

DATA SCIENCE AS A FORCE FOR GOOD

DSPT DAY 2019 | PORTO | 26.10.2019

1

WHY DATA SCIENCE AS A FORCE FOR SOCIAL GOOD?



90% of all data has been created in the **last two years**.

Source: IBM



Today it would take a person **approximately 181 million years** to download all the data from the internet.

Source: Physics.org



90% of all data has been created in the **last two years**.

Source: IBM



Today it would take a person **approximately 181 million years** to download all the data from the internet.

Source: Physics.org



By 2020, every person will generate **1.7 megabytes in just a second**.

Source: DOMO

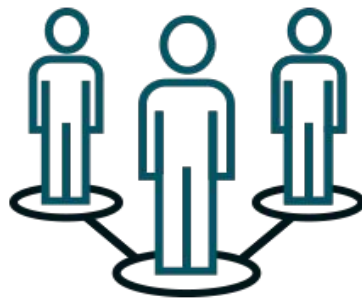
**Private-sector
organisations**

have been **building and
deploying data science
capabilities** for many years



**Public sector, civic and
non-profit organisations**

most public sector, civic and
non-profit organisations are
far behind



Everyone **produces data** *but* not everyone **benefits from it**

But when these public sector, civic and non-profit organisations are empowered with



- ✓ talent
- ✓ tools
- ✓ knowledge

Data Science can really work as a force for good

2

WHAT DO WE MEAN WITH USING DATA SCIENCE AS A FORCE FOR GOOD?

Data science projects that tackle **real world problems**

1 The **problem** we are solving is important and generates **social impact**

2 **Data can play** a role in solving the problem

3 The **organization is ready** to tackle this problem and **take actions** based on the work

Education

Health

Energy

Children
Social Care

International
development

Criminal
justice

Employment

Financial
inclusion

Public Safety

3

INSPIRING EXAMPLES



INCREASING THE EFFICIENCY OF HEART FUNCTION ASSESSMENT AND DIAGNOSIS THROUGH ECHOCARDIOGRAPHY



INCREASING THE EFFICIENCY OF HEART FUNCTION ASSESSMENT AND DIAGNOSIS THROUGH ECHOCARDIOGRAPHY

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



Cardiology Department of the University of Salamanca in Spain



INCREASING THE EFFICIENCY OF HEART FUNCTION ASSESSMENT AND DIAGNOSIS THROUGH ECHOCARDIOGRAPHY

ORGANISATION

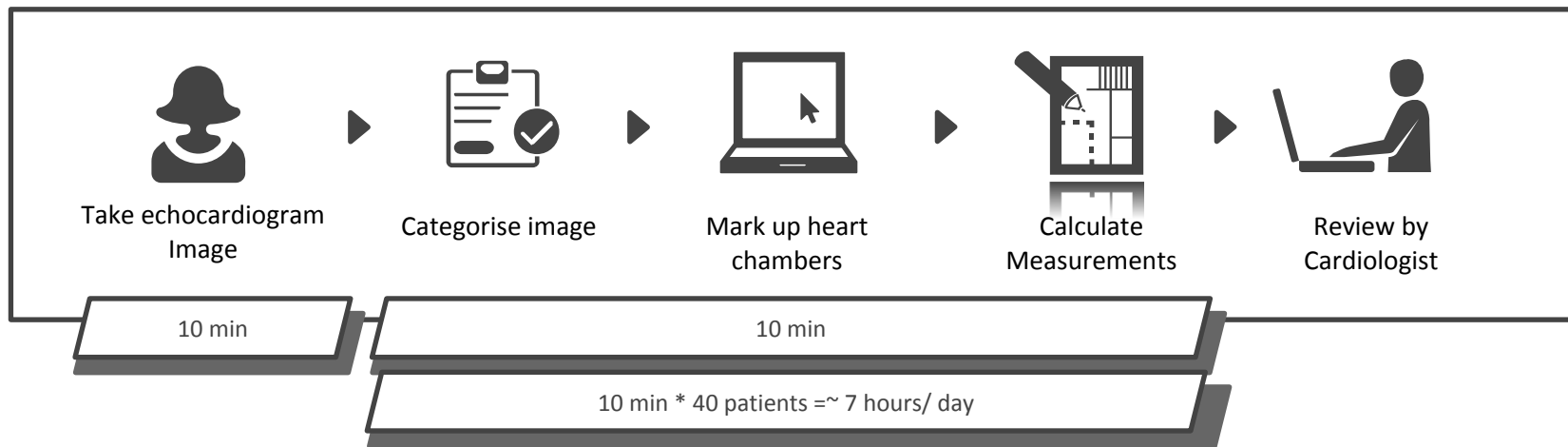
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



Doctors spend a lot of time measuring heart chambers, often for healthy patients



INCREASING THE EFFICIENCY OF HEART FUNCTION ASSESSMENT AND DIAGNOSIS THROUGH ECHOCARDIOGRAPHY

ORGANISATION

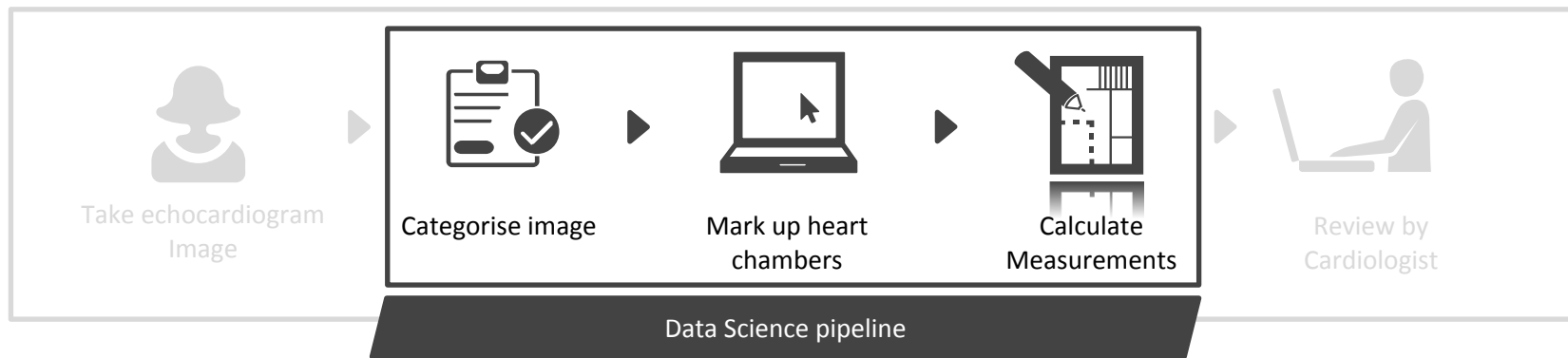
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



A data science pipeline that automates the estimation of measurements to suggest patient status so cardiologists can identify and spend more time with sick patients.



INCREASING THE EFFICIENCY OF HEART FUNCTION ASSESSMENT AND DIAGNOSIS THROUGH ECHOCARDIOGRAPHY

ORGANISATION

PROBLEM

SOLUTION

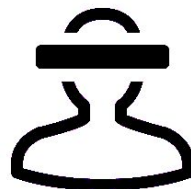
DATA

APPROACH

IMPACT



25,000 echocardiograms
(10 TB of videos)



Anonymized
database of measurements



INCREASING THE EFFICIENCY OF HEART FUNCTION ASSESSMENT AND DIAGNOSIS THROUGH ECHOCARDIOGRAPHY

ORGANISATION

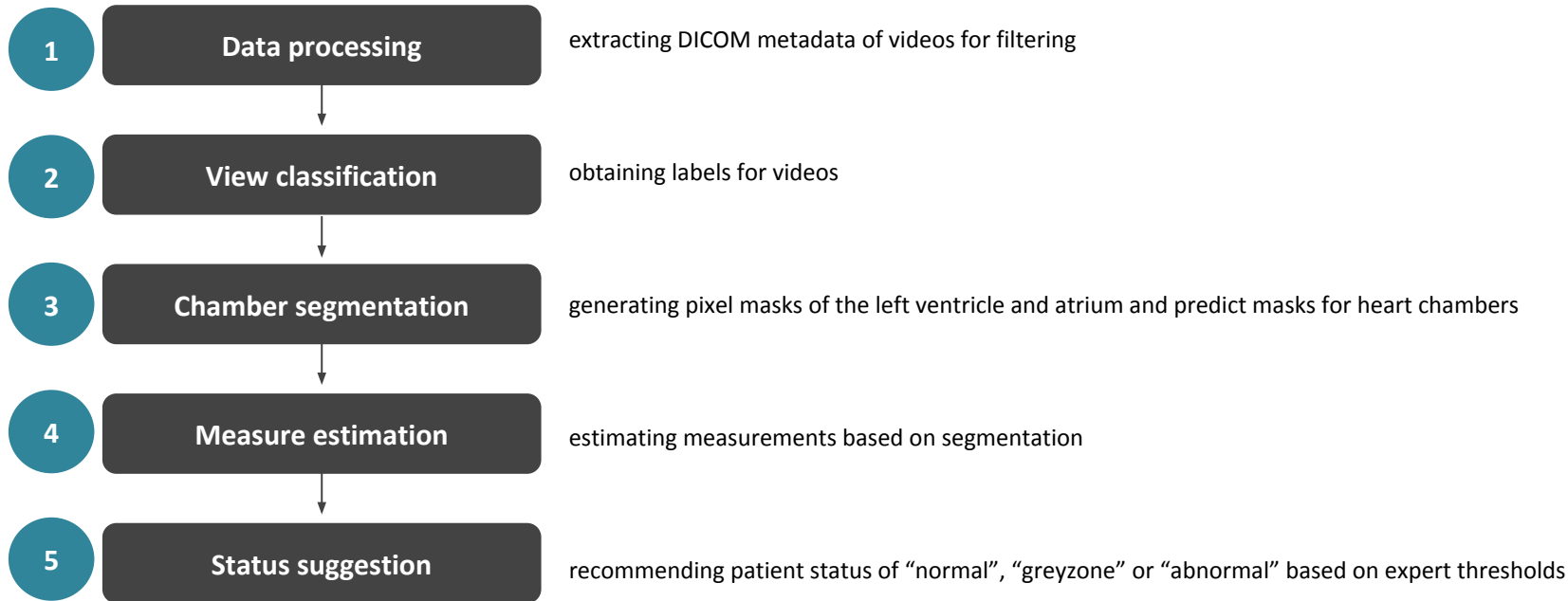
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT





INCREASING THE EFFICIENCY OF HEART FUNCTION ASSESSMENT AND DIAGNOSIS THROUGH ECHOCARDIOGRAPHY

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT

Free up 20% of
Cardiologist's time

4-6 more patients a day
at Univ.Hospital of
Salamanca alone

More time for specialists
to spend with the
patients that need them
most



MATCHING JOBSEEKERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL



MATCHING JOBSSEEKERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



INSTITUTO DO EMPREGO
E FORMAÇÃO PROFISSIONAL

IEFP is the government institute responsible for reducing unemployment in Portugal



MATCHING JOBS E E KERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



Job Center



Action Plan



Interventions



Job

Which interventions should job counsellors recommend to job seekers to help them find meaningful employment?



MATCHING JOBSSEKERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL

ORGANISATION

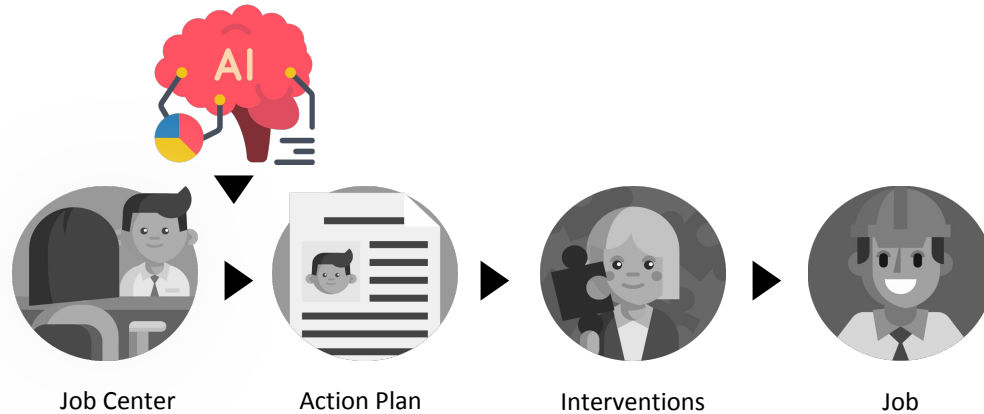
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



A system which suggests relevant and effective interventions for a job seeker



MATCHING JOBBEEKERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT

Intervention Recommendations

UTE ID

235245

	Code	Intervention Description	Predicted time to employment
Option 1	134	MEDIA TRAINING COURSE	6 months ?
	7	CV WORKSHOP (CREATIVE)	
	57	JOB SEARCH TECHNIQUES	
Option 2	134	MEDIA TRAINING COURSE	11 months ?
	13	INTERVIEW WORKSHOP (CREATIVE)	
	156	ADVERTISING COURSE	

Change priority

More options



MATCHING JOBSSEKERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

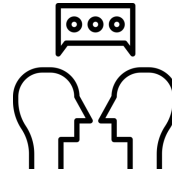
IMPACT



9 years of IEFP data
(2010-2019)



3.1m job seekers



101m transactions



255 training courses



MATCHING JOBBEEKERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL

ORGANISATION

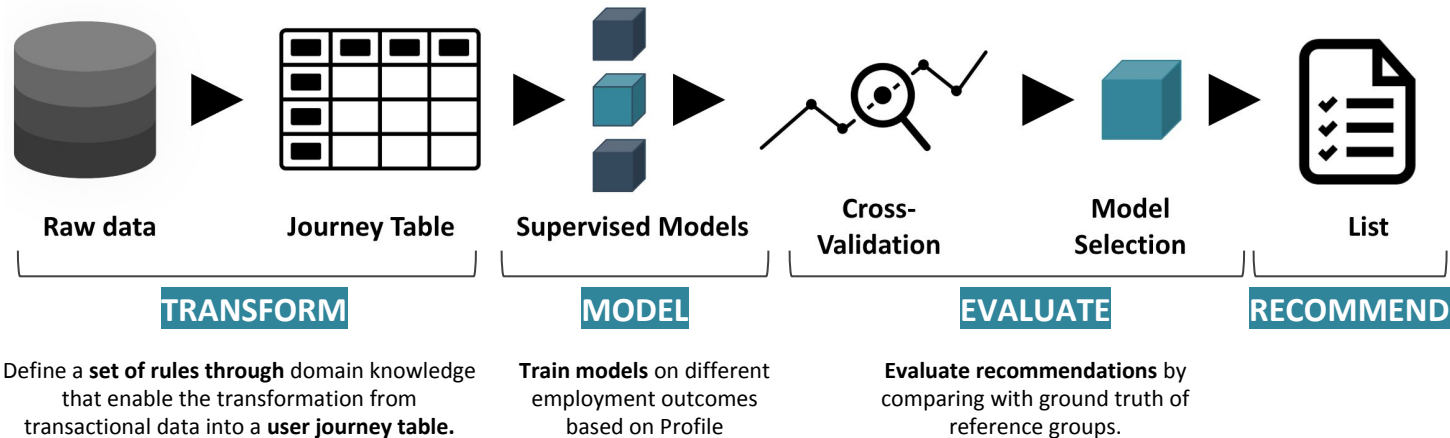
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT





MATCHING JOBSSEKERS WITH INTERVENTIONS TO IMPROVE EMPLOYMENT OUTCOMES IN PORTUGAL

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT

Help **300k job seekers**
find employment
faster

Assist 1500 job
counsellors to **work more**
efficiently and effectively

Expand IEPF's
data driven initiatives



**MAPPING JOURNEYS OF VULNERABLE CHILDREN ACROSS SERVICES TO IDENTIFY
MOST AT RISK CASES EARLY ON AND BETTER TARGET PREVENTATIVE SERVICES**



MAPPING JOURNEYS OF VULNERABLE CHILDREN ACROSS SERVICES TO IDENTIFY MOST AT RISK CASES EARLY ON AND BETTER TARGET PREVENTATIVE SERVICES

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



Newcastle City Council in the United Kingdom



MAPPING JOURNEYS OF VULNERABLE CHILDREN ACROSS SERVICES TO IDENTIFY MOST AT RISK CASES EARLY ON AND BETTER TARGET PREVENTATIVE SERVICES

ORGANISATION

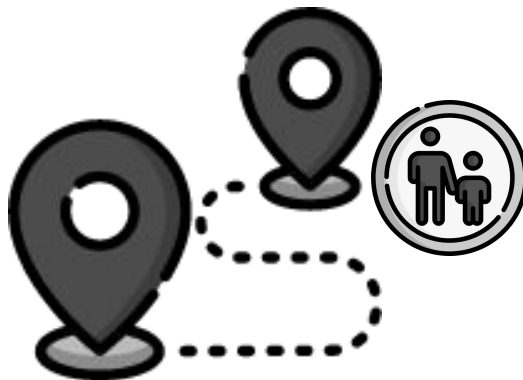
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



Service leads struggle to understand the journeys of the most vulnerable children across the different services offered by the council



MAPPING JOURNEYS OF VULNERABLE CHILDREN ACROSS SERVICES TO IDENTIFY MOST AT RISK CASES EARLY ON AND BETTER TARGET PREVENTATIVE SERVICES

ORGANISATION

PROBLEM

SOLUTION

DATA

APPROACH

IMPACT

Person centred data model that allowed journey and segmentation analysis
to understand population most at risk



MAPPING JOURNEYS OF VULNERABLE CHILDREN ACROSS SERVICES TO IDENTIFY MOST AT RISK CASES EARLY ON AND BETTER TARGET PREVENTATIVE SERVICES

ORGANISATION

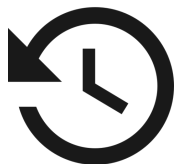
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT



20 years of data
(~1.3M records)



8+ datasets



MAPPING JOURNEYS OF VULNERABLE CHILDREN ACROSS SERVICES TO IDENTIFY MOST AT RISK CASES EARLY ON AND BETTER TARGET PREVENTATIVE SERVICES

ORGANISATION

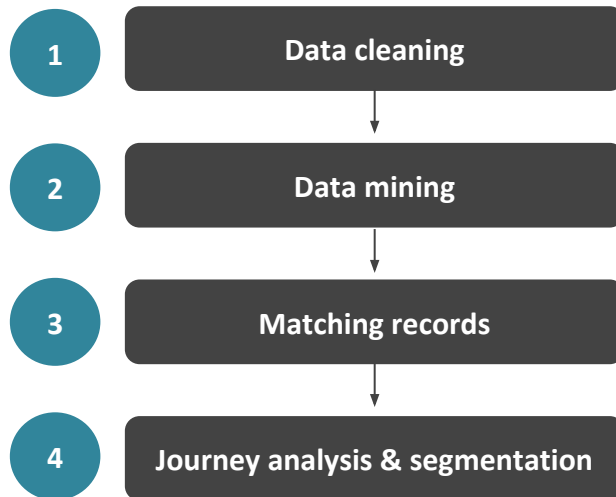
PROBLEM

SOLUTION

DATA

APPROACH

IMPACT





MAPPING JOURNEYS OF VULNERABLE CHILDREN ACROSS SERVICES TO IDENTIFY MOST AT RISK CASES EARLY ON AND BETTER TARGET PREVENTATIVE SERVICES

ORGANISATION

PROBLEM

SOLUTION

DATA















APPROACH

IMPACT

Informed the following
year's LA budget
allocation for
prevention services

Helped **identify young
people** most likely to
benefit from prevention
services

Measure effectiveness of
services working with
specific profiles

GOAL	AREA	COUNTRY	PARTNER
 <p>Mapping journeys of vulnerable children across services to identify early on and better target preventative services</p>	Children Services		Newcastle City Council
 <p>Reducing response times to citizen legal questions across Africa</p>	Legal		Barefoot Law
 <p>Increasing the efficiency of heart function assessment and diagnosis through echocardiography</p>	Health		CIBERCV (Biomedical Research Networking Centers) and Cardiology Department of the University Hospital of Salamanca
 <p>Improving Outcomes for Repeat/Frequent 911 Callers to Emergency Services</p>	Health		Memphis Fire Department
 <p>Quantifying traffic dynamics to better estimate and reduce air pollution exposure in London</p>	Health		Imperial College London, Transport for London, City of London, King's College London
 <p>Matching jobseekers with interventions to improve employment outcomes in Portugal</p>	Unemployment		Public Agency for Employment and Vocational Training (IEFP)
 <p>Identify children most at risk of not becoming school readiness</p>	Education		Wigan City Council

4

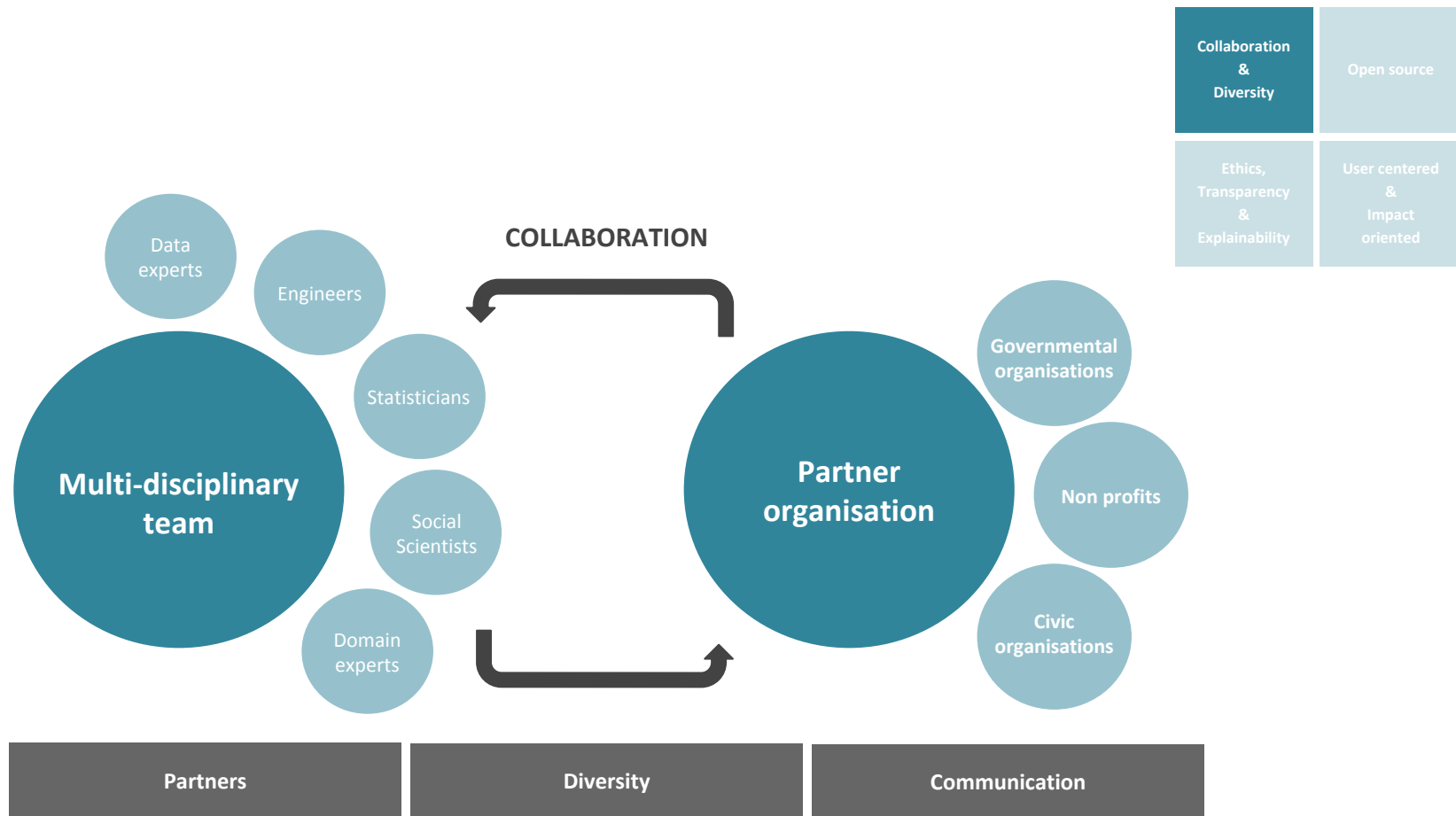
HOW DOES THIS COMPARE WITH A TYPICAL DATA SCIENCE PROJECT?

Collaboration
&
Diversity

Open source

Ethics, Transparency &
Explainability

Human centered
&
Impact oriented





Source: www.pinterest.com

Collaboration & Diversity	Open source
Ethics, Transparency & Explainability	User centered & Impact oriented

There is an opportunity to share learnings and scale solutions for common problems



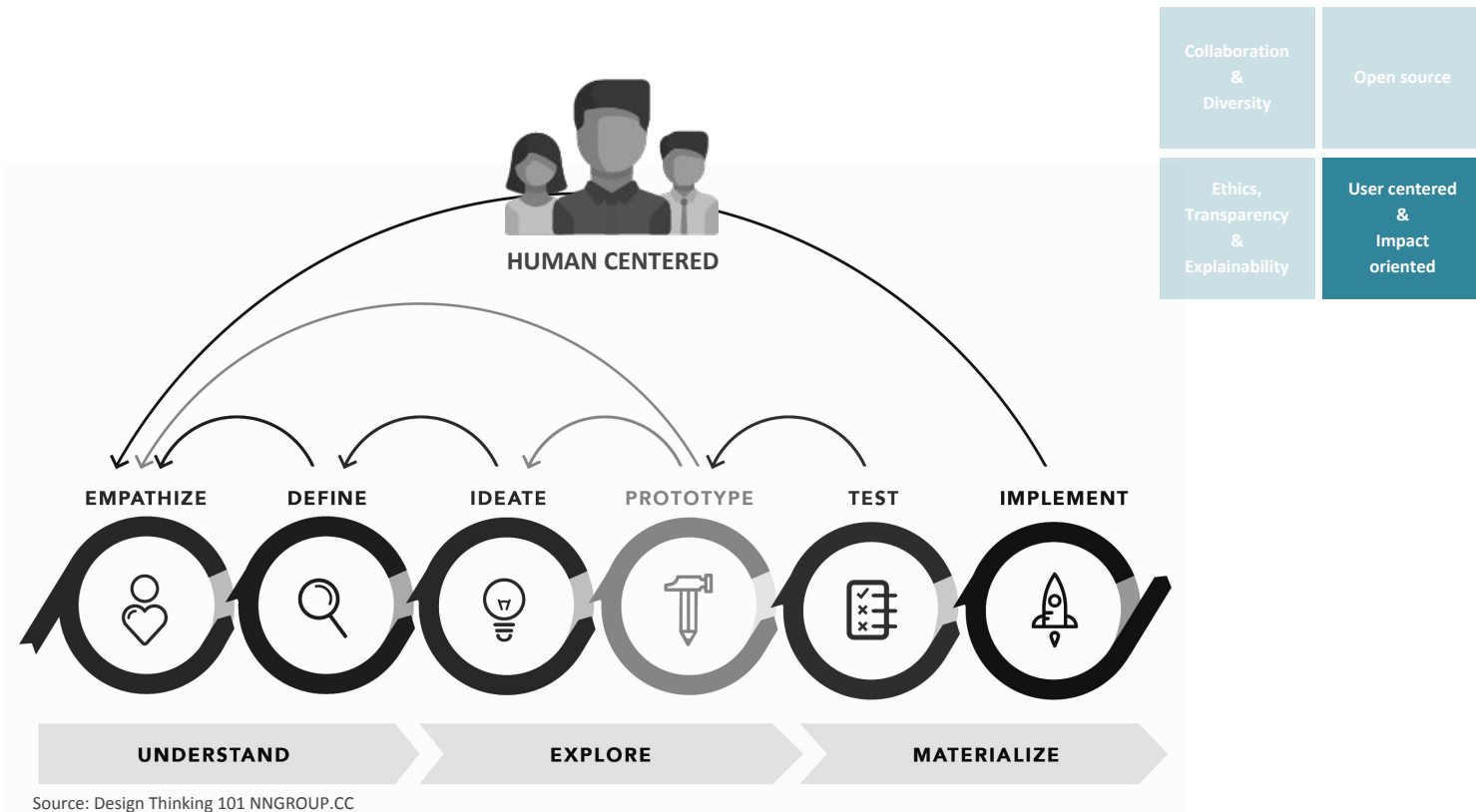
Collaboration
&
Diversity

Open source

Ethics,
Transparency
&
Explainability

User centered
&
Impact
oriented

It's our responsibility to make it as transparent as possible and communicate how it works to all the stakeholders



To achieve impact one needs to put humans at the heart of the project

5

BUILDING A **COMMUNITY** OF DATA SCIENTISTS FOR GOOD

Data Science For Social Good

Summer Fellowship



DATA SCIENCE FOR SOCIAL GOOD
PORTUGAL



The ROCKEFELLER
FOUNDATION



Center
for Inclusive
Growth

DataKind
USING DATA IN THE SERVICE OF HUMANITY



**Data for
Democracy**



AI for Good

**SOCIAL FINANCE
DIGITAL LABS**

Technology for social change



Source: Data Science for Social Good summer fellowship 2019



Source: Gianpietro Pucciariello

There is a growing community of people that want to use data science as a force for good

**“Once you start to have an impact on people’s lives
it is really hard not to continue to do that.”**

(Rayid Ghani)

DATA SCIENCE AS FORCE FOR GOOD

DSPT DAY 2019 | PORTO | 26.10.2019



SARA GUERREIRO DE SOUSA



sararogs@gmail.com



saraguerreirodesousa



SARA GUERREIRO DE SOUSA

ACADEMIC BACKGROUND

- **BSc in Applied Mathematics and Computation** @ Instituto Superior Técnico (IST)
- **Msc in Finance** @ NOVA School of Business and Economics
- **Data Science certification** @ General Assembly London

PROFESSIONAL JOURNEY

- **Teaching Assistant** @ NOVA School of Business and Economics
- **Analyst** @ MAZE (2015)
- **Data & Insights Associate** @ Social Finance Digital Labs (2016 - 2019)
- **Project Manager at Data Science for Social Good** summer fellowship @ Imperial College London (2019)