JOEL BECKER

PERSONAL INFORMATION

email	mail@joelbecker.ca
website	joelbecker.ca
phone	(519) 502 1517

SKILLS

Data Analysis
Discovering powerful, reproducible insights using statistical inference and data visualization in Python and R.

Software Development
Designing flexible and coder-friendly Python packages for data extraction and data integration.

Research Design
Developing research questions, methodologies, and implementation strategies to understand complex social phenomena.

Data Preparation
Extracting data from text corpora and APIs. Integrating data from disparate sources. Data cleaning and normalization.

Technical
Experienced Python and R programmer. Also proficient in C++, SQL, and Matlab, with basic knowledge of HTML.

WORK EXPERIENCE

2016–Present   **Software Developer and Research Technician**

NetLab
Developed Python packages that have enabled our interdisciplinary research team's leading-edge work in large-scale collaboration networks by streamlining the process of extracting and integrating data from multiple sources. Contributions include expanding features and increasing performance of existing tools; developing new Python packages; implementing data extraction and integration solutions; and developing coding tutorials.
Reference: John McLevey, Assistant Prof. · john.mclevey@uwaterloo.ca

2013–2017   **Research Assistant**

Conrad Business, Entrepreneurship, and Technology Centre
Collaborated with Dr. Mark Weber on all aspects of research in behavioural ethics and online self-presentation. Contributions include supplementary data collection using social media APIs, as well as developing hypotheses, study materials, and research methodology.
Reference: Mark Weber, Director · mark.weber@uwaterloo.ca

2012–2016   **Staff**

Silver Lake Mennonite Camp
Held a variety of leadership positions involving: representing staff in capital, financial, and personnel decisions on the board of directors (*2016*), coordinating summer camp programming (*2016*), designing and implementing a two-part six-week leadership training program (*2015*), staff training (*2014*), and evaluating junior staff (*2014*). Other responsibilities included youth supervision, crisis management, and emergency preparedness.
Reference: Mark Weber, Chair of the SLMC Board of Directors · mark.weber@uwaterloo.ca

EDUCATION

2012-2017	Bachelor of Knowledge Integration
The University of Waterloo	JOINT HONOURS KNOWLEDGE INTEGRATION AND PSYCHOLOGY COMPUTER SCIENCE MINOR
Achievements	HONOURS THESIS · 97%
	MAJOR AVERAGES · 90% / 93%
	DEAN'S HONOURS LIST

RESEARCH

2016-2017	Do just actions beget just evaluations for supervisors with intersectional identities?

Honours Thesis My thesis investigated whether the race and gender of workplace supervisors changes how their subordinates react to fair or unfair treatment. In a large-scale multi-phase field survey of 400+ workers, I found that supervisors' race and gender do not affect how subordinates evaluated supervisors who they perceive treat them with dignity and respect; that white women and black men are evaluated unusually harshly when their subordinates feel poorly treated; but that black female supervisors do not experience this kind of bias. My findings added to our growing understanding of the complex relationship between race, gender, and workplace outcomes.

SUPERVISORS · Professor Ramona BOBOCEL & Assistant Professor Winny SHEN

SOFTWARE DEVELOPMENT

2017	`recordlinkage`

Python Package `recordlinkage` is a powerful open-source toolkit for intelligently matching and merging data from multiple sources. I dramatically improved performance by identifying inefficiencies in a major performance bottleneck, making several key algorithms *up to 97 times faster*. A collaborator and I also contributed important new features, designing and implementing an all-new workflow for fusing matched records and resolving conflicting data, as well as implementing two new record comparison algorithms.

2017	`tidyextractors`

Python Package I created the `tidyextractors` package to simplify data extraction for both coders in a hurry and researchers new to Python. The current version makes it easy to extract ready-to-analyze datasets from `git` repositories, `mbox` mail archives, and Twitter user data via the Twitter API, with end-to-end `pandas` support. The tool is built to be extended, with modules for extracting data from FACTIVA queries and GitHub issues planned for the near future.

OTHER INFORMATION

Languages	ENGLISH · Native Lanuage
	FRENCH · Intermediate

Interests	Da Capo Chamber Choir · Exhibit Design · Piano · Volleyball

August 23, 2017