

## OVIDIU DAN

### PRESENT ADDRESS

10 Duh Dr, Apt 113  
Bethlehem, PA 18015  
+1 484 893 0253  
[contact@ovidiudan.com](mailto:contact@ovidiudan.com)

### PERMANENT ADDRESS

3 Govora, Tg-Mures  
Mures, Romania, 540031  
+40 265 260620  
[www.ovidiudan.com](http://www.ovidiudan.com)

### EDUCATION

[Lehigh University](#), Bethlehem, PA, USA  
PhD in Computer Science, expected 2014  
Working with Prof. Davison and the [WUME lab](#)

[INHOLLAND University](#), Amsterdam, The Netherlands

Bachelor of Engineering in IT, February 2009  
Graduated first in year, with an A on my thesis

### RESEARCH INTERESTS

Data mining in general, web mining in particular: social networks, personalization, ranking, online news collection and analysis, sentiment analysis, geographical named entities disambiguation, text clustering and classification, scalable crawling infrastructure, scaling data mining algorithms.

### RESEARCH/PROJECTS

- 2009 - Now Research in **social networks**: ranking, personalization, recommendations.
- 2011 Summer internship at **Microsoft Research** and **Bing**. My mentors were Pavel Dmitriev from Bing, Ryen White from MSR, and Xin Fu from Bing.
- 2007 - 2011 Conducting research in **mining online news** for [Newistic.com](#). I designed a scalable infrastructure that crawls, analyzes and stores online news articles in several languages. I wrote scalable algorithms for **web wrapper induction, distributed text clustering, classification, geographical named entities disambiguation, sentiment analysis, ranking**, and others.
- 2010 Internship at **AT&T Labs Research, Florham Park, NJ**. Worked with Junlan Feng and Bernard Renger on filtering social media messages for Social TV.
- 2007 - 2009 Wrote a **geographically aware DNS server**. Unlike standard DNS servers, it can redirect visitors to the best/closest mirror of a distributed web service based on their country of origin, city, ISP, datacenter load, and server uptimes.
- 2006 Built a **distributed data storage** framework. It was used as a storage component for data mining research conducted in the summer of 2006.

### PUBLICATIONS

- 2011 Ovidiu Dan, Junlan Feng, Bernard Renger, Book chapter: *Enhancing Social TV through Social Media Mining, TV Content Analysis*, CRC Press
- 2011 Ovidiu Dan, Junlan Feng, Brian Davison, *A Bootstrapping Approach to Identifying Relevant Tweets for Social TV, ICWSM 2011*
- 2011 Ovidiu Dan, Junlan Feng, Brian Davison, *Filtering microblogging messages for Social TV, WWW 2011*
- 2011 Liangjie Hong, Ovidiu Dan, Brian Davison, *Predicting Popular Messages in Twitter, WWW 2011, best poster award*
- 2009 Ovidiu Dan, Horatiu Mocian, *Scalable Web Mining with Newistic*, The 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-09), April 2009
- 2009 Ovidiu Dan, Horatiu Mocian, *Newistic: a distributed news gathering and analysis platform*. In Proceedings of the International Conference on Web Intelligent Systems '09, Chennai, India, January 2009
- 2006 Ovidiu Dan, Horatiu Mocian, *StorCH: A Distributed Self-Organizing Storage System*, 2006

## SELECTED WORK EXPERIENCE

- 2007 - Now **PHP/MySQL/Linux** - Wrote [geoscaling.com](#), a mission critical geographically-aware DNS server and built an infrastructure that spans three countries to host it.
- 2005 - 2011 **Java** - CEO and (from 2010) CTO at [Newistic.com](#), a startup that provides technology for collecting and analyzing online news.
- 2007 - 2009 **Linux administrator and programmer** for several clients from US and Romania.
- 2008 **PHP (Scalability) - Funded by Google** to work on an open source project via their [Google Summer of Code 2008](#) program. I extended the Coppermine Photo Gallery to allow flexible mirroring and distributed photo storage.
- 2008 **Plone/Python** - 5-month internship at the [Asian Human Rights Commission](#) in Hong Kong. I helped them develop their new website (not yet online) using the Plone platform, which is written in Python.
- 2008 - 2009 **Networking/Linux** - Built a gigabit network backbone for the "Papiu" high-school, Romania. I took care of networking equipment, wiring, and Linux server administration.
- 2002 - 2007 **PHP/MySQL** - CEO, CTO and eventually consultant for Playfuls Entertainment. During my time with the company I wrote a web site and I created a distributed web serving infrastructure that could handle 3 million visitors / month. I sold the company in 2006 to pay for college, but remained as a technical consultant until 2007. The website had 3 million visitors / month at its peak.
- 2002 - 2007 **PHP/MySQL/Linux** - Wrote the backend and CMS for [tvkon.com](#), a video streaming website. The system can encode any kind of video files into Flash Video streaming files.
- 2001 - 2004 **PHP/MySQL/Linux** - Worked as a web programmer at the STG NGO. My job was to write PHP frameworks for portals and e-commerce websites.

## SKILLS

### • Programming

- **Java** - 4 years experience in writing distributed web mining applications in Java
- **PHP/MySQL/HTML** - 6+ years experience in web programming in PHP for high traffic websites
- Python, C#, C/C++, Assembly

### • Data mining / web mining

- **Parallel data processing** - 2 years experience with **MapReduce**-like systems ([Hadoop](#) and [Scope/COSMOS](#)). I worked on clusters ranging from 40 machines to 6000 machines, and worked with datasets ranging from a few gigabytes to tens of terabytes. I also installed and managed three Hadoop clusters between 8 and 40 machines.
- **Algorithms** - Implemented algorithms for text clustering, text classification, distributed web crawling, information extraction from web pages, geographical named entities disambiguation, frequent itemset mining / association rule mining, and several other data mining algorithms.
- **Indexing/storage** - 4 years experience with [Lucene](#) and [SOLR](#), including distributed index serving, multiple cores, near real-time indexing, faceting.

### • Networking / Administration

- **Linux** - 7+ years experience in Linux server administration. I worked with several flavors of Linux and configured software such as PHP, MySQL, Apache, PostgreSQL, Samba, lighttpd, Squid, BIND, tinyDNS, Varnish, BerkeleyDB, QMail, Sendmail, Plone, and many others
- **Cisco** - Completed the CCNA 1-4, CCNP 1-2 and Wireless academic path courses
- Skilled in many desktop software packages under Linux and Windows

### • Language skills - Romanian - native, English - fluent (14 years of training, several certifications), Hungarian - basic conversational (I'm part Hungarian), French - very basic

### • Project Management / People skills - 2 years experience in managing small to medium-sized teams, 1 years experience teaching PHP to a high-school class

### • Driving license - Category C in the US, category B in Romania