

State of Data

Arató Bence
BI Consulting

Technology



The cardinal rules of DBMS development

Rule 1: Developing a good DBMS requires 5-7 years and tens of millions of dollars.

That's if things go extremely well.

Rule 2: You aren't an exception to Rule 1.

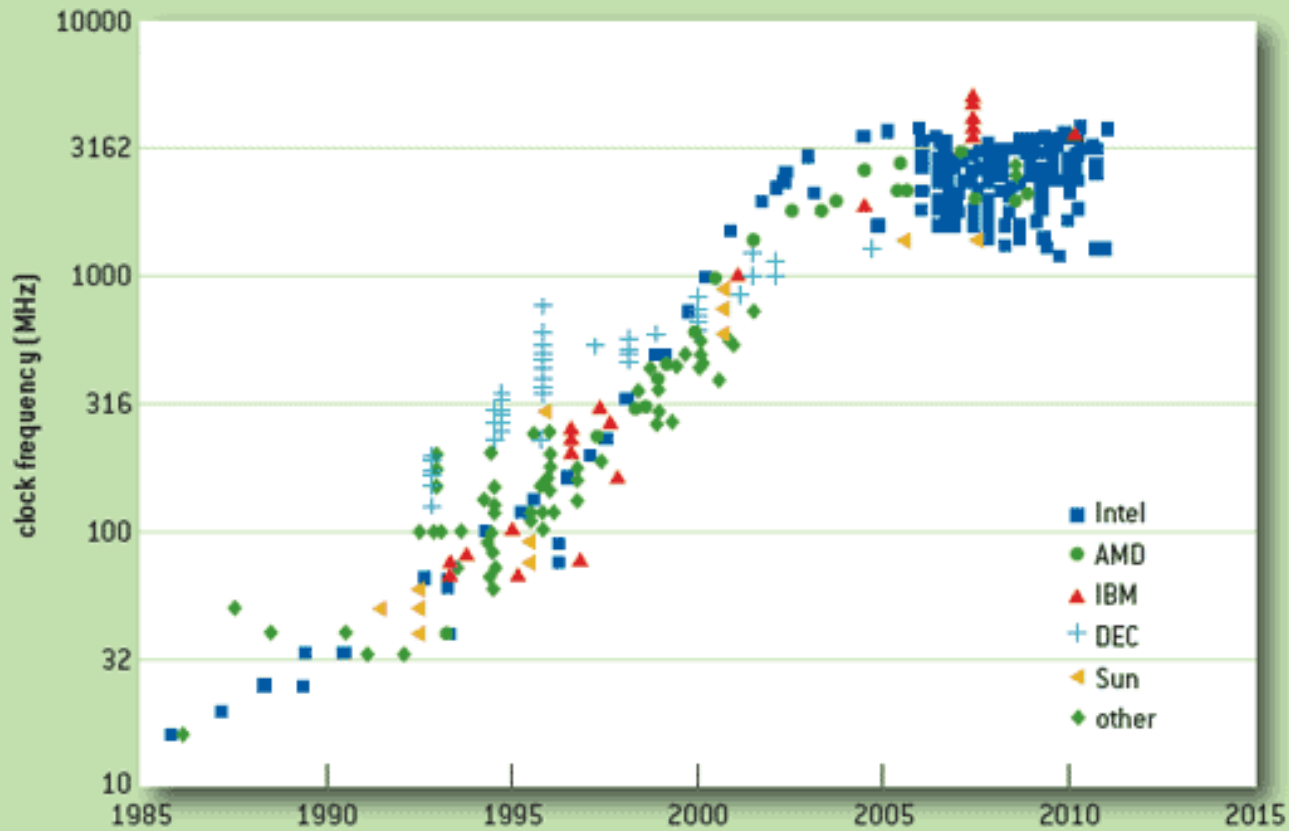
Let's talk about
size

Hadoop HW evolution

	CPU	RAM	Disk
2009	2x 4 cores	6 GB	3 TB
2010	2 x 4 cores	16-24 GB	4 x 1 TB
2011	2 x 4-6 cores	24-36	6-8 x 1-2TB
2013	2 x 6 cores	64 GB	12 x 1-2 TB
2014	2 x 8 cores	128 GB	12 x 2 -3TB

FIGURE 7

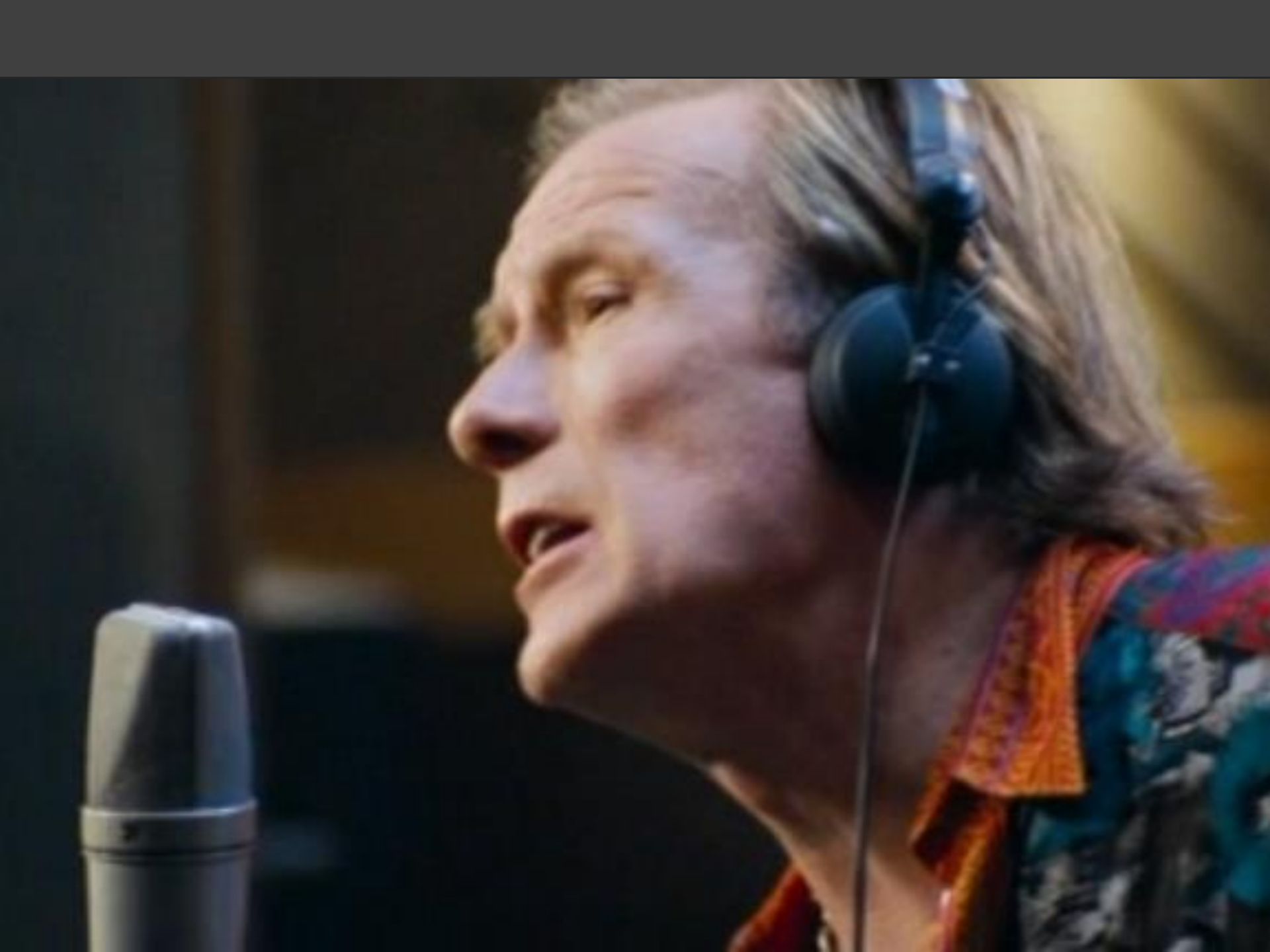
Processor Frequency Scaling Over Time



Cost of Data

	Price / TB	Examples
2008	\$100K	Vertica
2009	\$50-60K	Oracle, Teradata, Netezza
2010	\$20K	Netezza
2013	\$1-2K	General Hadoop
2013	\$1-3K	Amazon Redshift





Data is all
around us



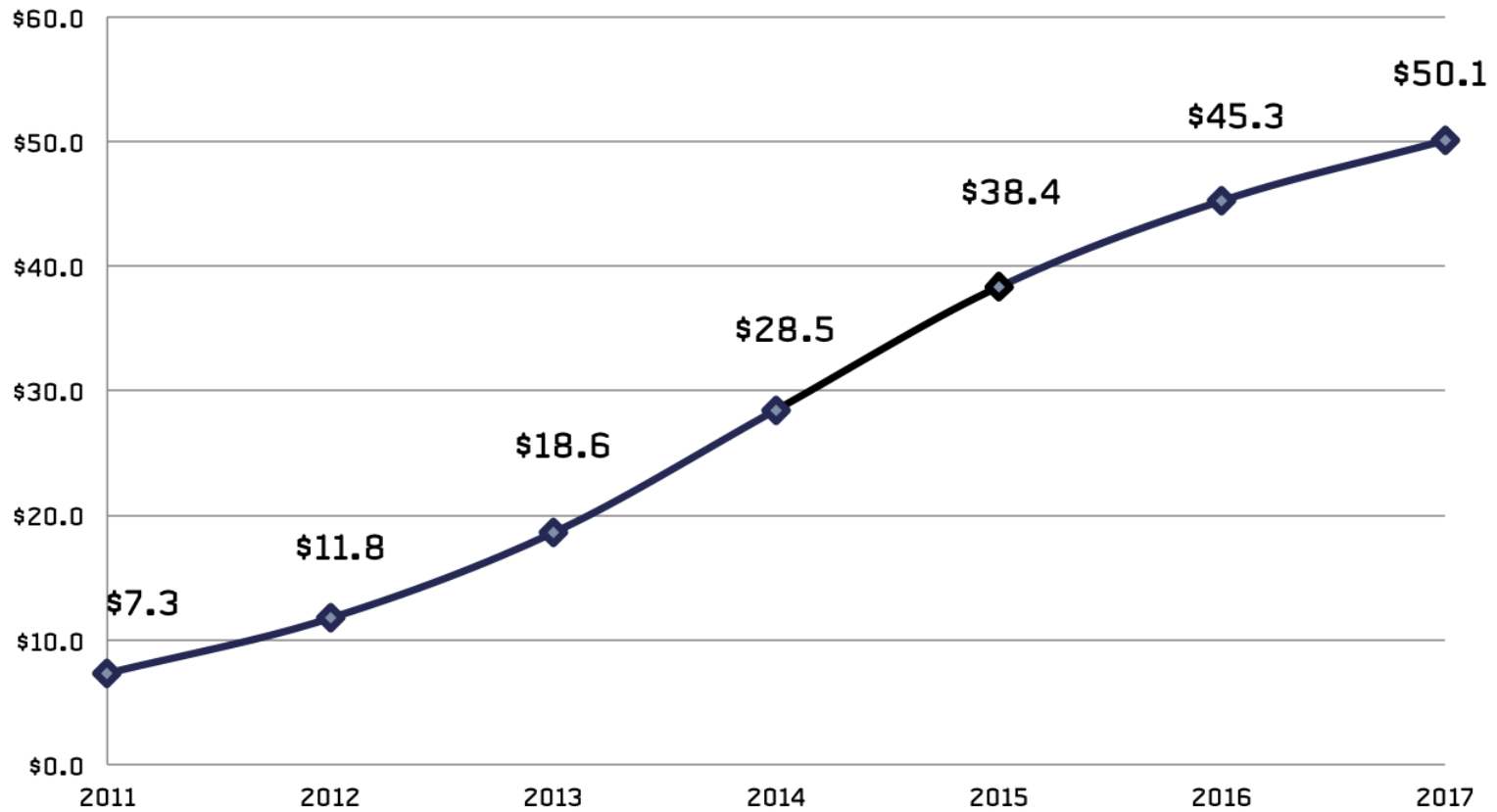
Big Market

The Big Data technology and services market will grow at a 27% compound annual growth rate to \$32.4 billion through 2017 - or about six times the growth rate of the overall ICT market

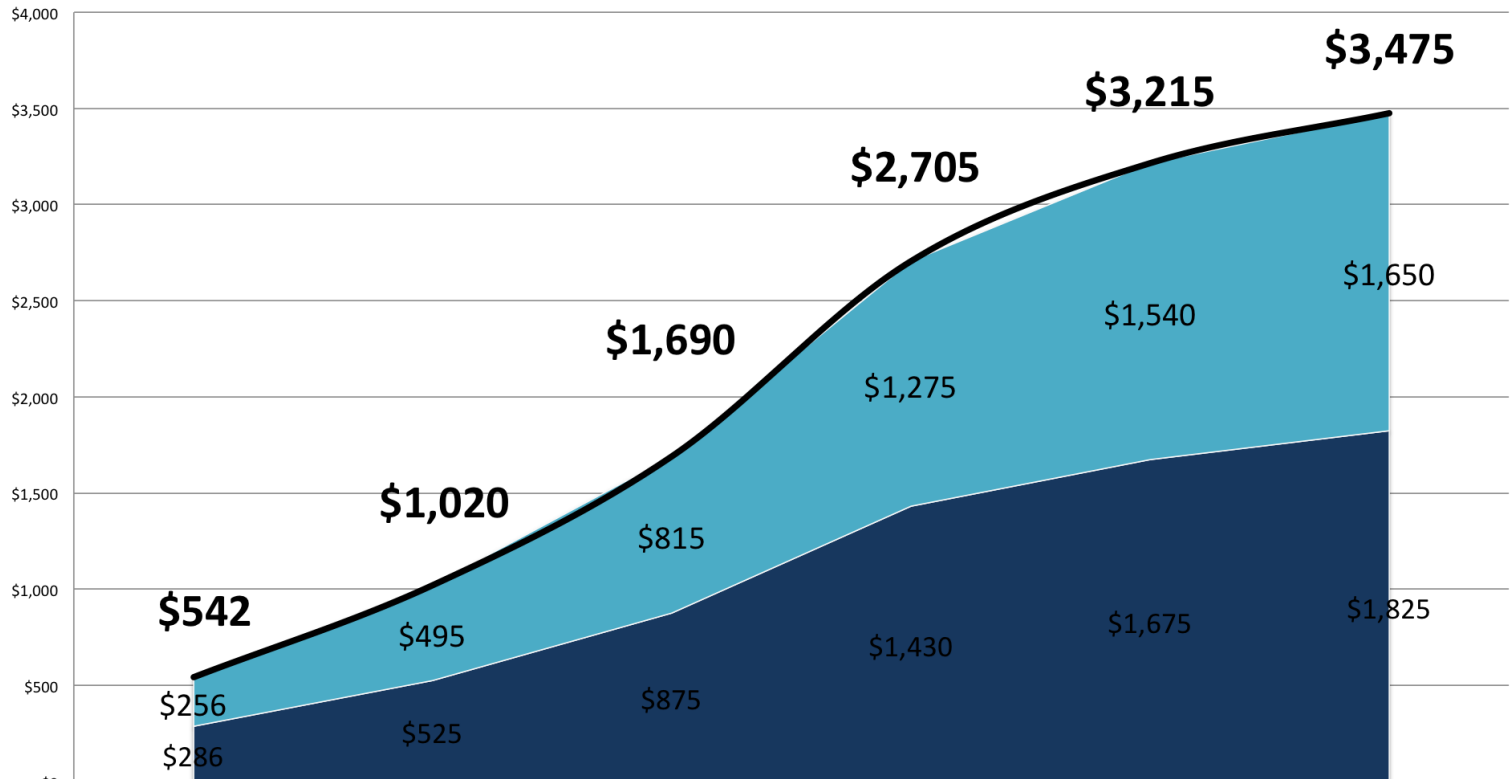
IDC Worldwide Big Data Technology and Services
2013-2017 Forecast, Dec 2013



Big Data Market Forecast, 2011-2017 (in \$US billions)



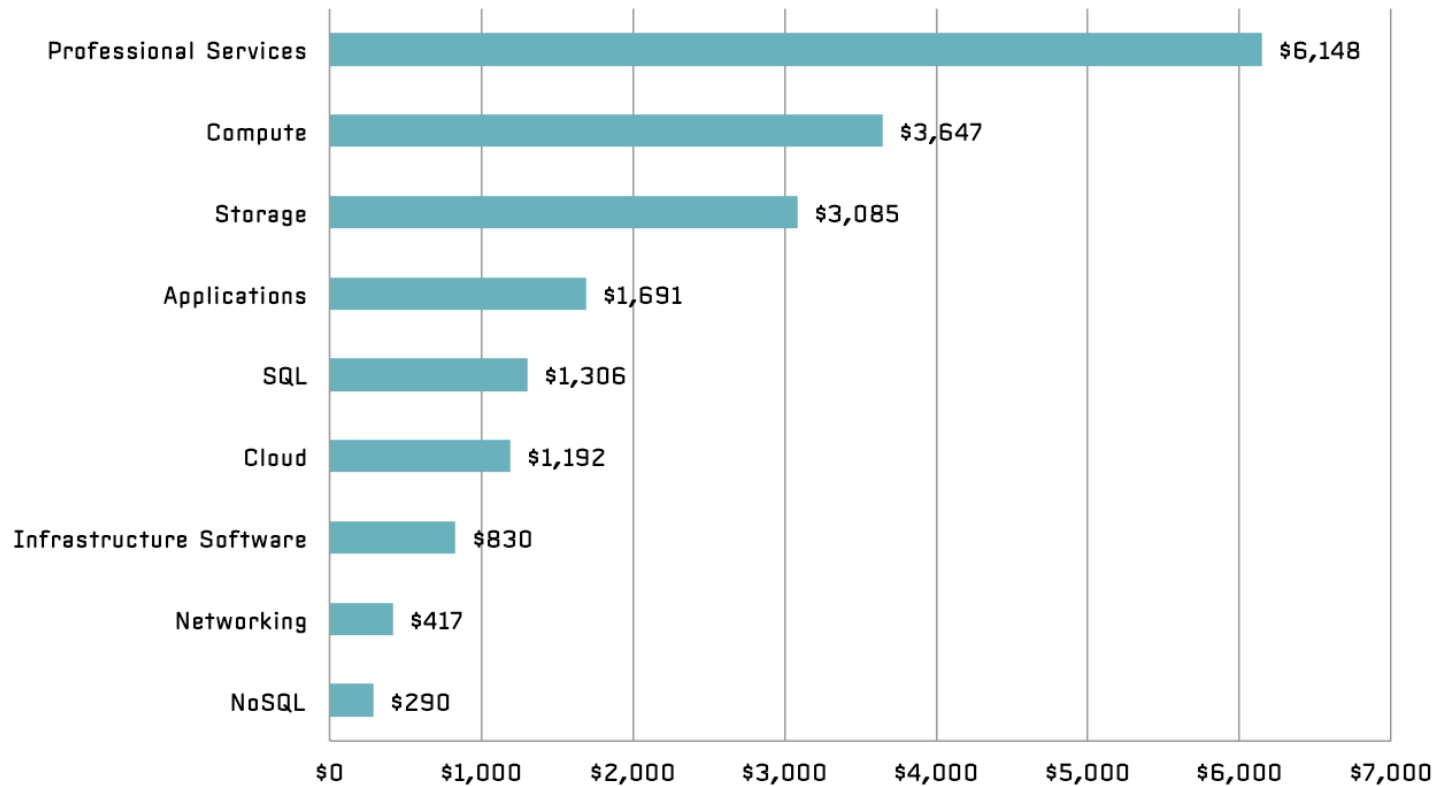
Hadoop & NoSQL Software/Services Revenue Projection, 2012-2017 (\$M)



	2012	2013	2014	2015	2016	2017
■ Hadoop	\$256	\$495	\$815	\$1,275	\$1,540	\$1,650
■ NoSQL	\$286	\$525	\$875	\$1,430	\$1,675	\$1,825
— Total	\$542	\$1,020	\$1,690	\$2,705	\$3,215	\$3,475



Big Data Revenue by Sub-Type, 2013 (in \$US millions) (n=\$18,814)



Big Data People

LinkedIn  Account Type: Basic

[Home](#) [Profile](#) [Contacts](#) [Groups](#) [Jobs](#) [Inbox](#) **26** [Companies](#) [News](#) [More](#)

Keywords:



First Name:

Last Name:

Title:

11 results



Tibor Kiss 1st

Senior Software Engineer at Omixon

Hungary · Information Technology and Services

▶ 2 shared connections · Similar ·  66



Zoltán Prekopcsák 1st

Data Scientist

Hungary · Research

▶ 10 shared connections · Similar ·  185

Hungary 

33 results

Sort by: Relevance  V**Karoly Kasa** 1st

CTO and co-owner at WebLib Kft.

Hungary · Information Technology and Services

▶ 13 shared connections · Similar ·  79Message **Tibor Kiss** 1st

Senior Software Engineer at Omixon

Hungary · Information Technology and Services

▶ 3 shared connections · Similar ·  84Message **Endre Jofoldi**   1st

Search Specialist, CEO of WebLib

Hungary · Computer Software


▶ 24 shared connections · Similar ·  431Message **Péter Hellinger** 2nd

Software Engineer

Hungary · Information Technology and Services

▶ 7 shared connections · Similar

Connect 

 PREMIUM Advanced

SEARCH
Advanced >

All
People
More...

Keywords

First Name


Last Name




Title

Company


School



103 results for **hadoop**




Janos Matyas   **1st**
CTO, Founder at SequenceIQ
Hungary · Information Technology and Services
[3 shared connections](#) · [Similar](#) ·  **238**



Current: CTO, Founder at SequenceIQ
Past: Solutions Architect at EPAM Systems
Weblogic Consultant at BEA Systems
Java EE Developer at KFKI Ltd.



Tibor Kiss  **1st**
Senior Software Engineer
Hungary · Information Technology and Services
[3 shared connections](#) · [Similar](#) ·  **166**

Past: Senior Software Engineer at Omixon Biocomputing
- massive parallel processing of full genome NGS data with
Hadoop - bioinformatic pipelines for NGS...



Tamas Foldi  **1st**
Director of IT Development
Hungary · Information Technology and Services
[51 shared connections](#) · [Similar](#) ·  **286**

Current: Director of IT Development at Starschema
..., **hadoop**) - Design state of art analytic solutions... , and some
other special ERP/CRMs - **Hadoop**, Storm, Mahout...

SEARCH

142 results for **data scientist**

[Save search](#) •

Advanced >

All

People

More...

Keywords

First Name

Last Name

Title

Company

School

Location

Located in or near:

Country

Hungary

Search
Reset

	<p>Istvan Jarasi 1st</p> <p>Data Scientist at HiFly Labs</p> <p>Hungary • Information Technology and Services</p> <p>▶ 35 shared connections • Similar • 189</p>	Message ▼
	<p>Levente Török 2nd</p> <p>Data Scientist at blinkbox music</p> <p>Hungary • Internet</p> <p>▶ 12 shared connections • Similar</p>	Connect ▼
	<p>Gábor Fodor 1st</p> <p>Data Scientist at realeyesit.com</p> <p>Hungary • Telecommunications</p> <p>▶ 18 shared connections • Similar • 113</p>	Message ▼
	<p>Eszter Csernai 2nd</p> <p>Data Scientist at BalaBit IT Security</p> <p>Hungary • Information Technology and Services</p> <p>▶ 3 shared connections • Similar</p>	Connect ▼
	<p>Gábor Simon 2nd</p> <p>Data Scientist at Dmlab</p> <p>Hungary • Computer Software</p> <p>▶ 6 shared connections • Similar</p>	Connect ▼
	<p>Péter Pach 2nd</p> <p>Lead Data Scientist at Bay Zoltán Nonprofit Ltd. for Applied Research</p> <p>Hungary • Research</p> <p>▶ 7 shared connections • Similar</p>	Connect ▼
	<p>Gabor Hamori 2nd</p> <p>Research Director, Data Scientist, STAT4U</p> <p>Hungary • Information Technology and Services</p> <p>▶ 6 shared connections • Similar</p>	Connect ▼

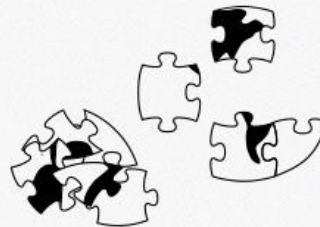
Doing Good with Data



Let's use data to change the world. Here's how.

**The Puzzle:**

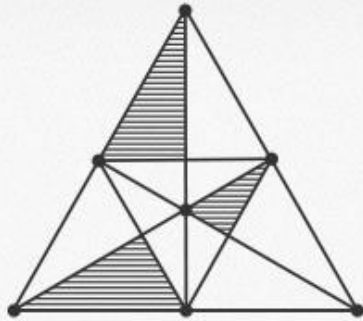
NGOs and other mission-driven organizations are faced with an onslaught of decisions on how to best collect, analyze and visualize the data that will help them make the world a better place.

**The Solution:**

We create teams of pro bono data scientists who work together with these organizations to understand the need, find a solution, and collaborate to tackle the problem with data science.

**The Big Picture:**

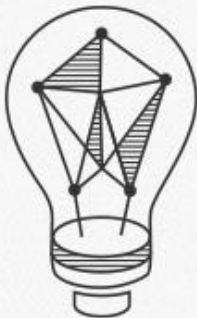
Organizations get a solution they can use along with an education in data science, our volunteers get real world experience and the warm fuzzies, and we all get to live in a better world.



DataCorps

The DataCorps™ is an elite group of data scientists, technologists, project managers, and designers who have been vetted by DataKind and engage with social organizations on a pro bono basis. These projects last between one to six months and are structured so that members can work in their spare time.

[Read more...](#)



DataDives

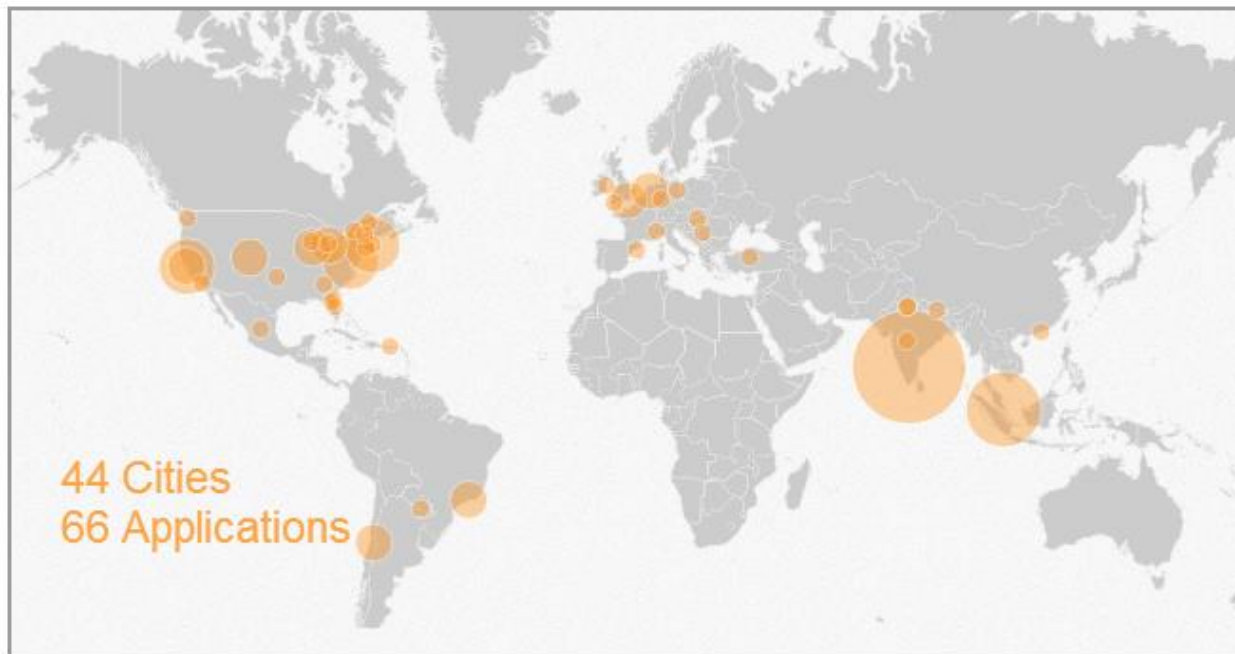
A DataDive™ is a weekend event that teams selected social organizations that have well-defined data problems with volunteer data scientists to tackle their data challenges. These events are completely free and voluntary and serve to energize the base, provide direct services to organizations, and to enlighten social sector groups to the power of using data in their programs.

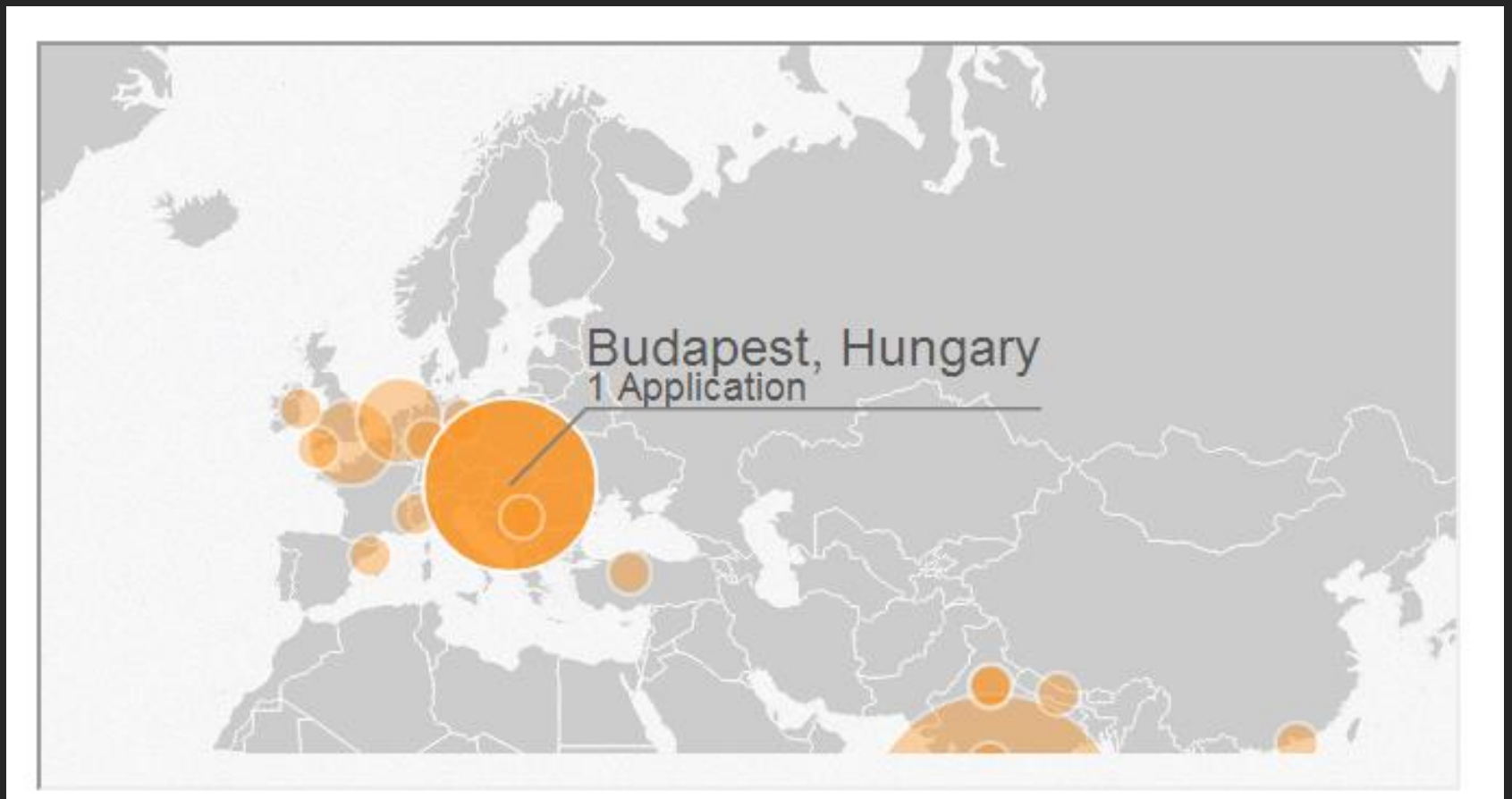
[Read more...](#)

DataKind, Coming to a City Near You?



If you haven't heard the news, we're inviting data scientist leaders to apply for the opportunity to [launch a DataKind chapter](#) in their city.





And finally



Pro Bono Analytics

[Home](#) [Members](#) [Photos](#) [Discussions](#) [More](#)

[Join us!](#)

Budapest, Hungary

Founded Jul 16, 2013

[Members](#) 29

[Our calendar](#) 

Organizer:
Arató Bence



[Contact](#)

We're about:

We are BI and analytics professionals, ready to help non-profit organizations to deal with data-related challenges. Inspired by [DataKind](#).

Join us if you would like to volunteer your free time and help us to help.

[Join us](#)

Join us and be the first to know when new Meetups are scheduled

[Who do I know here?](#)

Log in with Facebook to find out


By creating a Meetup account, you agree to the [Terms of Service](#)



Doing Data for Good

Upcoming

What's new

 NEW MEMBER

[Kovács László](#)
joined

April 30

