

DTL facts and figures

Nikolaos Laoutaris

Telefonica & DTL

Reframing Data Transparency, London June 2016



Birth of



the Data Transparency Lab

Participants included:

Northeastern University, MIT Human Dynamics Lab, Microsoft, Telefonica Innovation, Max Planck Institute for Software Systems, Mozilla, and more.



Downloads:

[DTL2014 Workshop Summary Report](#)

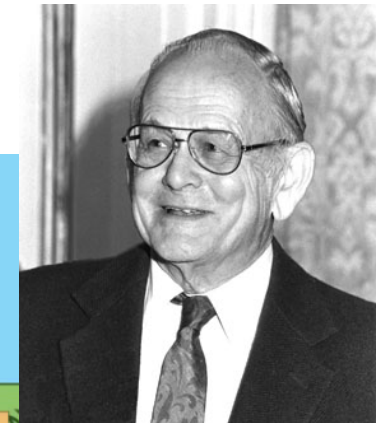
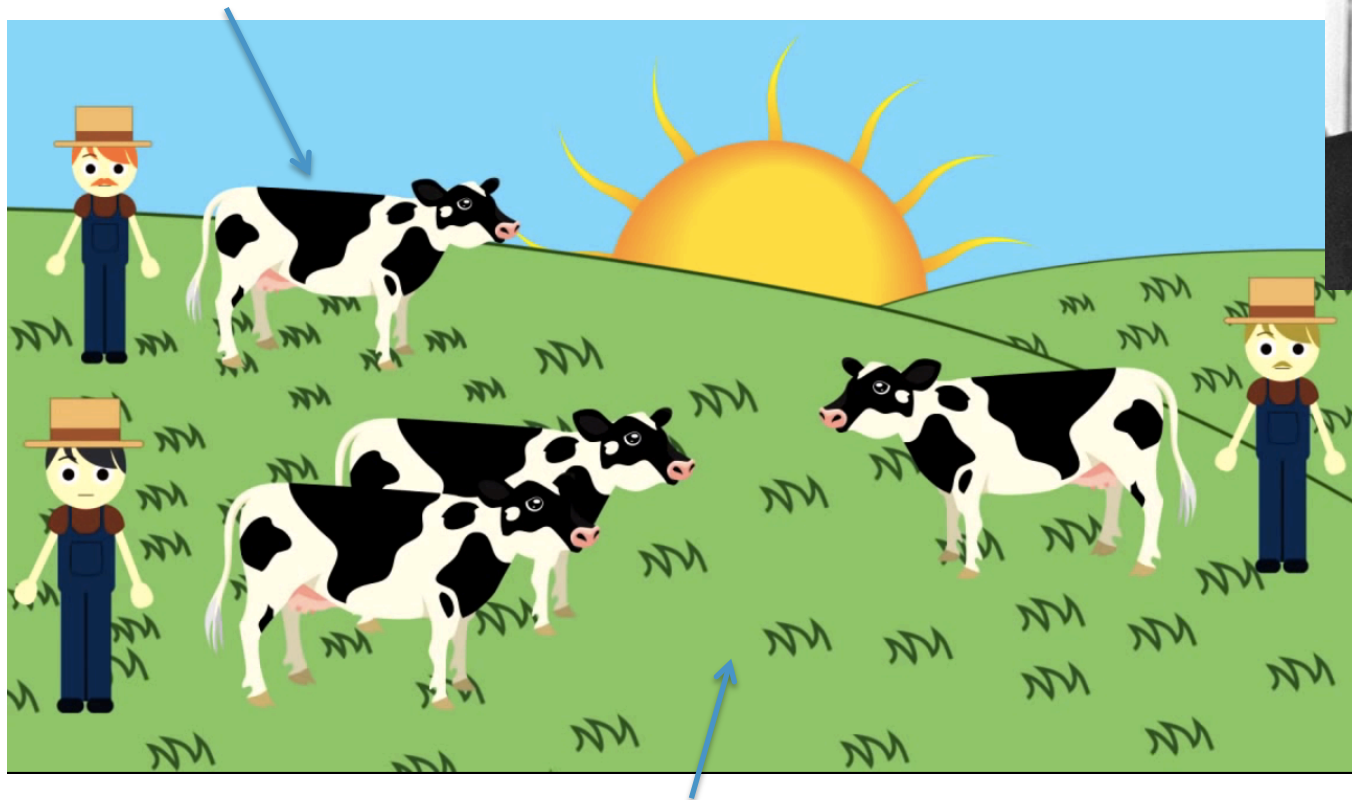
[DTL Vision](#)

[DTL Organization](#)

An impending “Tragedy of the commons on the web”

Garrett Hardin, 1968

Internet company in
Web Economy ... crossing **privacy red lines**



The “commons”: consumer trust on the web and it's business models



**“Sunlight is the
best disinfectant.”**

-LOUIS BRANDEIS



**“Publicity is justly commended as a
remedy for social and industrial diseases.
Sunlight is said to be the best of disinfectants;
electric light the most efficient policeman.”**

**—U.S. Supreme Court Justice Louis D. Brandeis,
“What Publicity Can Do,” *Harper’s Weekly*,
December 20, 1913**

after 1.5 years

... its time to move on

... we still have vision

... but also work to show

... here come some Facts and Figures

12

grants awarded in 2 rounds

DATA
TRANSPARENCY
LAB



Grants Program

DTL is a collaborative effort between universities, businesses and institutions to support research in tools, data, and methodologies for shedding light on the use of personal data by online services, and to empower users to be in control of their personal data online. In order to support research in these areas, DTL will award research grants to academic institutions worldwide.

Grantees

- ▶ Grantees 2016
- ▶ Grantees 2015

Call for Proposals 2016 (Now Closed)

The call for proposals ended on Apr 30th, 2016.

Grant awardees have been announced.

50K

of euros per grant

- Very low admin/reporting overhead
- But expecting to see running code ...
- 600K awarded in total so far

2015 Grantees

“WEB PRIVACY CENSUS”



“DIGITAL HALO” TOOL



Technical University of Denmark

REVEALING AND
CONTROLLING MOBILE
PRIVACY LEAKS



Northeastern University

BALANCE BETWEEN PRIVACY
AND FUNCTIONALITY



DATA-DRIVEN PRIVACY
ANALYSIS FOR END-USERS

Carnegie
Mellon
University

FACEBOOK DATA
VALUATION TOOL



Universidad
Carlos III de Madrid

DATA TRANSPARENCY LAB - EXAMPLE 1: FACEBOOK DATA VALUATION TOOL

Permits users to estimate how much money Facebook is making on them

INFORMS INDIVIDUAL USERS



A plugin for your browser that combines your online activity with Facebook's Public APIs to estimate your advertising value

DEMO VIDEO: <https://youtu.be/QPfc-gXGdjl>

REVEALS SOME MORE GENERAL TRENDS



Demonstrates how factors like country, status, studies, etc. impact on a user's advertising value

LIVE DEMO
https://acrumin.cartodb.com/viz/75d6d052-0648-11e6-8923-0e3ff518bd15/public_map

DATA TRANSPARENCY LAB - EXAMPLE 2: PRIVACY CENSUS

Privacy Census

A CENSUS THAT IDENTIFIES WEBSITES THAT TRACK USERS

Sites with canvas fingerprinting scripts

In a crawl conducted during January 2016, these websites were found to run scripts on their homepages that used the Canvas API to fingerprint users.

Show 25 entries Search:

Showing 1 to 25 of 15,089 entries

Alexa Rank	Site URL	Fingerprinting Domain
11	http://taobao.com	alicdn.com
29	http://tmall.com	alicdn.com
97	http://dropbox.com	dropboxstatic.com
115	http://bbc.com	doubleverify.com
143	http://cnzz.com	tbedn.cn
153	http://detail.tmall.com	alicdn.com
178	http://avito.ru	avito.st
219	http://washingtonpost.com	doubleverify.com

Privacy Auditor

👁️ You are being tracked. 👁️

[dropbox.com](#) has been spotted during the last month using the following fingerprinting techniques:

- ! Canvas Fingerprinting

Check my browser history vs. fingerprinting

Share your results:

[Twitter](#) [Facebook](#)

What is fingerprinting?

[About this tool](#)

TO SHED LIGHT ABOUT HOW TRACKING IS USED AND BY WHOM

MIT Technology Review

Computing

Largest Study of Online Tracking Proves Google Really Is Watching Us All

Google's Web trackers are present on the majority of the Web's top million sites.

by Tom Simonite May 18, 2016



Audio fingerprinting being used to track web users, study finds

Posted May 19, 2016 by [Natasha Lomas](#) (@riptari)

Data transparency must combine cutting-edge research with community involvement

Arvind Narayanan, Assistant Professor of computer science at Princeton.

26 May 2016

[Key to the success of data transparency...](#)

Steve Englehardt and I recently made available our draft paper [Online tracking: A 1-million-site measurement and analysis](#), funded in part by a DTL grant. It is part of the [Web Transparency & Accountability Project](#) at Princeton, and it's the most detailed look at online tracking conducted so far. Among our findings was the fact that the Audio, Battery, and WebRTC APIs in HTML5 are all being abused by third-party scripts for fingerprinting. There's been some press coverage [here](#) and [here](#).

1000000

15089

103

sites crawled by Privacy Census every month
sites doing canvas fingerprinting (P. Census)
of new Apps detected to leak PII (ReCon)

2016 Grantees

FINDING WALDO IN A
HAYSTACK OF INFORMAL
WRITING STYLES



CHARACTERIZING INDIRECT
PRIVACY LEAKS IN MOBILE APPS



PRIVACYMETER: REAL-TIME
PRIVACY QUANTIFICATION
FOR THE WEB



DETECTION AND
CIRCUMVENTION OF AD-BLOCK
DETECTORS



ANTMONITOR: ON DEVICE
NETWORK MONITORING



















TRANSPAD: BRING
TRANSPARENCY TO
TARGETED ADVERTISEMENT



30

2015 committee

Name ▼	Email	Affiliation
<input type="checkbox"/> John Byers 	byers@cs.bu.edu	Boston U
<input type="checkbox"/> Claude Castelluccia 	claudio.castelluccia@inria.fr	INRIA
<input type="checkbox"/> Augustin Chaintreau 	augustin@cs.columbia.edu	Columbia U
<input type="checkbox"/> Jon Crowcroft 	jon.crowcroft@cl.cam.ac.uk	Cambridge U
<input type="checkbox"/> Paul Francis 	francis@mpi-sws.org	MPI-SWS
<input type="checkbox"/> Krishna Gummadi (Chair) 	gummadi@mpi-sws.org	MPI-SWS
<input type="checkbox"/> Tristan Henderson 	tnhh@st-andrews.ac.uk	University of St Andrews
<input type="checkbox"/> Christian Kreibich 	christian@icir.org	ICSI
<input type="checkbox"/> Balachander Krishnamurthy 	bala@research.att.com	AT&T
<input type="checkbox"/> Nikolaos Laoutaris (Chair)	nikolaos.laoutaris@telefonica.com	Telefonica
<input type="checkbox"/> Alan Mislove 	amislove@ccs.neu.edu	Northeastern U
<input type="checkbox"/> Sue Moon 	sbmoon@kaist.edu	KAIST
<input type="checkbox"/> Arvind Narayanan 	arvindn@cs.princeton.edu	Princeton University
<input type="checkbox"/> Andrew Odlyzko 	odlyzko@umn.edu	U Minnesota
<input type="checkbox"/> Nuria Oliver 	nuria.oliver@telefonica.com	Telefonica
<input type="checkbox"/> Salvatore Ruggieri 	ruggieri@di.unipi.it	University of Pisa
<input type="checkbox"/> Matthew Smith 	smith@cs.uni-bonn.de	U Bonn



Krishna Gummadi
Max Planck Institute

2016 committee

Name ▼	Email	Affiliation
<input type="checkbox"/> Ernst Biersack	eure@e-biersack.eu	Eurecom
<input type="checkbox"/> Jeff Brueggeman	jeff.brueggeman@att.com	AT&T
<input type="checkbox"/> John W. Byers	byers@cs.bu.edu	Boston University
<input type="checkbox"/> Claude Castelluccia	claudc.castelluccia@inria.fr	INRIA
<input type="checkbox"/> Augustin Chaintreau	augustin@cs.columbia.edu	Columbia University
<input type="checkbox"/> David Choffnes	choffnes@ccs.neu.edu	Northeastern University
<input type="checkbox"/> Daniel Coloma (PC, sysadmin)	daniel.colomabaiges@telefonica.com	Telefónica
<input type="checkbox"/> Emiliano De Cristofaro	e.decrisofaro@ucl.ac.uk	e.decrisofaro@ucl.ac.uk
<input type="checkbox"/> Josep Domingo-Ferrer	josep.domingo@urv.cat	Universitat Rovira i Virgili
<input type="checkbox"/> Krishna Gummadi	gummadi@mpi-sws.org	Max-Planck Institute for Software Systems
<input type="checkbox"/> Tristan Henderson	tnhh@st-andrews.ac.uk	University St. Andrews
<input type="checkbox"/> Balachander Krishnamurthy (Chair)	bala@research.att.com	AT&T
<input type="checkbox"/> Nikolaos Laoutaris (Chair)	nikolaos.laoutaris@telefonica.com	Telefónica
<input type="checkbox"/> Marco Mellia	mellia@polito.it	Politecnico di Torino
<input type="checkbox"/> Ionel Naftanaila	naftanaila@iabeuropa.eu	IAB Europe
<input type="checkbox"/> Nick Nikiforakis	nick@cs.stonybrook.edu	Stony Brook University
<input type="checkbox"/> Chris Payne	C.Payne@wfanet.org	World Federation of Advertisers
<input type="checkbox"/> Chris Riley	mchris@mozilla.com	Mozilla
<input type="checkbox"/> Vincent Toubiana	vtoubiana@cnil.fr	CNIL - Commission Nationale de l'Informatique et des Libertés
<input type="checkbox"/> Sara Watson	samariewatson@gmail.com	Tow Center and Berkman Center



Balachander Krishnamurthy
AT&T Labs--Research

* *"All hat and no cattle",*

Texan quote

116

received submissions in 2 rounds

- we only sent emails, twitted, and gave talks
- submission open for 1-2 months

... and its not like we are giving away millions

350

473

#Reviews: 190 (2015), 160 (2016)

#Comments: 137 (2015), 336 (2016)

3

of DTL Conferences

DATA TRANSPARENCY LAB: DTL ANNUAL CONFERENCE

A unique opportunity to connect different type of profiles working on Transparency

THE BEGINNING: BARCELONA NOVEMBER 2014



PARTICIPATION

- 80+ attendants
- World-Class research, industry and policymakers participation
- Definition of the Data Transparency Areas of Research and proposal of the Framework for Collaboration in DTL (Memorandum of Understanding)

DATA TRANSPARENCY LAB: DTL ANNUAL CONFERENCE

A unique opportunity to connect different type of profiles working on Transparency

HELD AT THE MIT MEDIA LAB (NOV'15)



MEDIA IMPACT



PARTICIPATION

- 130+ attendants (plus 400+ remote audience)
- Multi-disciplinary reach:
 - Policy Making and Economic Development organisations: OECD, Open Data Institute, Centre for Information Policy Leadership...
 - Industry and design: Ghostery, AT&T; NewsUK, Design School at Politecnico di Milano, Office for Creative Research...
 - Research: 6 Grantees + 15 poster sessions.

KEY TAKEAWAYS

- Internet is evolving from being “user-centric” to being “data-centric”
- Opportunities for companies to demonstrate best practices around usage of data and be transparent with end-users
- Policy-makers and technologists must work together as technology is way more advanced than the current laws

DATA TRANSPARENCY LAB: DTL ANNUAL CONFERENCE

**2016 Columbia University, NYC, November
16th-19th 2016**



FOCUS

Practical steps towards transparency
in industry, society and research

Co-Located Workshops

- ACM 1st Data Transparency Workshop
- FAT-ML: Fairness Accountability and
Transparenyc in Machine Learning

Sessions

- **Research:** presentation of ongoing research by DTL grantees
- **Industry:** “Transparency from within”. How to organize a company for personal data transparency
- **Society:** Data Transparency for Social Good
- **Economy:** the Economics of Data Transparency
- **Users:** user-centric Design for Personal Data Transparency
- **Awareness:** Computational Journalism and Transparency

Hackathon

18-hour hakcathon to explore new possibilities in data transparency
With the support and collaboration of Mozilla

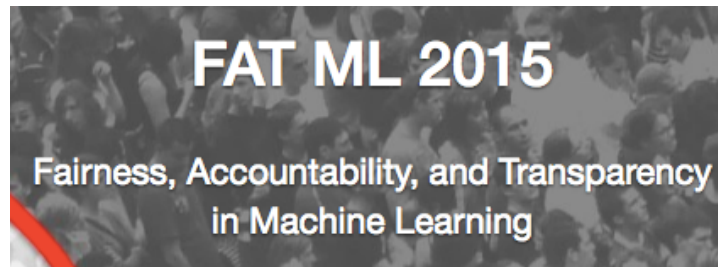
3.5

days we'll spend in NYC in Nov'16

data transparency lab



1.5 days of DTL Conf



1 Machine Learning & Privacy

1 Systems/Measurements & Privacy



I am done

Hope to see you in New York

Thank you!
Let's make the web economy sustainable!



To probe further

- \$heriff extension available at <http://sheriff.dynu.com/views/home>
- DTL at <http://www.datatransparencylab.org/>
- Publications
 - J. Mikians, L. Gyarmati, V. Erramilli, N. Laoutaris, “Crowd-assisted Search for Price Discrimination in E-Commerce: First results,” **ACM CoNEXT.13**. [\[full version\]](#)
 - J. Mikians, L. Gyarmati, V. Erramilli, N. Laoutaris, “Detecting price and search discrimination on the Internet,” in Proc. of **ACM HotNets'12**. [\[pdf\]](#)
- Blog
 - Cows, privacy, and tragedy of the commons on the web
<http://www.thedigitalpost.eu/2016/channel-data-economy/cows-privacy-tragedy-commons-web>