–325.		13	
Zwitter, A. (2014). Big Data Ethics. Big Data & Society, 1(2).		6	
Totals			
	409	1682	

Collective Intelligence

1 2 3 4 5 6 7 Active participation Module description Literature Exam

Welcome to the Moodle course page for Collective Intelligence (CI), a 5-ECTS elective course on the 8th semester of both IS and KOM, taught by Florian Meier (FM, course responsible) and Toine Bogers (TB). It is centered on the concept of collective intelligence: combining the behavior, preferences, or ideas of a group of people to create novel insights. Students will learn how to interpret and organize the enormous amounts of user-generated content on the Web to produce novel insights about user experience, marketing, personal tastes, and human behavior in general. We will focus especially on a range of different sophisticated techniques and algorithms for harnessing such collective intelligence automatically. Teaching will be a combination of lectures and lab sessions, where the students will apply some of these techniques to real-world data sets.

Teaching consists of 20 hours of lectures and lab sessionsis divided over eleven two-hour sessions, most of which will be spent on lab sessions where you will work on programming exercises and tasks. Before each lecture, students can send clarification questions to the teachers which will then be explained in more detail. The remainder of the time will be spent on practical programming exercises. If you have any questions about the course, please send an e-mail to fmeier@hum.aau.dk.

Announcements

Course pitch slides

Hidden from students

Testimonials (2014-2017)

Hidden from students

1. Introduction + Wisdom of crowds [TB, 2 hrs]

DATE @ TIME

This lecture will introduce students to concept of 'wisdom of crowds' (also known as 'collective intelligence') and will discuss why and when a crowd is smarter than individuals. The accompanying active participation assignment has students perform their own wisdom-of-crowds experiment, analyze and present the results during one of the other sessions in the course.

Required reading

- Malone, T.W., Laubacher, R., and Dellarocas, C. (2010). The Collective Intelligence Genome. *IEEE Engineering Management Review*, 38(3), 21-31. (link) [11 pages / Clear article that provides a framework for categorizing different applications of collective intelligence]
- Woolley et al. (2010). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, 330(6004), 686-688. (link) [3 pages / Article describing a collection of studies that try to determine what influences group intelligence: what is it about a group that makes it intelligent?]

Supplementary reading

- Page, S. E. (2007). *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton: Princeton University Press. [456 pages / Comprehensive & quantitative approach to collective intelligence and how you can show mathematically that diversity benefits the outcome of many groups under specific conditions]
- Surowiecki, J. (2005). *The Wisdom of Crowds*. New York: Anchor Books [336 pages / Best introduction to wisdom of crowds and collective intelligence. Read this first if you want something interesting to read on vacation]

Active participation requirement(s)

- Designing, executing, analyzing, and presenting a Wisdom of Crowds experiment [15 hours / Students (ideally in groups of two) should design their own wisdom of crowds task—more specifically a so-called cognition task—collect responses from participants, and analyze the results. Each of the other course lecture will then feature 1-2 groups presenting their results]

Slides ('Introduction')

Hidden from students

Slides ('Wisdom of Crowds')

Hidden from students

M&M spreadsheet

Hidden from students

Assignment 1 ('Wisdom of Crowds')

Hidden from students

Assignment 1 ('Presentation slots')

Hidden from students

Assignment 1 ('Slide uploads')

Hidden from students

Literature

Malone et al. (2010).pdf

Woolley et al. (2010).pdf

Download folder

2. Online matchmaking / Machine learning [TB, 4 hrs]

DATE @ TIME

This lecture will introduce students to the basics of machine learning. Students will learn how computers can automatically learn how to solve problems with a particular focus on decision tree learning. In the lab session, students will apply decision tree learning to the problem of online matchmaking on online dating sites.

Required reading

- Abbott, D. (2014). Chapter 8: Predictive Modeling. In *Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst* (pp. 213-222). Indianapolis, IN: Wiley. [10 pages / Useful introductions to decision tree learning (and several other algorithms.]
- Rokach, L., & Maimon, O. (2015). Chapter 1: Introduction to Decision Trees. In *Data Mining with Decision Trees: Theory and Applications* (2nd ed., Vol. 81, Series in Machine Perception and Artificial Intelligence, pp. 1-16). London: World Scientific Publishing. [16 pages / Part of the introduction to decision tree learning.]
- Rokach, L., & Maimon, O. (2015). Chapter 2: Training Decision Trees. In *Data Mining with Decision Trees: Theory and Applications* (2nd ed., Vol. 81, Series in Machine Perception and Artificial Intelligence, pp. 17-21). London: World Scientific Publishing. [5 pages / Part of the introduction to decision tree learning.]
- Yee, S., & Chu, T. (2015, July 27). A visual introduction to machine learning. Retrieved August 19, 2017, from http://www.r2d3.us/visual-intro-to-machine-learning-part-1/ [5 pages / Amazingly intuivive visual explanation of how decision tree learning works; watch this before anything else.]
- Yee, S., & Chu, T. (2018, June 18). A visual introduction to machine learning - part II: Model Tuning and the Bias-Variance Tradeoff. Retrieved June 19, 2018, from http://www.r2d3.us/visual-intro-to-machine-learning-part-2/ [7 pages / Great visual explanation of the problem of overfitting in evaluating machine learning, also known as the bias-variance trade-off; watch this after part 1.]

Supplementary reading

- Akser, M. (2015). Old and New Methods for Online Research: The Case of Online Dating. In I. A. Degim, J. Johnson, & T. Fu (Eds.), *Online Courtship: Interpersonal Interactions Across Borders* (Vol. 16, Theory on Demand, pp. 29-35). Amsterdam: Institute of Network Cultures. [7 pages / Article from an anthology on online dating, which is worth checking out if you want to read more academic perspectives on online interaction. This particular article looks at online dating and the different methods used to investigate it. Good introduction if you want to know more about studying online dating]
- Fu, T. (2015). What are the Shengnv Looking for in Online Heterosexual Dating and Courtship? A Content Analysis of Shanghainese Women's Personal Profiles on Jiayuan.com. In I. A. Degim, J. Johnson, & T. Fu (Eds.), *Online Courtship: Interpersonal Interactions Across Borders* (Vol. 16, Theory on Demand, pp. 186-207). Amsterdam: Institute of Network Cultures. [22 pages / Another article from that same anthology. This one describes what women look for and how they describe themselves, which could be useful when thinking about which features are relevant to predict who to date. It might also shed some light on the differences and similarities between online dating in Asia vs. Western countries]
- Mitchell, T. M. (1997). Chapter 3: Decision tree learning. In *Machine Learning* (pp. 52-60). New York: McGraw-Hill [9 pages / More mathematical and formal introduction to decision tree learning.]
- Rokach, L., & Maimon, O. (2015). Chapter 7: Popular Decision Trees Induction Algorithms. In *Data Mining with Decision Trees: Theory and Applications* (2nd ed., Vol. 81, Series in Machine Perception and Artificial Intelligence, pp. 77-83). London: World Scientific Publishing. [7 pages / Overview of the different types of decision tree algorithms.]
- Reichert. R. (2015). Dating Maps: Mapping Love in Online Dating Communities. In I. A. Degim, J. Johnson, & T. Fu (Eds.), *Online Courtship: Interpersonal Interactions Across Borders* (Vol. 16, Theory on Demand, pp. 36-46). Amsterdam: Institute of Network Cultures. [11 pages / And yet article from that same anthology. This one takes a social network analysis perspective]
- Segaran, T. (2007). *Programming Collective Intelligence*. Beijing: O'Reilly Media [362 pages / If you want to learn more practical information about how many collective intelligence techniques work and how to implement them in Python, then this is a good place to start.]

Active participation requirement(s)

- Analysis assignment on real-world data on one of the following three topics: (1) machine learning, (2) social network analysis, or (3) clustering & visualization [20 hours / The analysis assignment typically requires you to apply some of the techniques you've learned for one of the three topics (machine learning, social network analysis, or clustering & visualization) to a data set and write a report about the results and its analysis. More details about the actual assignment will be provided after each of these lectures.]

Slides ('Online matchmaking / Machine learning')

Hidden from students

Assignment 2a ('Machine learning')

Hidden from students

Literature

Abbott (2014), chapter 8.pdf

Akser (2015).pdf

Fu (2015).pdf

Mitchell (1997), chapter 3.pdf

Reichert (2015).pdf

ht

Rokach et al. (2015), chapter 1.pdf
Rokach et al. (2015), chapter 2.pdf
Rokach et al. (2015), chapter 7.pdf
Download folder

Lab session
Background (Evaluating machine learning).pdf
Feedback from previous years.pdf
matchmaker.tab
Tutorial (Online matchmaking).pdf
Download folder

Hidden from students

3. Clustering & visualization [FM, 4 hrs]

DATE @ TIME

This lecture will introduce students to the basic concepts behind clustering and visualization of data: how can we detect and visualize groups of similar items in a large data set? The lab session has students gain hands-on experience with clustering & visualization using specialized software.

Required reading

- Tan, P.-N., Steinbach, M. & Kumar, V. (2006). Chapter 8: Cluster Analysis: Basic Concepts and Algorithms. In *Introduction to Data Mining* (pp. 487-568). Harlow: Addison Wesley. (link) [82 pages / Introduction to clustering algorithms.]

Active participation requirement(s)

- Analysis assignment on real-world data on one of the following three topics: (1) machine learning, (2) social network analysis, or (3) clustering & visualization [20 hours / The analysis assignment typically requires you to apply some of the techniques you've learned for one of the three topics (machine learning, social network analysis, or clustering & visualization) to a data set and write a report about the results and its analysis. More details about the actual assignment will be provided after each of these lectures.]

Slides ('Clustering & visualization')
Hidden from students

Assignment 2c ('Clustering and Visualisation')
Hidden from students

Tutorial ('Clustering & visualization')
Hidden from students

Link to 'zoo' dataset
Hidden from students

4. Sentiment analysis [FM, 2 hrs]

DATE @ TIME

This lecture will introduce students to sentiment analysis. Sentiment analysis aims to determine the attitude of a speaker, writer, or other subject with respect to some topic or the overall contextual polarity or emotional reaction to a document, interaction, or event.

Required reading

- Wright, A. (2009). Our Sentiments, Exactly. *Communications of the ACM*, 52(4), 14-15 [3 pages / Overview paper about sentiment analysis.]
- TBA [? pages]

5. Social network analysis [FM, 4 hrs]

DATE @ TIME

This lecture will teach students about social networks, network theory and social network analysis (SNA) and its applications. In the lab session, students will map the social network of their fellow Master's students and analyze it using specialized software.

Required reading

ht
- Golbeck, J. (2013). *Analyzing the Social Web*, chapters 1-5. Boston, MA: Morgan-Kaufmann. [74 pages / These five chapters will introduce you to the standard terminology and concepts in (social) network analysis. They are well-written and include a lot of visualizations. Read these chapters first before anything else]

- Haythornthwaite, C. (1996). Social Network Analysis: An Approach and Technique for the Study of Information Exchange. *Library & Information Science Research*, 18(4), 323-342. (link) [20 pages / Article describing how social network analysis can be used in the field of information science (and library science).]

Supplementary reading

- Gol Fan, W., & Gordon, M. D. (2014). The Power of Social Media Analytics. *Communications of the ACM*, 57(6), 74-81. (link) [8 pages / High-level introductory article on what social media analytics can do and what it can be used for (which includes social network analysis)]
- Golbeck, J. (2013). *Analyzing the Social Web*, chapters 7 + 10. Boston, MA: Morgan-Kaufmann. [34 pages / Two additional chapters that cover a few more topics that are relevant in social network analysis for those of you that would like to work more with this.]
- Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social Media Analytics: An Interdisciplinary Approach and Its Implications for Information Systems. *Business & Information Systems Engineering*, 6(2), 89–96. (link) [8 pages / For those of you who want the bigger picture, this article is probably a first read (compared to Gol Fan et al.). it casts social media analytics as an interdisciplinary research field, describes its main components (one of which is SNA) and suggests future research directions.]

Active participation requirement(s)

- Analysis assignment on real-world data on one of the following three topics: (1) machine learning, (2) social network analysis, or (3) clustering & visualization [20 hours / The analysis assignment typically requires you to apply some of the techniques you've learned for one of the three topics (machine learning, social network analysis, or clustering & visualization) to a data set and write a report about the results and its analysis. More details about the actual assignment will be provided after each of these lectures.]

Slides ('Social network analysis')

Hidden from students

Literature

- Fan et al. (2014).pdf
- Golbeck (2013), chapter 1.pdf
- Golbeck (2013), chapter 2.pdf
- Golbeck (2013), chapter 3.pdf
- Golbeck (2013), chapter 4.pdf
- Golbeck (2013), chapter 5.pdf
- Golbeck (2013), chapter 7.pdf
- Golbeck (2013), chapter 10.pdf

Download folder

Lab session

- karate-club.gexf
- lastfm-artists.gexf
- Tutorial (SNA with Gephi).pdf

Download folder

Hidden from students

Assignment 2

- Assignment 2b (Social network analysis).pdf
- enron-email.500.gexf
- enron-email.1000.gexf
- enron-email.2000.gexf
- lastfm-artists.gexf
- lastfm-users.250.gexf
- lastfm-users.500.gexf
- lastfm-users.2728.gexf

Download folder

Hidden from students

6. Recommender systems [TB, 2hrs]

DATE @ TIME

This lecture will introduce students to the recommender systems: systems that recommend new and interesting products to users based on their past interests and those of others.

ht**Required reading**

- Resnick, P. and Varian, H.R. (1997). Recommender Systems. *Communications of the ACM*, 40(3), 56-58 (link) [3 pages / Short, high-level introduction to what recommender systems are]
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (2011). Introduction to Recommender Systems Handbook (chapter 1). In *Recommender Systems Handbook* (pp. 1-29). Berlin: Springer [29 pages / More up-to-date and comprehensive introduction to recommender systems, the different types of approaches and the current challenges in recommender systems research]

Slides ('Recommender systems')

Literature

Resnick et al. (1997).pdf

Ricci et al. (2011).pdf

Download folder

7. Brainstorm session [FM/TB, 2hrs]

DATE @ TIME

In the final lab session, students will brainstorm about applying the different CI techniques covered in the course to a specific case.

Active participation requirement(s)

- Brainstorm assignment (Putting CI into practice) [20 hours / The goal of this exercise is brainstorm about how the methods and techniques you've learned about in the course can be applied to an existing website to improve it. For the first part of the brainstorm assignment each group is required to write a 2-3-page report describing their proposed improvement(s) to a selected website. For the second part, each group has to critically peer-review another group's report. Finally, each group should incorporate the feedback they receive from another group. More details to follow in class.]
- Collective course evaluation [1 hour / The final active participation requirement is participating in the collective course evaluation. More details to follow in class]

Slides ('Brainstorm session')

Assignment 4 - Brainstorm session

Assignment 4: Draft report uploads (deadline Nov 23 @ 23:59)

Assignment 4: Peer review uploads (deadline Nov 30 @ 23:59)

Assignment 4: Final report uploads (deadline Dec 07 @ 23:59)

Feedback schedule

Overview of active participation assignments

For your convenience, here is an overview of all the active participation assignments:

- Designing, executing, analyzing, and presenting a Wisdom of Crowds experiment [15 hours / Students (ideally in groups of two) should design their own wisdom of crowds task—more specifically a so-called cognition task—collect responses from participants, and analyze the results. Each of the other course lecture will then feature 1-2 groups presenting their results]
- Analysis assignment on real-world data on one of the following three topics: (1) machine learning, (2) social network analysis, or (3) clustering & visualization [20 hours / The analysis assignment typically requires you to apply some of the techniques you've learned for one of the three topics (machine learning, social network analysis, or clustering & visualization) to a data set and write a report about the results and its analysis. More details about the actual assignment will be provided after each of these lectures]
- Brainstorm assignment (Putting CI into practice) [20 hours / The goal of this exercise is brainstorm about how the methods and techniques you've learned about in the course can be applied to an existing website to improve it. For the first part of the brainstorm assignment each group is required to write a 2-3-page report describing their proposed improvement(s) to a selected website. For the second part, each group has to critically peer-review another group's report. Finally, each group should incorporate the feedback they receive from another group. More details to follow in class.]
- Collective course evaluation [1 hour / The final active participation requirement is participating in the collective course evaluation. More details to follow in class]

Module description

Placement: 7th semester

Module coordinator: Florian Meier

Type and language: Elective / English

Objectives

At the end of the module, students are expected to be able to:

Knowledge

- understand the phenomenon of 'wisdom of the crowds', and its success conditions
- be familiar with techniques and algorithms for harnessing collective intelligence automatically, such as social network analysis, machine learning for online matchmaking, clustering & visualization of large data sets, and recommender systems

- explain how the Web 2.0 paradigm enables the application of collective intelligence
- identify and compare the basic techniques available for harnessing the collective intelligence represented large groups of users
- discuss and apply central hypotheses, theories, concepts, methods, and processes involved in using collective intelligence approaches

Competences

- relate theories and concepts and methods from collective intelligence to real world cases

Academic content and basis

The objective of the course is to present and utilize the concept of collective intelligence: combining the behavior, preferences, or ideas of a group of people to create novel insights. Students will learn how to interpret and organize the enormous amounts of user-generated content on the Web to produce novel insights about user experience, marketing, personal tastes, and human behavior in general. We will focus especially on a range of different sophisticated techniques and algorithms for harnessing such collective intelligence automatically. Example topics include:

- Wisdom of the crowds (Why and when is the crowd smarter than individuals?)
- Sentiment analysis (How can we automatically determine the sentiment expressed in text?)
- Online matchmaking (How can we use machine learning to automatically match people in online dating sites)?
- Clustering & visualization (How can we detect and visualize groups of similar items in a large data set?)
- Recommender systems (How can we recommend new and interesting products or media to users based on their past interests and those of others?)

Teaching will be a combination of lectures and lab sessions, where the students will apply some of these techniques to real-world data sets.

Scope and expectation

The module comprises 5 ECTS, which corresponds to 137.5 working hours. Teaching will be a combination of 20 hours of lectures and lab sessions.

Literature

	Required (pages)	Supplementary (pages)	Digital upload
1. Introduction + Wisdom of crowds			
Malone, T.W., Laubacher, R., and Dellarocas, C. (2010). The Collective Intelligence Genome. IEEE Engineering Management Review, 38(3), 21-31.	11		
Woolley et al. (2010). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. Science, 330(6004), 686-688.	3		
Page, S. E. (2007). The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton: Princeton University Press.		456	
Surowiecki, J. (2005). The Wisdom of Crowds. New York: Anchor Books		336	
2. Online matchmaking / Machine learning			
Abbott, D. (2014). Chapter 8: Predictive Modeling. In Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst (pp. 213-222). Indianapolis, IN: Wiley.	10		
Rokach, L., & Maimon, O. (2015). Chapter 1: Introduction to Decision Trees. In Data Mining with Decision Trees: Theory and Applications (2nd ed., Vol. 81, Series in Machine Perception and Artificial Intelligence, pp. 1-16). London: World Scientific Publishing.	16		
Rokach, L., & Maimon, O. (2015). Chapter 2: Training Decision Trees. In Data Mining with Decision Trees: Theory and Applications (2nd ed., Vol. 81, Series in Machine Perception and Artificial Intelligence, pp. 17-21). London: World Scientific Publishing.	5		
Yee, S., & Chu, T. (2015, July 27). A visual introduction to machine learning. Retrieved August 19, 2017, from http://www.r2d3.us/visual-intro-to-machine-learning-part-1/	5		
Yee, S., & Chu, T. (2018, June 18). A visual introduction to machine learning - part II: Model Tuning and the Bias-Variance Tradeoff. Retrieved June 19, 2018, from http://www.r2d3.us/visual-intro-to-machine-learning-part-2/	7		
Akser, M. (2015). Old and New Methods for Online Research: The Case of Online Dating. In I. A. Degim, J. Johnson, & T. Fu (Eds.), Online Courtship: Interpersonal Interactions Across Borders (Vol. 16, Theory on Demand, pp. 29-35). Amsterdam: Institute of Network Cultures.		7	
Fu, T. (2015). What are the Shengnv Looking for in Online Heterosexual Dating and Courtship? A Content Analysis of Shanghainese Women's Personal Profiles on Jiayuan.com. In I. A. Degim, J. Johnson, & T. Fu (Eds.), Online Courtship: Interpersonal Interactions Across Borders (Vol. 16, Theory on Demand, pp. 186-207). Amsterdam: Institute of Network Cultures.		22	
Mitchell, T. M. (1997). Chapter 3: Decision tree learning. In Machine Learning (pp. 52-60). New York: McGraw-Hill.		9	
Rokach, L., & Maimon, O. (2015). Chapter 7: Popular Decision Trees Induction Algorithms. In Data Mining with Decision Trees: Theory and Applications (2nd ed., Vol. 81, Series in Machine Perception and Artificial Intelligence, pp. 77-83). London: World Scientific Publishing.		7	
Reichert. R. (2015). Dating Maps: Mapping Love in Online Dating Communities. In I. A. Degim, J. Johnson, & T. Fu (Eds.), Online Courtship: Interpersonal Interactions Across Borders (Vol. 16, Theory on Demand, pp. 36-46). Amsterdam: Institute of Network Cultures.		11	
Segaran, T. (2007). Programming Collective Intelligence. Beijing: O'Reilly Media.		362	
3. Clustering & visualization			
Tan, P.-N., Steinbach, M. & Kumar, V. (2006). Chapter 8: Cluster Analysis: Basic Concepts and Algorithms. In Introduction to Data Mining (pp. 487-568). Harlow: Addison Wesley.	82		

	Required (pages)	Supplementary (pages)	Digital upload
4. Sentiment analysis			
Wright, A. (2009). Our Sentiments, Exactly. Communications of the ACM, 52(4), 14-15.	3		
TBA			
5. Social network analysis			
Golbeck, J. (2013). Analyzing the Social Web, chapters 1-5. Boston, MA: Morgan-Kaufmann.	74		
Haythornthwaite, C. (1996). Social Network Analysis: An Approach and Technique for the Study of Information Exchange. Library & Information Science Research, 18(4), 323-342.	20		
Gol Fan, W., & Gordon, M. D. (2014). The Power of Social Media Analytics. Communications of the ACM, 57(6), 74-81.		8	
Golbeck, J. (2013). Analyzing the Social Web, chapters 7 + 10. Boston, MA: Morgan-Kaufmann.		34	
Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social Media Analytics: An Interdisciplinary Approach and Its Implications for Information Systems. Business & Information Systems Engineering, 6(2), 89–96.		8	
6. Recommender systems			
Resnick, P. and Varian, H.R. (1997). Recommender Systems. Communications of the ACM, 40(3), 56-58.	3		
Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (2011). Introduction to Recommender Systems Handbook (chapter 1). In Recommender Systems Handbook (pp. 1-29). Berlin: Springer.	29		
Totals			
	268	1260	