

SOCIOMETRICS AND HUMAN RELATIONSHIPS

Analyzing Social Networks
to Manage Brands,
Predict Trends, and
Improve Organizational
Performance

SOCIOMETRICS AND HUMAN RELATIONSHIPS

Analyzing Social Networks
to Manage Brands,
Predict Trends, and
Improve Organizational
Performance

BY

PETER A. GLOOR

*MIT Center for Collective Intelligence,
Massachusetts Institute of Technology,
Cambridge, MA, USA*



United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2017

Copyright © 2017 Peter A. Gloor

Reprints and permissions service
Contact: permissions@emeraldinsight.com

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-1-78714-113-1 (Print)
ISBN: 978-1-78714-112-4 (Online)
ISBN: 978-1-78714-725-6 (Epub)



ISOQAR
REGISTERED

Certificate Number 1985
ISO 14001

ISOQAR certified
Management System,
awarded to Emerald
for adherence to
Environmental
standard
ISO 14001:2004.



INVESTOR IN PEOPLE

CONTENTS

Acknowledgments xi

1. Introduction	1
1.1. Roadmap	6
1.2. Key Takeaways of This Book	15
1.3. Study Plan for a One-Semester Course	18
1.4. Sample Course Syllabus	20

PART I. TREND PREDICTION BY MEASURING SOCIAL NETWORKS

2. Coolfarming Organizations	27
2.1. Knowledge Flow Optimization through Organizational Social Network Analysis	29
2.2. The Coolfarming Data Collection and Analysis Process	31
3. Coolhunting and Trend Forecasting on the Web	37
3.1. Measuring Collective Awareness	37
3.2. The Coolhunting Process — Finding Trends by Finding Trendsetter	39
4. The Six Honest Signals of Collaboration	45
4.1. The Honest Signals Have Different Meanings for Different Organizations	54
4.2. Virtual Mirroring Leads to Change	56
4.3. Dealing with Privacy Concerns	56

4.4.	How to Apply Knowledge Flow Optimization	58
4.5.	Four Examples	61
4.6.	Areas of E-Mail-Based SNA	64
4.7.	Improving Financial Capital through Optimizing Social Capital	65
5.	Essentials of Social Network Analysis and Statistics	69
5.1.	Basics of Social Network Analysis (SNA)	70
5.2.	Basics of Statistics	75
6.	How Ideas Spread in Online Social Networks — Readings	85
6.1.	Theories of Information Diffusion	89
6.2.	Spreading Ideas on Facebook	95
6.3.	Finding Fake Reviews through Machine Learning	96
6.4.	Measuring Financial Performance	97
6.5.	Calculating Demographic Information	99
6.6.	Predicting Election Outcome	103
 PART II. ANALYZING STRUCTURE, DYNAMICS, AND CONTENT OF NETWORKS WITH CONDOR 		
7.	The Four-Step Analysis Process	111
7.1.	Social Media Fetchers	115
7.2.	Social Media Filters	116
7.3.	Social Media Visualizers	116
7.4.	Social Media Exporters	118
8.	Getting Started with Condor	121
8.1.	Analyzing the Facebook Wall with Condor	125

8.2.	Sample Four-Step Analysis with Twitter	130
8.3.	Measuring the Importance of Brands through Betweenness of Actors in Bipartite Graphs	136
8.4.	Pruning the Leaves in a Graph	137
8.5.	Degree-of-Separation Search with Google CSE	141
8.6.	Degree-of-Separation Search with Twitter	146
8.7.	Wikipedia Search	150
9.	Analyzing E-Mail with Condor	153
9.1.	Creating a Virtual Mirror of Your Own Mailbox	154
9.2.	Finding COINs through Community Detection	185
9.3.	Creating a Virtual Mirror of an Organization	192
9.4.	Analyzing Hillary Clinton's Mail	219
9.5.	Organizational Aspects of E-Mail-Based SNA	228
9.6.	Follow-on Exercises	232
9.7.	(Partial) List of E-Mail Studies Conducted by the Author in Various Organizations	232
10.	Calculating Personality Characteristics from E-Mail	241
10.1.	Calculating Correlations between FFI and E-Mail	242
10.2.	Developing a General Prediction Formula	244
10.3.	Adding Gender, Ethnicity, and Nationality as Control Variables	254
10.4.	Follow-on Exercises	260

11. Predicting Criminal Intent from E-Mail — Analyzing the Enron E-Mail Archive	263
11.1. Exploratory Analysis	264
11.2. Identifying Criminal Actors through Their Honest Signals of Collaboration	273
11.3. “Tribefinder” — Identifying Criminals through Machine Learning in Condor	280
11.4. Follow-on Exercises	291
12. Coolhunting on the Internet with Condor	295
12.1. Expert Analysis — Websites and Blogs	298
12.2. Swarm Analysis — Wikipedia	311
12.3. Analysis of the Crowd — Twitter	322
12.4. Follow-on Exercises	334
12.5. (Partial) List of Internet Coolhunting Studies	335
13. Coolhunting — Francogeddon	339
13.1. Follow-on Exercises	347
14. Coolhunting the US Presidential Elections	349
14.1. Bernie Sander’s Presidential Campaign — The Perfect COIN	353
14.2. Coolhunting Bernie Sanders, Hillary Clinton, Jeb Bush, and Donald Trump	356
14.3. Tribefinder on Twitter (Using Machine Learning)	366
14.4. Follow-on Exercises	383

**PART III. AUTOMATIC MEDIA INSIGHTS
COIN ASSESSMENT (AMICA)**

15. Inside Media Individual Collaboration (IMIC)	391
15.1. IMIC Annotation Process	401

16. Outside Media Individual Collaboration (OMIC)	405
16.1. OMIC Annotation Process	414
17. Inside Media Organizational Collaboration (IMOC)	419
17.1. IMOC Annotation Process	423
18. Outside Media Organizational Collaboration (OMOC)	425
18.1. OMOC Annotation Process	429
18.2. Follow-on Exercises	430
19. Survey of Individual and Organizational Collaboration (SIC & SOC)	431
19.1. Survey of Individual Collaboration (SIC)	431
19.2. Survey of Organizational Collaboration (SOC)	439
19.3. Sample Download	444

PART IV. APPENDIX — USEFUL MACHINE LEARNING AND GRAPH ANALYSIS TOOLS

<i>Appendix A: Identifying Anti-Vaxxers through Machine Learning Using KNIME</i>	447
<i>Appendix B: Generating Nice Graph Pictures with Gephi</i>	459
<i>Appendix C: Sample Mid-Term Exam</i>	465
<i>Appendix D: References</i>	469
<i>Biography</i>	483
<i>Index</i>	485

ACKNOWLEDGMENTS

The tools and methods described in this book have been developed and tested over the last 12 years in the Collaborative Innovation Networks (COINs) seminar. I am deeply grateful to all my instructor colleagues, and of course to the hundreds of students from the United States, Finland, Germany, Switzerland, Chile, Italy, South Korea, and China who have contributed many creative ideas, and have taught me what works, and what does not.

The COINs seminar was started at MIT Sloan in spring 2005. In fall of the same year, the seminar morphed into a virtual distributed course joined by students from Helsinki, supervised by Maria Paasivaara and Casper Lassenius, students from Cologne tutored by Detlef Schoder and Kai Fischbach, and students from Savannah College of Art and Design (SCAD) lectured by Christine Z. Miller. In the meantime, the seminar has also repeatedly been taught at Pontificia Universidad Catolica Santiago de Chile coached by Cristobal Garcia Herrera, and University of Applied Sciences Northwestern Switzerland, where Michael Henninger has been the indispensable instructor. Since 2011, the students from Cologne have been coached first by Johannes Putzke, and since 2014 by Gloria Volkmann, while at

University of Bamberg, students have been instructed by Kai Fischbach and Matthaeus Zylka.

The software tool Condor that is the basis of this course was started in 2003, when the Center for Digital Strategies at Dartmouth College under the leadership of Hans Brechbühl and Eric Johnson agreed to support Yan Zhao's software development efforts as part of her Master's thesis supervised by Fillia Makedon. For the next three years, Yan, ably supported by the algorithm genius of her husband Song Ye, built the first two versions of Condor, originally called TecFlow. End of 2006, she passed the baton to Renauld Richardet, who added Apache Lucene's text processing capabilities. In 2008, Condor development continued in Switzerland at galaxyadvisors, funded by the Swiss Commission for Technology and Innovation CTI. Michael Henninger, Hauke Fuehres, Martin Stangl, Lucas Broennimann, Marton Makai, and Kevin Zogg from the University of Applied Sciences Northwestern Switzerland (FHNW) worked on building a fundamentally revised version of Condor in the team of Manfred Vogel and André Csillaghy at the Institute for 4D technologies i4ds. Since 2013, Condor development is done by my colleagues at galaxyadvisors, Marton Makai, Hauke Fuehres, and Joao Marcos Da Oliveira, supported from 2014 to 2015 by Karsten Packmohr.

This book is the product of many people working together over 14 years, building the tools and methods described here. First of all, I am grateful to Ken Riopelle and Michael Henninger, who have been essential in making the social media analysis tool Condor accessible to a wider audience beyond programmers and

statisticians. Ken created the first Condor videos, and wrote a comprehensive manual, the precursor of this book. Michael wrote the first tutorial for Condor in the COINs seminar at University of Applied Sciences Northwestern Switzerland. Ken Riopelle, Michael Henninger, and Lucas Broennimann provided valuable feedback on earlier versions of this manuscript. Ken also contributed the last section of Chapter 3 of Part II.

My sincerest thanks to all of you, without your creative ideas, didactical talent, and Java development and software architecting skills, both the COINs course and Condor would not exist.

INTRODUCTION

Imagine being able to spot if a customer is becoming really unhappy with your product and service — and do something about it before they actually leave you.

Imagine finding out what the constituency of a politician or political party really thinks.

Imagine finding out what your customers love and hate about your product.

Imagine being able to identify your most creative employees, your external innovators, and lead users — and help them become even more creative.

Imagine being able to predict who wants to leave your company, your department, or your project team — and not just identify them, but help them become happy and motivated workers again.

Imagine identifying potentially fraudulent or risky behavior among your employees before they actually commit anything illegal.



If you are looking for answers to these and similar questions, read on. This book gives you a framework to analyze your organization from the inside, by mining e-mail, skype, and calendar data, and from the outside, by crunching Twitter, Wikipedia, and blog data.

From your and your organization's e-mail, skype, and calendar data, you can:

- Find out about the happiness of your employees (see Section 9.3).
- Find out about the satisfaction of your customers (see Section 9.3).
- Find out who might be leaving your company (see Section 9.3).
- Find your most creative and motivated employees (see Chapter 10).
- Find out about the willingness of your employees to take unnecessary risks (see Section 11.3).

From Twitter, Wikipedia, and blog interaction data, you can:

- Find out about what your customers and prospects really think about your company and your brands (see Chapter 12).
- Measure the strength of your brand (see Chapter 12).

- Find out about the demographic profile of the customers and aficionados of your company and brands (see Section 14.3).
- Forecast the popularity and voter share of a politician (see Section 14.2).
- Find out about the demographic profile of the voters of a politician (see Section 14.3).

These are just a few use cases that we will address to study how humans communicate and collaborate inside the organization, through e-mail, chat, videoconferencing, and face-to-face communication, and outside on online social media. Better communication leads to better collaboration, which leads to more and better innovation! This book describes algorithms and tools to find and support collaboration within and between organizations. Our approach puts a lens to the organization by mining electronic communications such as e-mail, sociometric badges, telephone, chat, online meeting, Web/videoconferencing, and calendars to make existing communication patterns visible. The Condor software tool, which has been developed over the past decade at the MIT Center for Collective Intelligence and the University of Applied Sciences Northwestern Switzerland, mines these electronic archives and generates a broad range of structural, temporal, and content-based social network metrics which can be used to calculate and forecast all of these real-world insights mentioned above ([Figure 1](#)).

This book provides a practical guide to Coolhunting and Coolfarming on online social media. It explains how to “Coolhunt” — to find cool trends by finding the trendsetters on Twitter, Facebook, Wikipedia, blogs,