IAB-Workshop
Nutzerschulungen: Bedarf und Umsetzung bei Datenanbietern – Beispiele aus der Praxis
12. April 2018
IPSDS project
We are pleased to announce the launch of the International Program in Survey and Data Science (IPSDS). Fundamental changes in the nature of data, their availability, the way in which they are collected, integrated, and disseminated are a big challenge for all those working with designed data from surveys as well as organic data. IPSDS was developed in response to the increasing demand from researchers and practitioners for the appropriate methods and right tools to face these changes. We offer a multidisciplinary curriculum, world-class faculty, and a web-based learning environment that allows you to take courses from anywhere in the world.
Key elements:

- Multidisciplinary curriculum
- Modularized – adapt to prior skills and work needs
- Mix of faculty from academia and industry
The project on which this report is based was promoted with funds from the Federal Ministry of Education and Research under the reference number [16OH22064]. Responsibility for the contents of this publication lies with the author.
IPSDS curriculum
Problems we tried to solve – in brief

Key elements:
- Multidisciplinary curriculum
- Modularized – adapt to prior skills and work needs
- Mix of faculty from academia and industry
Data Generating Process

- Research Question
- Fundamentals of Survey and Data Science 6 ECTS

Data Curation/Storage

- min. 6 ECTS
- Database Management I-III 2 ECTS each
- Data Munging I-III 2 ECTS each

Data Analysis

- min. 10 ECTS
- Generalized Linear Models 4 ECTS
- Analysis of Complex Data I-III 2 ECTS each
- Propensity Score/Statistical Matching 4 ECTS
- Machine Learning I-III 2 ECTS each
- Item Nonresponse and Imputation 2 ECTS

Data Output/Access

- min. 6 ECTS
- Ethics 2 ECTS
- Data Confidentiality and Statistical Disclosure Control 4 ECTS
- Visualization I-II 2 ECTS each

Total: 75 ECTS
Master Thesis: 15 ECTS
IPSDS structure
Problems we tried to solve – in brief

Key elements:
- Flexible web-based learning environment
- Live (video) interaction with faculty and students
- Face-to-face networking meetings
Onsite/Online

Onsite (Connect@IPSDS)

Online
Asynchronous/Synchronous

Asynchronous

• Pre-recorded lectures
• Readings/Other materials
• Assignments/Quizzes
• Discussion forums

Synchronous

• Small virtual classrooms
• Weekly discussions led by the instructor
• Obligatory component
IPSDS courses
Advantages of Flipped Teaching

- more opportunities for interactivity in (online) discussions
- more personalized guidance
- more time for feedback
- deeper learning

Source: Derek Bruff at https://goo.gl/Nrt1xA
Further Design Features

Pacing of the instruction

• Course length: 4 to 12 weeks (2 to 6 ECTS)
• Theme based units per calendar week
• Weekly assessment of topic/progress

Assessment

• Submit questions for online discussions
• Quiz
• Programing assignment
• Presentation
• Etc.

Feedback Mechanisms

• Automated feedback in quizzes
• Detailed from instructor for assignments and in online discussions
• Peer to peer feedback in some classes

Level of Interaction

• Weekly meetings
• Feedback for assignments
• Discussion forum and/or e-mails
• Group assignment
Welcome to SURV 736!

This short course provides a condensed overview of web technologies and techniques to collect data from the web in an automated way. To this end, students will use the statistical software R. The course introduces fundamental parts of web architecture and data transmission on the web. Furthermore, students will learn how to scrape content from static and dynamic web pages and connect to APIs from popular web services. Finally, practical and ethical issues of web data collection are discussed.

Course and Learning Objectives

- have an overview of state-of-the-art research that draws on web-based data collection,
- have a basic knowledge of web technologies,
- be able to assess the feasibility of conducting scraping projects in diverse settings,
- be able to scrape information from static and dynamic websites as well as web APIs using R,
- and be able to tackle current research questions with original data in their own work.

Before you start working on Unit 1, please make sure to thoroughly read the syllabus.

New units auto-display each week. Each unit includes:

- Readings (Note reference to book chapter, URLs, PDFs)
- Slides
- Lecture videos
- (Link to external resources)
- (Additional material)
- Zoom link for online meeting + date and time
- Discussion forum for submitting questions/student-instructor interaction
- Homework
- Quiz (autograded)
- Assignment submission (time restrictions)
Moodle

- Course description/General Information
  Topics covered, syllabus, additional resources

- New units auto-display each week. Each unit includes:
  - Readings (Note reference to book chapter, URLs, PDFs)
  - Slides
  - Lecture videos
  - (Link to external resources)
  - (Additional material)
  - Zoom link for online meeting + date and time
  - Discussion forum for submitting questions/student-instructor interaction
  - Homework
    - Quiz (autograded)
    - Assignment submission (time restrictions)
Moodle

- **Course description/General Information**
  Topics covered, syllabus, additional resources

- **New units auto-display each week. Each unit includes:**
  - Readings (Note reference to book chapter, URLs, PDFs)
  - Slides
  - Lecture videos
  - (Link to external resources)
  - (Additional material)
  - Zoom link for online meeting + date and time
  - Discussion forum for submitting questions/student-instructor interaction
  - Homework
    - Quiz (autograded)
    - Assignment submission (time restrictions)
• Lectures, interviews and discussions with experts, demonstrations of specific techniques and software tools

• Lectures are broken into easily-digestible sessions to help students to better focus on the material.

• Students engage with the material at their own pace: e.g., replay parts that cover difficult concepts
Virtual Classrooms

- Weekly mandatory online meetings (50 minutes)
- Discuss students’ questions
- Review problems with assignments
- Collaborative problem solving
- Motivate students to persist in the course
- Break out rooms, (private and public) chats, polls...
IPSDS students
2 Test Cohorts

- 31 Participants (18 f + 13 m)
- 15 countries of residence
- Age: median=30.5 (min-23; max-56)
Test Cohorts 1 & 2

- **22** (71%) M.A./M.S.

- **41** hours/week
  (Min-10; Max-55)

- **8** students with family duties
IPSDS evaluations & studies
Evaluations

- Start of the program/yearly milestone survey/qualitative interviews of students
- Post-course student survey
- Learning analytics
- Qualitative interviews with instructors
- Pilot studies
- Curriculum and courses review
BMBF-Study: Online Communication

Synchronous Communication
Fundamentals of Survey and Data Sc.

• 11 online discussions (mandatory)
• 16 students

Asynchronous Communication
Data Collections Methods

• 1 introductory online meeting
• Discussion forums
• 15 students

• 12 weeks/6 ECTS
• Post questions
• Answer/comment questions of other students
BMBF-Study: Online Communication

Synchronous Communication
Fundamentals of Survey and Data Sc.
- Stronger sense of community and belonging
- Immediate feedback
- All students successfully finished the course

Asynchronous Communication
Data Collections Methods
- Greater flexibility
- More time for reflection
- Less workload
- 2 drop outs
BMBF-Study: Flexibility

Instructor-Paced Format

- Weekly online meetings
- Set dates for all assignments and final exam

Self-Paced Format

- 4 ECTS/ 8 + 1 weeks
- 8 participants

- 1 introductory online meeting
- Only 1 deadline: all assignments and final exam due in the ninth week
BMBF-Study: Flexibility

Self-Paced Format

- 4 drop outs
- Flexibility with deadlines for assignments appreciated
- Biggest challenge: time management
IPSDS lessons learning
Lessons Learning

• Modular approach much appreciated by working professionals
• Feasibility of combining studies with work and family (6 ECTS/term, 3 drop outs in 2 years)
• Biggest challenge: workload management (average workload: 10 hours/week)
• Balancing flexibility and consistency (time management, self-discipline)
• Workplace orientation:
  • Hands-on application and working with data is key
  • Project-oriented courses/working on real-world projects
your QUESTIONS
#collaborations

- Business
- Economics
- Law
- Linguistics
- Mathematics
- Computer Science
- Political Science
- Psychology
- Sociology

Universities:
- University of Maryland
- Stockholm University
- Universidad de Chile
Universities

- University of Maryland
- University of Mannheim
- Catholic University of Santiago de Chile
- Australian National University
- Beijing University
- Ashoka University (expressed interest)
- University of Capetown (planned)

Others

- SRO - Michigan
- PEW
- German Record Linkage Center
- GESIS
- Bureau of Labour Statistics
- U.S. Census Bureau
- Statistics Netherlands
Was haben wir gelernt?

Durchschnittlich 6 ECTS/Term

IPSDS ermöglicht Weiterbildung und Familie zu koordinieren:
- 7 Teilnehmer/innen mit Kindern
- 3 Geburten während der ersten Förderphase

Kurs-Design berücksichtigt Motivationsförderung:
- synchrones Lernen
- Connect@IPSDS
- fest terminierter Kursrhythmus

Nur 3 Studienabbrecher (nach 1. Jahr)
IPSDS faculty

✓ Manfred Antoni
  (German Institute for Employment Research)
✓ Stefan Bender
  (Deutsche Bundesbank)
✓ Trent Buskirk
  (Vice President at Statistics and Methodology at Marketing Systems Group/University of Massachusetts Boston)
✓ Mario Callegaro
  (Google UK)
✓ Frederick Conrad
  (University of Maryland/University of Michigan)
✓ Jill Dever
  (RTI)
✓ Jörg Drechsler
  (German Institute for Employment Research/LMU Munich)
✓ Jeffrey Gonzalez
  (Bureau of Labor Statistics OSMR)
✓ Steven Heeringa
  (University of Maryland/University of Michigan)
✓ Matt Jans
  (UCLA Center for Health Policy Research/University of Maryland/University of Connecticut)
✓ James Lepkowski
  (University of Maryland/University of Michigan)
✓ Daniel Oberski
  (Utrecht University)
✓ Louis Rizzo
  (Westat)
✓ Jennifer Romano-Bergstrom
  (Facebook)
✓ Joseph Sakshaug
  (University of Michigan/German Institute for Employment Research)
✓ Richard Valliant
  (University of Maryland/University of Michigan)
Data

- Designed
  - Experiment
  - Survey
- Organic
  - Administrative
  - Aspirational
  - Transactional

Source: Roberto Rigobon
DOMIAN EXPERT
User, analyst, or leader with deep subject matter expertise related to the data, its appropriate use, and its limitations

RESEARCHER
Team member with experience applying formal research methods, including survey methodology and statistics

SYS ADMIN
Team member responsible for defining and maintaining a computation infrastructure that enables large scale computation

COMPUTER SCIENTIST
Technically skilled team member with education in computer programming and data processing technology

Usher 2015
Learn how to communicate results, distribute and store your data; Ethics

Learn a variety of analysis methods suited for different data types

Learn how to curate and manage data

Understand how to collect data yourself, and how data are generated through administrative and processes.

Learn how to formulate your research goal and which data are best suited to achieve this goal.