DATA SCIENCE FOR SOCIAL GOOD
The best minds of my generation are thinking about how to make people click ads. That sucks.

— Jeff Hammerbacher (Co-founder Facebook)
Social organizations need to allocate scarce resources for maximum impact.
· predicting where services will be needed
· prioritizing resources based on expected impact
  · forecasting trends and changes
· identifying supporters likely to donate
Basically regression problems
Social organizations often don’t have **DATA SCIENTISTS**
What if data scientists would donate their skills?
This is already happening...* 

* ...but not so much in Germany
A guided tour through the Data Science for Social Good movement
DSSG Fellowship

- supported by the Eric & Wendy Schmidt Foundation
- 6 month fellowship
- currently in 3rd year (started 2013)
- Partners: NGOs, Governments

http://dssg.io/
Examples (Project Headline Copy & Paste)

- World Bank Group – Prediction & Identification of Collusion in International Development Projects
- Chicago Public Schools – Student Enrollment Prediction for Budget Allocation
- Pecan Street, WikiEnergy – Building Open Source Tools to Analyze Smart Meter Data
More examples

visit http://dssg.io/projects/
DSSG Atlanta

- like DSSG Chicago
- mainly funded by Oracle and Georgia Tech
- started 2014 (one year after Chicago)

http://dssg-atl.io/
University of Washington just announced their DSSG Summer Program

http://escience.washington.edu/what-we-do/data-science-for-social-good
Bayes Impact (YC-backed)

Bayes Impact is a nonprofit that deploys data scientists to solve big social problems with civic and nonprofit organizations

- founded 2014
- started as (full-time-)fellowship
- now hiring long-term employees
Examples (Project Headline Copy & Paste)

- Increasing Graduation Rate And Optimizing Class Offerings For UC Riverside
- Improving Outcomes For Emotionally And Behaviorally Challenged Children With Youth Villages
- Stratification Of Parkinson's Disease Patients
- Optimizing Ambulance Response Times In Sf
Bayes Hack Nov 15th

One weekend, impact the world

http://bayeshack.challengepost.com/submissions
DRIVENDATA

- started this year (2014)
- currently 4 open competitions

http://www.drivendata.org/
MORE OPPORTUNITIES TO GET INVOLVED

We're tackling the world's biggest problems through data science.
— http://www.datakind.org

DataKind connects charities with data scientists by organizing two-day data dives where those data scientists help solve the charities’ data problems.
EXAMPLES
Analyzing poverty through satellite imagery

DataKind helped GiveDirectly – an NGO making unconditional cash transfers to poor households via mobile phones in Kenya and Uganda\(^1\) – to identify especially needy villages through satellite image analysis\(^2\).

\(^1\) [http://www.ted.com/talks/joysunshouldyoudonate_differently](http://www.ted.com/talks/joysunshouldyoudonate_differently)

View the presentation
or read the paper
Early warning system for human rights abuses

To help prioritize the many calls for help reaching Amnesty International’s Urgent Action Network DataKind volunteers have created a predictive model that analyzes messages for potential escalation.\(^3\)\(^4\)


Predicting demand for children’s hospices

Combining data from Shooting Star Chase, public data about the hospice and healthcare sector and demographic data DataKind volunteers calculated predicted demand against hospice capacity to reveal areas of possible shortage.  

5 http://www.datakind.org/finding-30000-missing-children/
Most of DataKinds projects have been tackled by volunteers on 2-day data dives.
What’s a Data Dive?
Before: Preparation of the data

(by voluntary data ambassadors in collaboration with the challenge partner – starting ~2 month before the data dive)

- anonymization/pseudonymization
- cleaning/fixing
- ensuring proper (machine readable) data formats
BEFORE:
Preparation of the Questions
Any data scientist worth their salary will tell you that you should start with a question, not the data.

— Jake Porway in https://hbr.org/2013/03/you-cant-just-hack-your-way-to/
The Data Dive

- Challenge partners pitch their problems
- Volunteers create analyses, models and visualizations (led by data ambassadors) in two intense days of hacking
- solutions are being presented at the end
Aftermath: Implementation

Social organizations still don’t have the expertise: data ambassadors must help implement the solutions
Aftermath: Impact Analysis

Did it help at all?
There is currently no organization in Germany comparable to DataKind.
There is currently no organization in Germany comparable to DataKind.
We want to start the Data Science for Social Good Movement in Germany
WHO WE ARE?

- Daniel Kirsch
- Marit Brademann
- Jana Kludas
- Richard Lawrence
- Georg Walther
Supporters

- Klaas Bollhöfer, Chief Data Scientist @ The Unbelievable Machine Company
- Adam Drake, Chief Data Officer @ Skyscanner
- Dr. Alexander Weiß, Head of Data Analytics @ Trademob
What we need
We need data ambassadors (how DataKind calls them)

- to prepare data before data dives
- lead teams at data dives
- help with the implementation afterwards
We need problems

The international of the Data Science for Social Good-movement shows that data scientists are eager to donate their skills.

Social organizations need to understand how we can help them. Are you in contact with NGOs? Spread the word!
Contact

http://dssg-berlin.org/
@dssgber

Daniel Kirsch
mail@danielkirs.ch
Thank you!

Questions?
Links:

- http://datakind.org
- http://dssg.io
- http://dssg-atl.io
- http://bayesimpact.org
- http://codefor.de
- http://datalook.io
Foto of Jeff Hammberbacher by Fred Brenenson licensed under CC BY 2.0