

# Arnaud Sahuguet, Ph.D.

New York, NY 10025  
+1 646-246-0942

[arnaud.sahuguet@gmail.com](mailto:arnaud.sahuguet@gmail.com)  
[@sahuguet](https://www.linkedin.com/in/sahuguet)  
[linkedin.com/in/sahuguet](https://www.linkedin.com/in/sahuguet)

## SUMMARY ↑

My areas of expertise include data – big, small, open, geographic – management and analysis, architecture of complex systems, product design and research & development.

My passion is to invent, architect and build products that leverage technology to solve meaningful problems and have a large social impact. I aim at empowering people and organizations to be more productive and collaborative through innovation.

I have also developed some strong working relationships with the academic research community, the non-profit sector and innovators at the city and government levels.

[Education](#) | [Experience](#) | [Add. information](#) | [Publications](#) | [External engagements](#) | [Patents](#)

## EDUCATION ↑

[University of Pennsylvania](#), Philadelphia, PA 1996–2001  
Ph.D. in computer science (advisor: [Val Tannen](#)).

[Ecole Normale Supérieure, Univ. Paris VI](#), Paris, France 1995–1996  
M.S. in algorithms and cryptography.

[Ecole Nationale des Ponts et Chaussées](#), Paris, France 1994–1996  
M.S. in civil engineering.

[Ecole Polytechnique](#), Palaiseau, France 1991–1994  
B.S. in engineering with major in computer science.

## EXPERIENCE ↑

[The Foundry @ Cornell Tech](#) 6/2014–present  
Director, New York,  

- helping faculty and students move from ideas to products and companies.

[The Governance Lab @ NYU](#) 6/2014–8/2015  
Chief Technology Officer and Principal, New York,  

- responsible for all IT and technology for a lab of 20 people.
- coordinating the technology aspect of all lab's projects and proposals.
- recruited and managing the engineering team (2 people).

[Google Inc.](#) 3/2006–3/2014  
Product manager, [Google.org](#), New York, 5/2011–4/2014  

- co-founded the Big Civic Data team (5 people) and the Collaborative Data Science project within Google Research; I conducted interviews with various cities to understand real needs; I identified the [iPython notebook](#) as a starting point for an integrated, interactive, visual and collaborative data science platform; I advocated the idea to internal teams (Chrome, Docs, Cloud) as well as [externally](#); the project has been open sourced at part of [project Jupyter](#).

- co-founded [OneToday](#), a mobile platform for micro-donations, launched in May 2013; highest rated Google app on [PlayStore](#); I was involved in all facets of the launch, including legal, marketing, partnership; I personally enrolled more than 500 non-profits to create enough project inventory for the launch.
- product lead for our effort on civic services, [open data](#) and e-government solutions; I co-authored a [schema.org vocabulary for government services](#) working with the Whitehouse Office of Science and Technology Policy (OSTP) and New York City.
- led our internal efforts on child protection as the Google liaison with the National Center for Missing and Exploited Children ([NCMEC](#)).

**Product manager, Geo team, New York, 5/2008–5/2011**

- co-architect for the next-generation unified infrastructure for Google Maps; I coordinated 20 engineers across 4 different teams.
- designed & implemented analytical and bug tracking tools for geo data support teams.
- 3 related [patents](#) filed.

**Product manager, Google Research and speech, New York, 08/2006–11/2008**

- defined strategy and product roadmap for speech efforts including dialogue systems, speech transcription and speech synthesis (foundations for Google Voice Search).
- launched Google Audio Indexing (GAudi) for the [2008 US political campaign](#).
- launched [GOOG-411](#), Google first voice product (US and Canada); secured funding from CEO and coordinated marketing campaign ([billboard](#), [taxi](#), etc.); negotiated deal with phone manufacturer for "GOOG-411 phone" (CES 2008).

**20% projects and others, New York, 03/2006–3/2014**

- internal champion for Google barcode efforts; co-founded the [ZXing](#) open source library for Android, used by BarcodeScanner, a featured T-Mobile application that appeared on TV ads; 2 related [patents](#).
- launched "[nearby street intersection](#)" feature for non-visual maps applications.
- Google *trusted interviewer*, with more than 400 interviews conducted.

**Bell Labs Research, Lucent Technologies**

2001–2006

**Member of technical staff, Murray Hill, NJ, 09/2001–03/2006**

- Lucent representative for 3GPP on Generic User Profile; key contributor for [TS 22.240](#), [TS 23.240](#) and [TS 29.240](#).
- researcher focusing on data federation, personalization, identity management and converged services; [publications and tutorials](#) presented at various database and web conferences.

**Tropea Inc.**

1999-2002

**Founder, CTO, Philadelphia, PA, 1999-2002**

- Founded [company](#) to commercialize [W4F](#) web extraction tools.
- Sold licenses to end-user customers and integrator companies.
- Released technology as open source project after joining Bell Labs.

**DigiCash bv**

1996

**Intern, Amsterdam, NL, 03/1996–07/1996**

- Implemented an electronic voting prototype for TeleDemocratie project, with Berry Shoenmakers, [David Chaum](#) and Marcel Bullinga.

**Parisbas Capital Markets** 1994  
**Intern**, London, UK, 09/1994–12/1994

- Rewrote and implemented of various pricing tools, with Eric Auld and [Bruno Dupire](#).

**French Armed Forces** 1991-1992

**Sous-lieutenant at 11e R.C.**, Carpiagne, France, 03/1992–08/1992

**Aspirant at EABC**, Saumur, France, 10/1991–02/1992

### ADDITIONAL INFORMATION <sup>↑</sup>

- US and French citizen.
- Spoken languages: French (native), English (fluent), Spanish (good).
- Programming languages: Python, Javascript, SQL, Java; functional languages.
- Sports: soccer, marathon (3h20 PR), ultra-marathon (Black Rock City, 2011), triathlon.

### SELECTED PUBLICATIONS <sup>↑</sup>

[h-index=17](#); [total citations=1,472](#); full list of publications available from [DBLP](#).

- [Identity Management for Converged Networks](#), a Sun/Lucent white paper, 2006.
- [Sub-document queries over XML with XSQirrel](#), with Bogdan Alexe, WWW 2005.
- [User Profile Management in Converged Networks \(Episode II\): "Share your Data, Keep your Secrets"](#), with Bogdan Alexe, Irimi Fundulaki, Pierre-Yves Lalilgand, Abdullatif Shikfa, Antoine Arnail, CIDR 2005.
- [An Electronic Patient Record "on Steroids": Distributed, Peer-to-Peer, Secure and Privacy-conscious](#), with Serge Abiteboul, Bogdan Alexe, Omar Benjelloun, Bogdan Cautis, Irimi Fundulaki, Tova Milo, VLDB 2004.
- [ubQL: A distributed query language to program distributed query systems](#), PhD dissertation, 2001.
- [Building intelligent Web applications using lightweight wrappers](#), with Fabien Azavant, in Data Knowledge Engineering, 2001; 289 citations.
- [Everything You Ever Wanted to Know About DTDs, But Were Afraid to Ask \(Extended Abstract\)](#), WebDB 2000.

### EXTERNAL ENGAGEMENTS <sup>↑</sup>

- [The future of High Performance Computing](#), Feb 2014, **invited speaker** for Department of Energy.
- [NYU Abu Dhabi International Hackathon for Social Good in the Arab world](#), Feb 2013, judge; Feb 2014, **co-organizer and judge**.
- [De Innovatione Rerum](#), July 2013, Clermont-Ferrand, France, **invited speaker** for Michelin.

- [Financial Empowerment Hackathon](#), March 2013, New York City, **judge**.
- [NYC BigApps](#), March 2013, **invited speaker**.
- [Reinvent NYC.gov](#), New York City government's first hackathon, July 2011, **winner** for user interface category.
- [Identity Management on Converged Networks: A Reality Check](#) panel at the WWW 2006 conference, **organizer**.
- academic research related duties: ACM SIGMOD 2009 (program committee), CIDR 2007 (chair and program committee).

## PATENTS <sup>↑</sup>

- "Reducing power consumption of mobile device through automated environment awareness" ([8504066](#)).
- "Proactive User-Based Content Correction And Enrichment For Geo Data" ([20130110847](#)).
- "Techniques for authenticating access to a private account at a public computing device using a user's mobile computing device" ([8392971](#)).
- "Managing information about entities using observations generated from user modified values" ([8676804](#)).
- "Systematic approach to query optimization" ([6567802](#)).
- "Machine-readable representation of geographic information" ([20090303036](#)).
- "Ratings using machine-readable representations" ([20090242620](#)).