

# Amaç Herdağdelen

herdagdelen.com



Researcher with over 15 years of experience in computational social science, machine learning, and natural language processing. Developed humanitarian datasets used by international organizations. Built large-scale data pipelines and machine learning models. Driven by a curiosity about methods, seeking novel ways to make sense of complex data.

## Positions

<b>Senior Research Scientist</b> , Core Data Science at Meta (Facebook)	2012–2022
<b>Senior Research Scientist</b> , Crimson Hexagon	2011–2012
<b>Postdoctoral Fellow</b> , New England Complex Systems Institute	2011
<b>Research Intern</b> , Google	2009
<b>Co-founder</b> of Çilek Ağacı, an Information Design Collective	2009–present

## Projects

### SOCIAL IMPACT

#### Created Movement Range Maps during COVID-19

- Meta's most downloaded humanitarian dataset
- Measured local mobility to verify stay-at-home policy effectiveness
- Informed the pandemic response of CDC, California, EU Commission, and others

#### Informed Community and Product Strategy

- Developed high-resolution social capital maps, using online engagement
- Analyzed immigrant community integration via friendship networks
- Developed social cohesion metrics for online communities
- Analyzed sharing and reciprocity in Facebook gifting groups

### LANGUAGE & AI

#### Owned language classification service at Facebook (2013-2015)

- Developed and deployed near-real-time language detection service for all Facebook content
- Scaled service infrastructure across global data centers
- Built evaluation frameworks and benchmarking methodology

### **Created large-scale social media language corpora with demographic metadata**

- Provided public resources for sociolinguistic research
- Demonstrated improved word recognition prediction in psycholinguistic studies

### **Early research on gender bias in AI systems**

- Analyzed gender stereotypes in web-derived corpora
- Developed methods to measure bias in word associations

## **NETWORKS & COMMUNITIES**

### **Identified coordinated harmful behavior on Facebook**

- Built graph analysis algorithms and embeddings to detect malicious networks

### **Improved Instagram account recommendations**

- Developed algorithms to identify subject matter experts
- Applied expertise signals to improve discovery

## **DATA VISUALIZATION**

**Co-founded Çilek Ağacı**, an information design collective working on popular and social issues

- Created data journalism projects featured in national media
- Received international awards for information design

# **Education**

**Ph.D.** | University of Trento

2007–2011

- Cognitive and Brain Sciences Doctoral School
- Thesis: "Collecting Common Sense from Text and People"

**M.S.** | Boğaziçi University

2004–2007

- Computer Engineering, GPA: 3.76 / 4.00
- Thesis: "A Cultural Market Model"

**B.S.** | Boğaziçi University

2000–2004

- Computer Engineering, GPA: 3.34 / 4.00

# **Awards and Patents**

**Best Doctoral Thesis Award**, University of Trento

**Bronze medal in Data Journalism** (Information is Beautiful Awards)

**11 granted patents** in community detection, content classification, search query understanding, language identification

# Publications

## COMPUTATIONAL SOCIAL SCIENCE

- Herdağdelen, A., Adamic, L. (2023). The Geography of Facebook Groups in the United States. *Proceedings of the International AAAI Conference on Web and Social Media*, 17, 351-362.
- Herdağdelen, A., Adamic, L., State, B. (2022). Community Gifting Groups on Facebook. *arXiv preprint arXiv:2211.09043*.
- Lee, M., Herdağdelen, A., Park, M., Martin, J.L. (2021). Measuring Mohr Social Capital. *Poetics*, 88, 101596.
- Kraut, R.E., Levine, J.M., Escobar, M.M., Herdağdelen, A. (2020). What Makes People Feel Close to Online Groups? The Roles of Group Attributes and Group Types. *Proceedings of the International AAAI Conference on Web and Social Media*, 14, 382-392.
- Herdağdelen, A., Dow, A., Bogdan, S., Payman, M., Pompe, A. (2020). Protecting Privacy in Facebook Mobility Data During the COVID-19 Response. *Facebook Research*.
- Herdağdelen, A. (2017). Personal Names in Modern Turkey. *arXiv preprint arXiv:1801.00049*.
- Herdağdelen, A., State, B., Adamic, L., Mason, W. (2016). The Social Ties of Immigrant Communities in the United States. *Proceedings of the 8th ACM Conference on Web Science*, 78-84.
- Herdağdelen, A., Zuo, W., Gard-Murray, A., Bar-Yam, Y. (2013). The Geography and Politics of News-Sharing Communities in Twitter. *Complexity*, 19(2), 10-20.
- Herdağdelen, A., Zuo, W., Gard-Murray, A., Bar-Yam, Y. (2012). An Exploration of Social Identity: The Geography and Politics of News-Sharing Communities in Twitter. *arXiv preprint arXiv:1202.4393*.
- Herdağdelen, A., Bingol, H. (2009). A Cultural Market Model. *International Journal of Modern Physics C*, 19(2), 271.
- Tasgin, M., Herdağdelen, A., Bingol, H. (2008). Community Detection in Complex Networks Using Genetic Algorithms. *arXiv preprint arXiv:0711.0491*.
- Herdağdelen, A., Aygün, E., Bingol, H. (2007). A Formal Treatment of Generalized Preferential Attachment and Its Empirical Validation. *Europhysics Letters*, 78(6), 60007.
- Herdağdelen, A., Aygün, E., Bingol, H. (2009). Measuring Preferential Attachment in a Hyper-Textual Dictionary Reference Network: Ekşi Sözlük.

## AI & LANGUAGE MODELS

- Herdağdelen, A., Marelli, M. (2017). Social Media and Language Processing: How Facebook and Twitter Provide the Best Frequency Estimates for Studying Word Recognition. *Cognitive Science*, 41(4), 976-995.
- Herdağdelen, A. (2013). Twitter N-gram Corpus with Demographic Metadata. *Language Resources and Evaluation*, 47, 1127-1147.
- Ciaramita, M., Herdağdelen, A., Mahler, D. (2013). Generalized Edit Distance for Queries. *US Patent 8,417,692*.

- Herdağdelen, A., Baroni, M. (2012). Bootstrapping a Game with a Purpose for Commonsense Collection. *ACM Transactions on Intelligent Systems and Technology*, 3(4), 1-24.
- Herdağdelen, A., Baroni, M. (2011). Stereotypical Gender Actions Can Be Extracted from Web Text. *Journal of the American Society for Information Science and Technology*, 69(2), 1741-1749.
- Herdağdelen, A. (2011). Collecting Common Sense from Text and People. University of Trento. PhD Thesis.
- Herdağdelen, A., Ciaramita, M., Mahler, D., Holmqvist, M., Hall, K., Riezler, S., Alfonseca, E. (2010). Generalized Syntactic and Semantic Models of Query Reformulation. *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 283-290.
- De Bona, F., Riezler, S., Hall, K., Ciaramita, M., Herdağdelen, A., Holmqvist, M. (2010). Learning Dense Models of Query Similarity from User Click Logs. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 474-482.
- Herdağdelen, A., Baroni, M. (2010). The Concept Game: Better Commonsense Knowledge Extraction by Combining Text Mining and a Game with a Purpose. *Proceedings AAAI Fall Symposium on Commonsense Knowledge, Arlington*.
- Herdağdelen, A., Erk, K., Baroni, M. (2009). Measuring Semantic Relatedness with Vector Space Models and Random Walks. *Proceedings of the 2009 Workshop on Graph-based Methods for Natural Language Processing*, 50-53.
- Herdağdelen, A., Baroni, M. (2009). BagPack: A General Framework to Represent Semantic Relations. *Proceedings of the Workshop on Geometrical Models of Natural Language Semantics*, 33-40.

## **MACHINE LEARNING & DATA SCIENCE APPLICATIONS**

- Weinsberg, E., Herdağdelen, A., Dimson, T.F., Xiao, B., de Sa Resende, D.T. (2022). Systems and Methods for Identifying and Grouping Related Content Labels. *US Patent 11,263,239*.
- Pal, A., Herdağdelen, A., Chatterji, S., Taank, S., Chakrabarti, D. (2020). Systems and Methods for Recommendation of Topical Authorities. *US Patent 10,803,068*.
- Herdağdelen, A., Boucher, M., Weinsberg, E. (2020). Systems and Methods to Predict Hashtags for Content Items. *US Patent 10,698,945*.
- Merl, D.M., Pal, A., Funiak, S., Park, S., Huang, F., Herdağdelen, A. (2019). Identifying Multiple Languages in a Content Item. *US Patent 10,180,935*.
- Chordia, S., Kai, R.E.N., Pal, A., Herdağdelen, A., Wang, T. (2018). Targeting Content Based on Inferred User Interests. *US Patent App. 15/482,447*.
- Herdağdelen, A., Green, B.R. (2018). Classifying Languages for Objects and Entities. *US Patent 10,002,131*.
- Herdağdelen, A., Green, B.R. (2017). Classifying Languages for Objects and Entities. *US Patent 9,740,687*.
- Firat, A., Brooks, M., Bingham, C., Herdağdelen, A., King, G. (2016). Systems and Methods for Calculating Category Proportions. *US Patent 9,483,544*.
- Pal, A., Herdağdelen, A., Chatterji, S., Taank, S., Chakrabarti, D. (2016). Discovery of Topical Authorities in Instagram. *Proceedings of the 25th International Conference on World Wide Web*, 1203-1213.

- Ciaramita, M., Herdağdelen, A., Mahler, D. (2016). Generalized Edit Distance for Queries. *US Patent 9,251,206*.
- Filiz, I.O., Herdağdelen, A., Chang, J. (2015). Inferring an Industry Associated with a Company Based on Job Titles of Company Employees. *US Patent App. 14/309,038*.
- Herdağdelen, A., Alpaydın, E. (2004). Dynamic Alignment Distance Based Online Signature Verification. *The 13th Turkish Symposium on Artificial Intelligence & Artificial Neural Networks*, 10-11.
- Yüçeturk, A.C., Herdağdelen, A., Uyanik, K. (1999). A Solution to the Problem of Local Minima in Backpropagation Algorithm. *Proceedings of the Fourteenth International Symposium on Computer and Information Sciences*, 1081-1083.