Data Science Community Meeting

Results of the Ideation Session

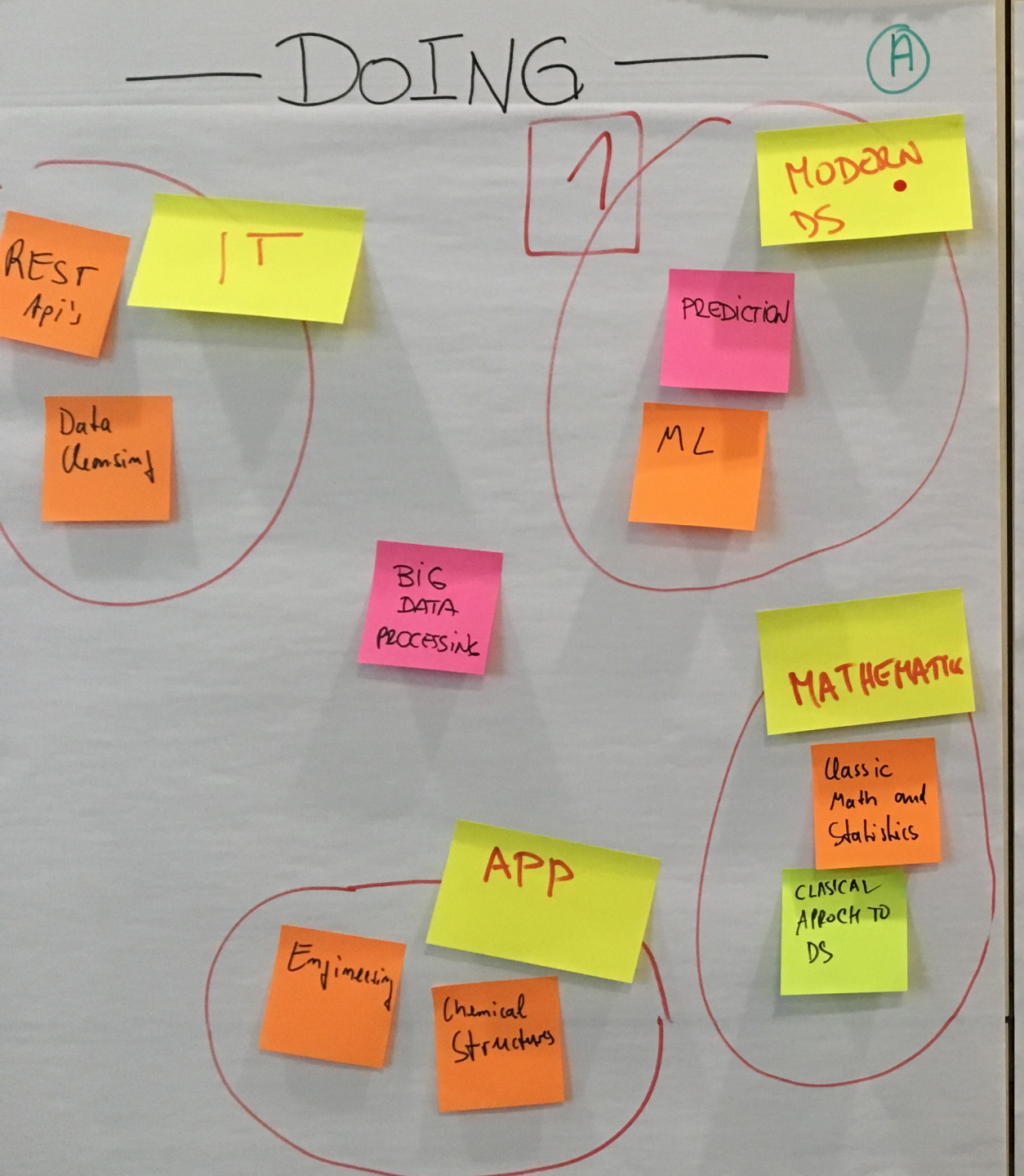
Hosted by

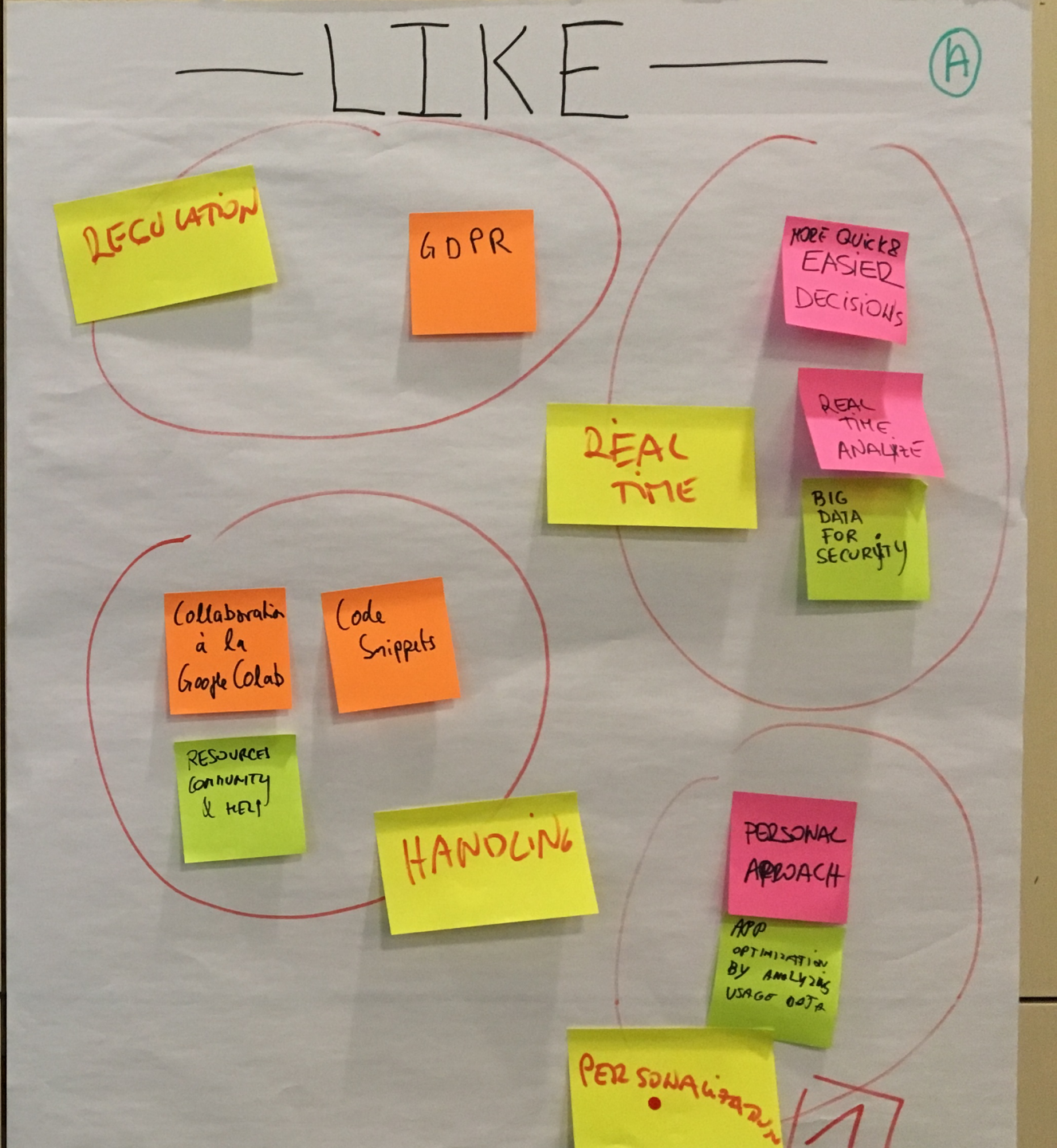
IBM Cloud Development Lab

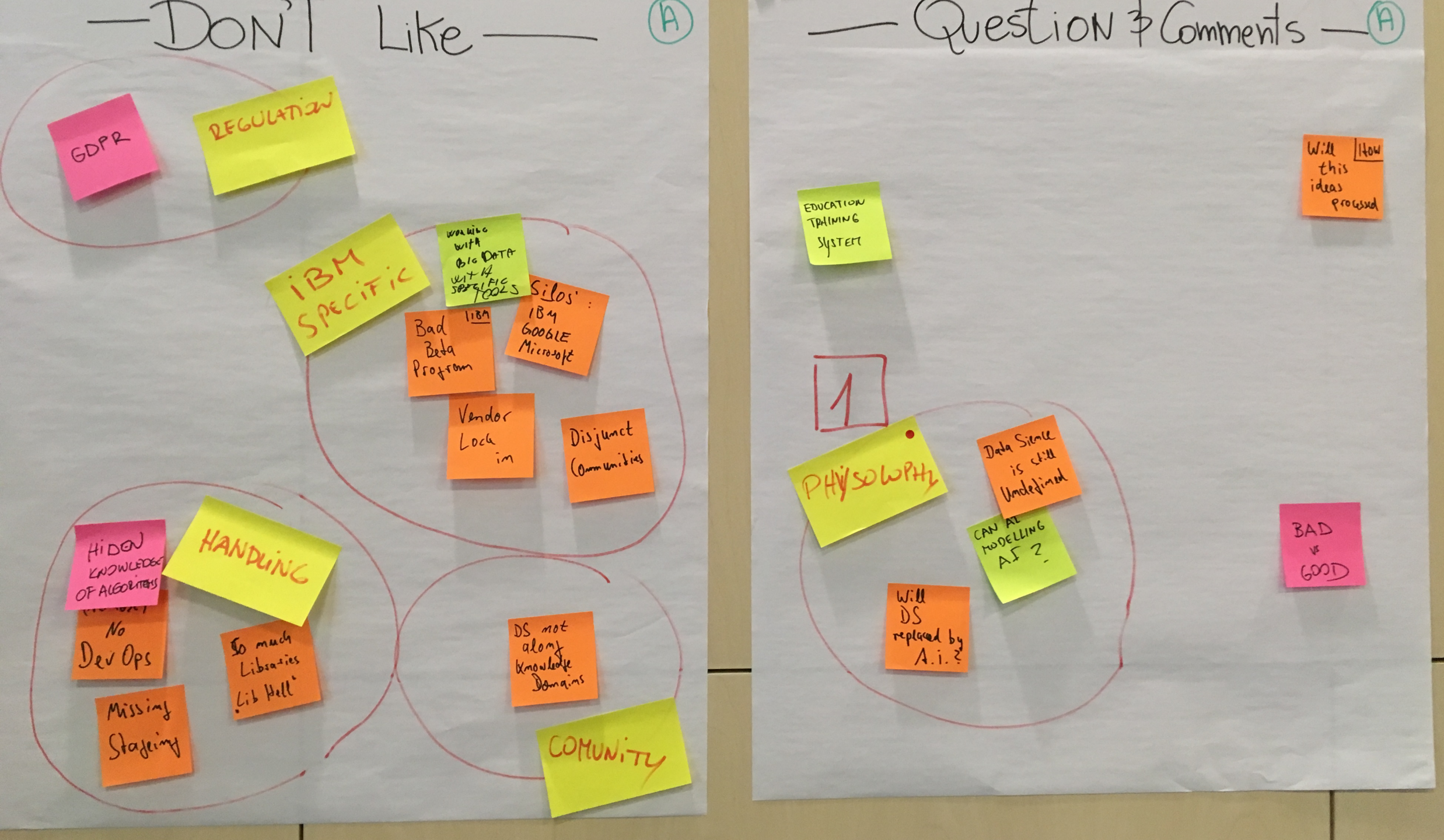
Germany

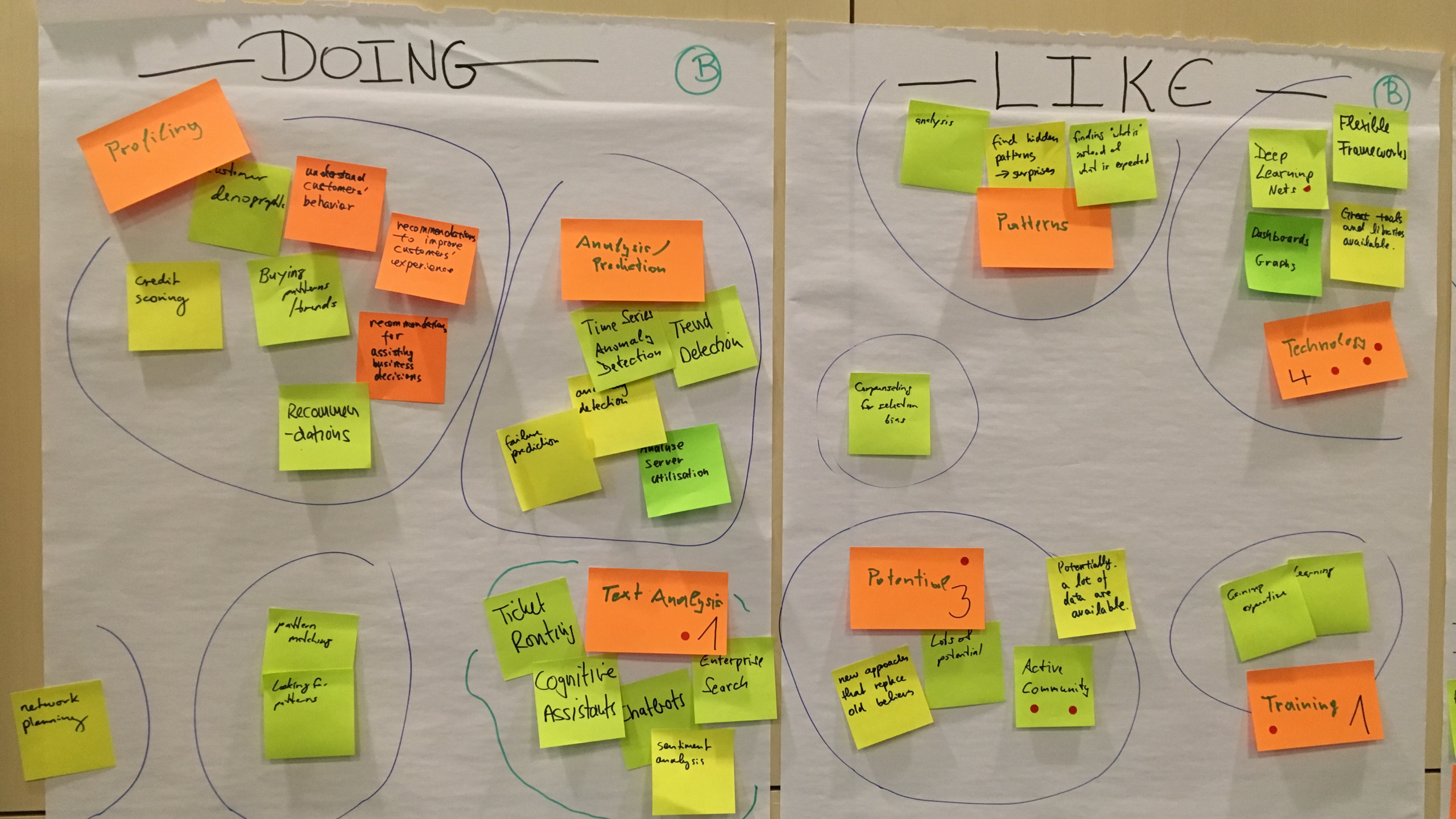
Data Science Today

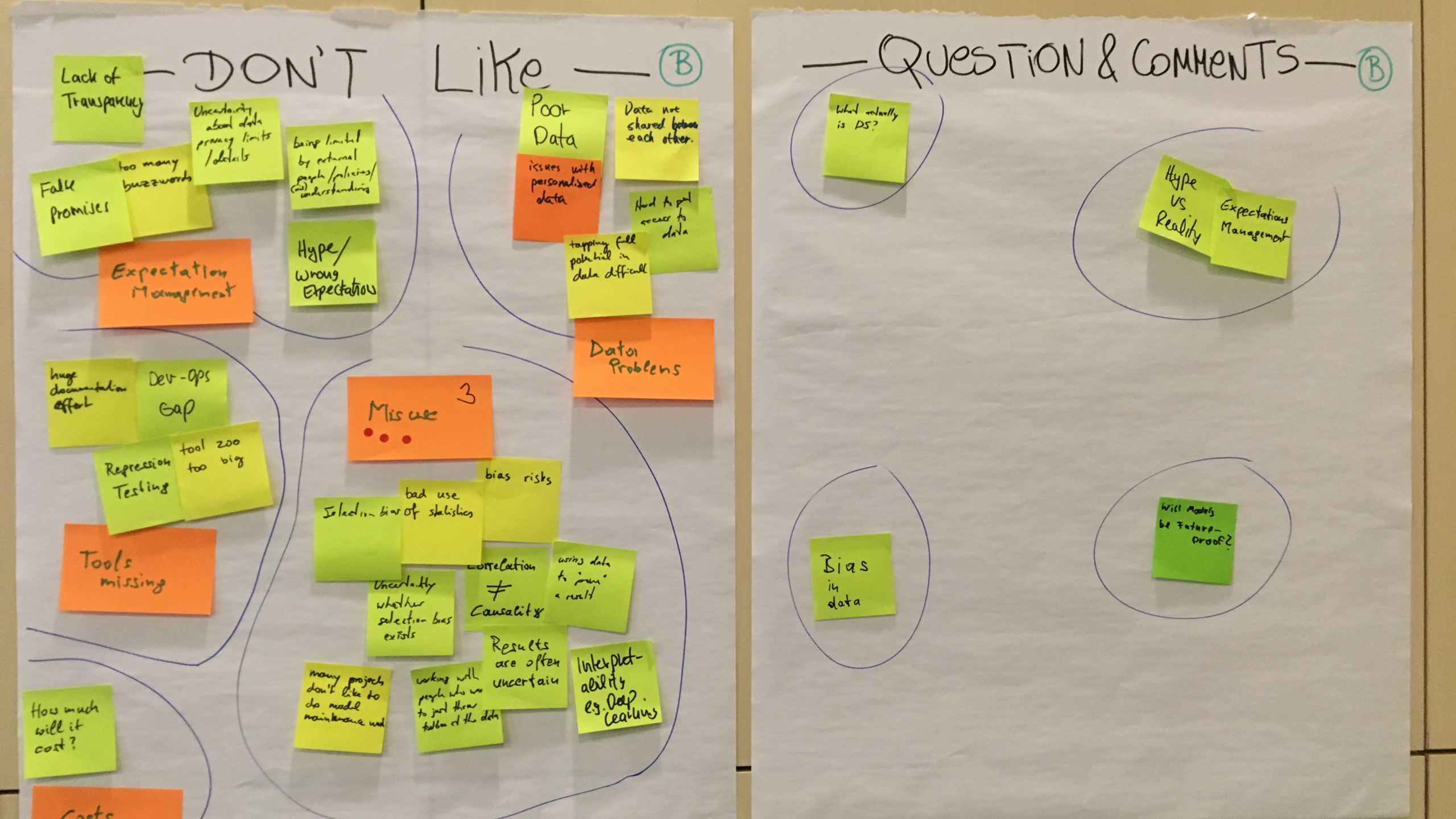
Doing / Like / Don‘t Like / Comments

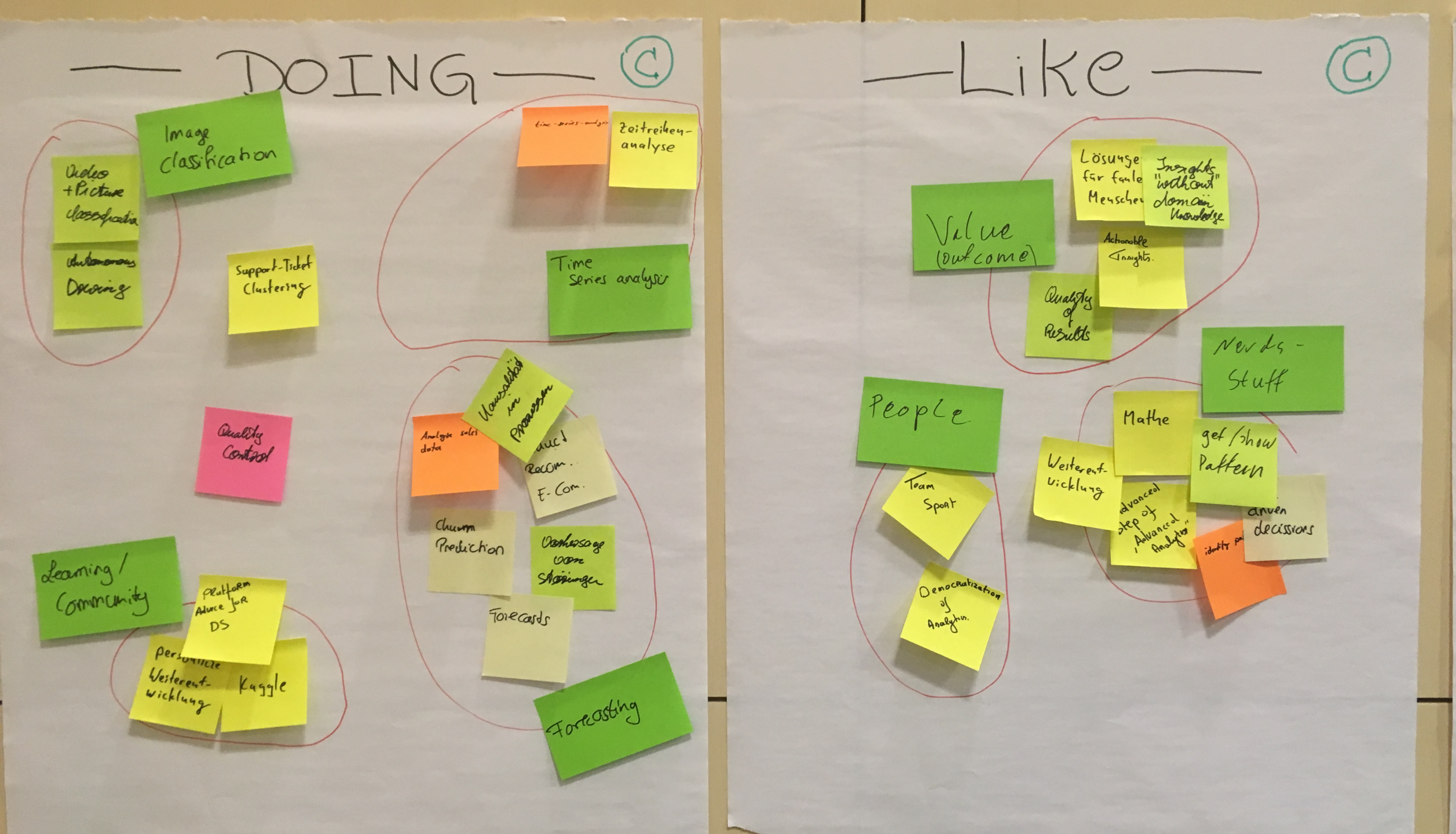

DOING
A
1
MODERN DS
REST Api's
IT
PREDICTION
Data Cleansing
ML
BIG DATA PROCESSING
MATHEMATICS
Classic Math and Statistics
APP
CLASICAL APROCH TO DS
Engineering
Chemical Structures

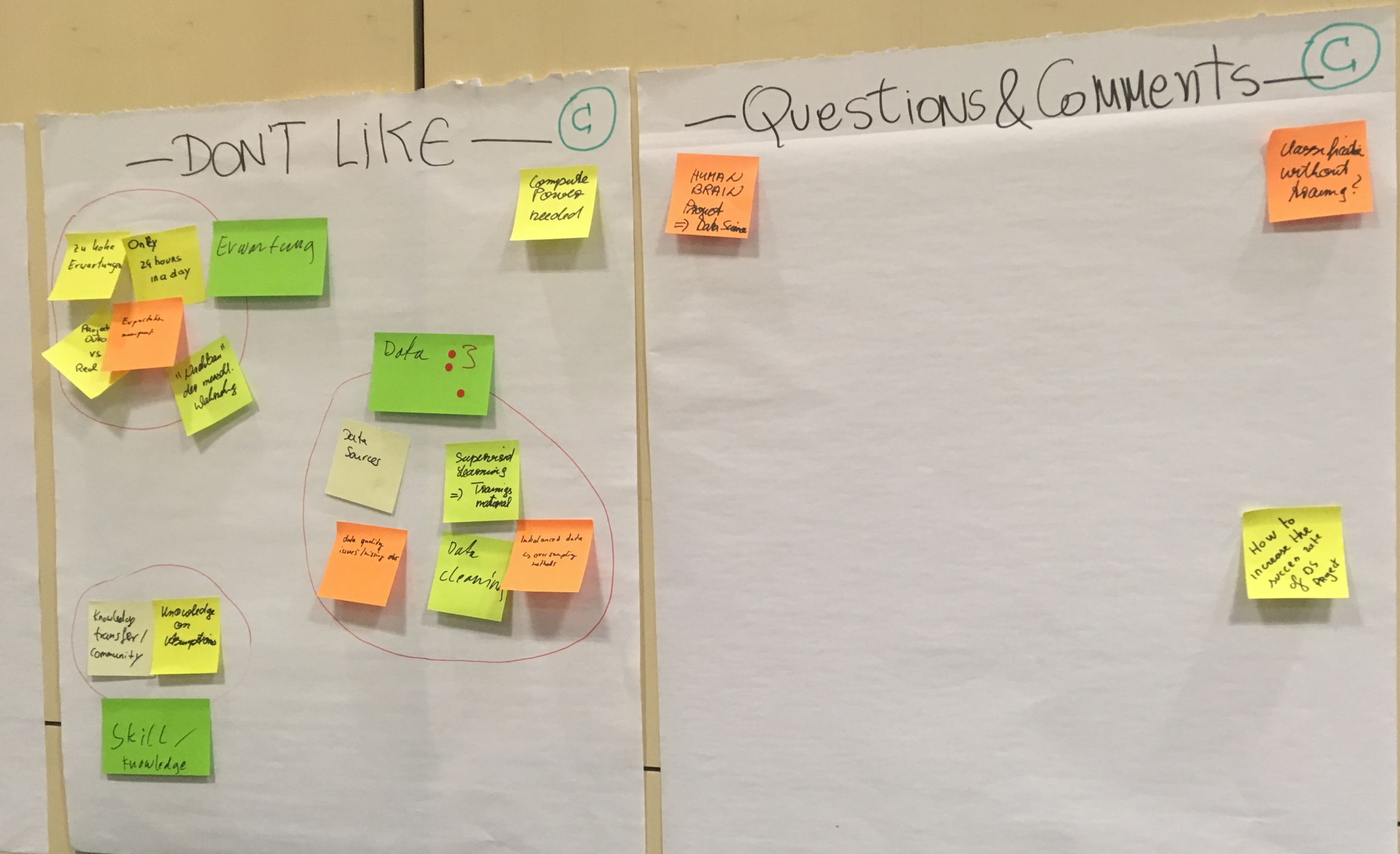

LIKE
A
REGULATION
GDPR
MORE QUICK & EASIER DECISIONS
REAL TIME
REAL TIME ANALYZE
BIG DATA FOR SECURITY
Collaboration à la Google Colab
Code Snippets
RESOURCES COMMUNITY & HELP
HANDLING
PERSONAL APROACH
APP OPTIMIZATION BY ANALYZING USAGE DATA
PERSONALIZATION
1

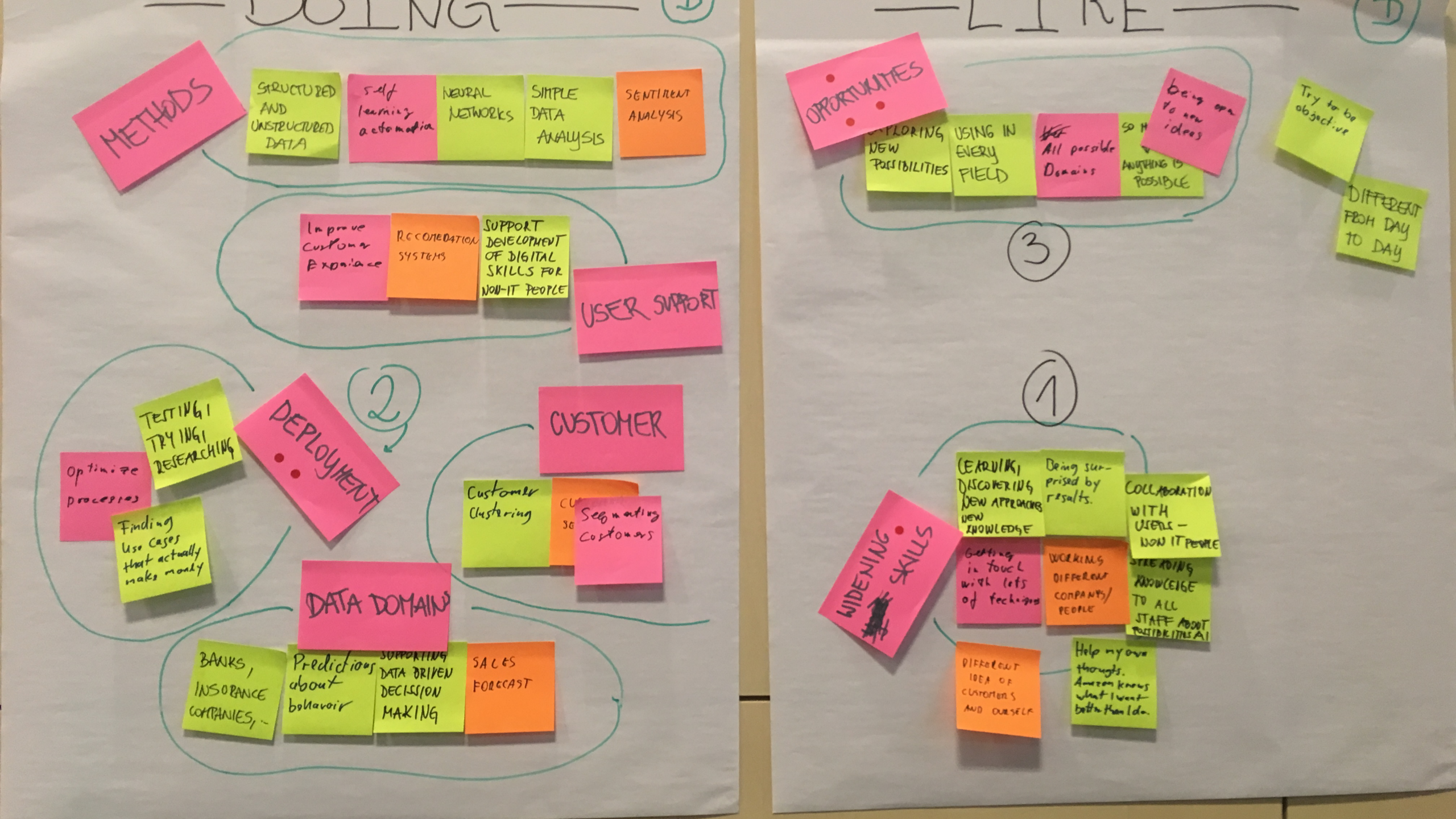

Don't Like
A
GDPR
REGULATION
IBM SPECIFIC
working with Big DATA with SPECIFIC TOOLS
Silos: IBM GOOGLE Microsoft
Bad Beta Program
IBM
Vendor Lock in
Disjunct Communities
HANDLING
HIDEN KNOWLEDGE OF ALGORITHMS
No Dev Ops
So much Libraries Lib Hell
Missing Stapling
DS not alone knowledge Domains
COMUNITY
Question & Comments
A
Will this ideas processed
How
EDUCATION TRAINING SYSTEM
1
PHYSOPHY
Data Sience is still Undefined
CAN AI MODELLING AI ?
Will DS replaced by A.I. ?
BAD vs GOOD

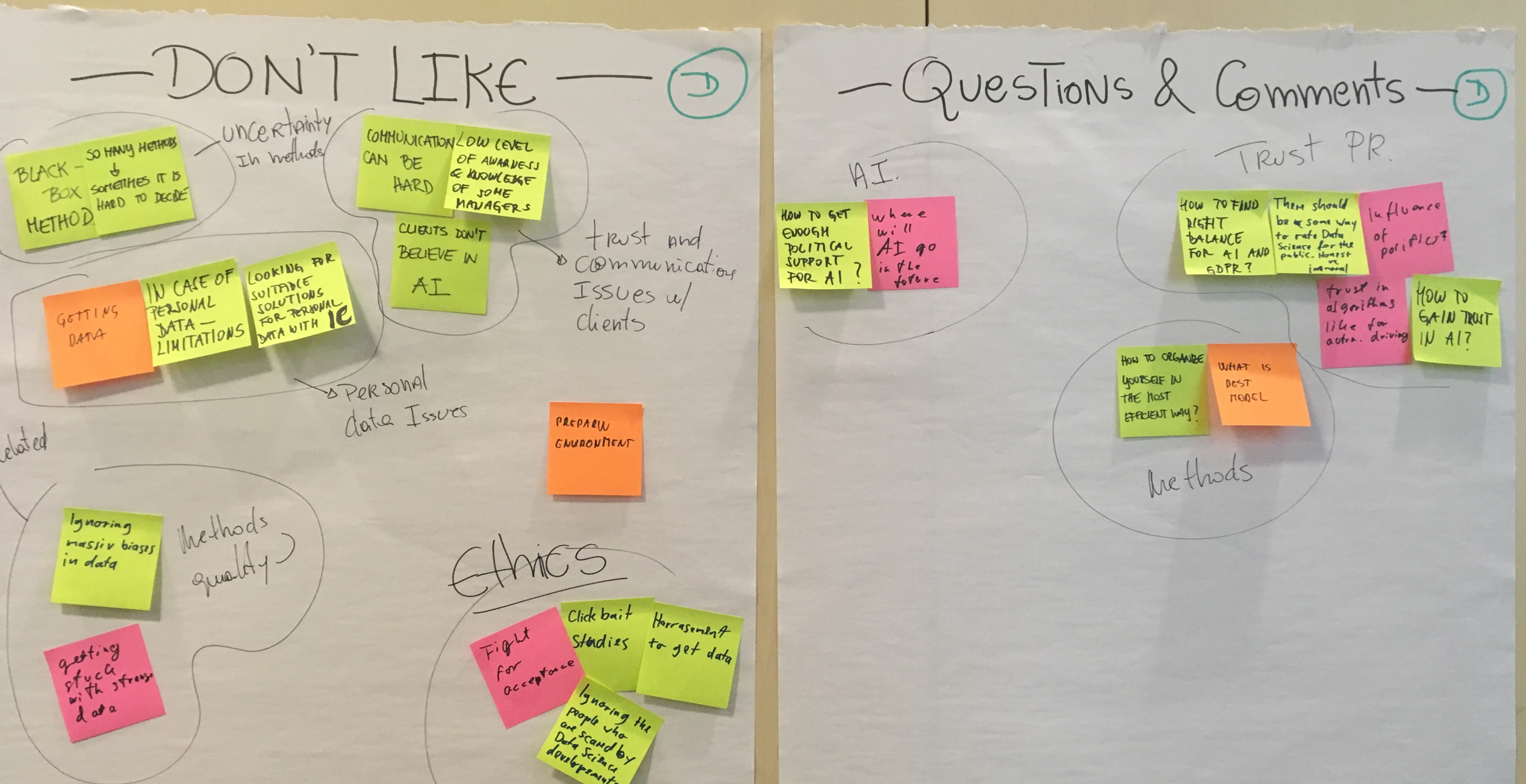

DOING
B
Profiling
customer demographic
understand customers' behavior
recommendations to improve customers' experience
Credit scoring
Buying patterns /trends
recommendations for assisting business decisions
Recommen-dations
Analysis / Prediction
Time Series Anomaly Detection
Trend Detection
anomaly detection
failure prediction
Analyse Server utilisation
network planning
pattern matching
Looking for patterns
Ticket Routing
Text Analysis 1
Cognitive Assistants
Chatbots
Enterprise Search
sentiment analysis
LIKE
B
analysis
find hidden patterns → surprises
finding "what is" instead of what is expected
Patterns
Deep Learning Nets
Flexible Frameworks
Dashboards Graphs
Great tools and libraries available.
Technology 4
Compensating for selection bias
Potential 3
Potentially a lot of data are available.
Lots of potential
new approaches that replace old believes
Active Community
Training 1
Training expertise
Learning


DON'T Like
B
Lack of Transparency
Uncertainty about data privacy limits / details
being limited by external people / policies / (old) understanding
too many buzzwords
Fake promises
Expectation Management
Hype / Wrong Expectation
Poor Data
issues with personalised data
Data not shared between each other.
Hard to get access to data
tapping full potential in data difficult
Data Problems
huge documentation effort
Dev-Ops Gap
Repression Testing
tool zoo too big
Tools missing
Misuse 3
bias risks
bad use of statistics
Correlation ≠ Causality
using data to prove a result
Uncertainty whether selection bias exists
Results are often uncertain
Interpretability e.g. Deep Learning
many projects don't like to do model maintenance
working with people who use to just throw tables at the data
How much will it cost?
QUESTION & COMMENTS
B
What actually is DS?
Hype vs Reality
Expectations Management
Bias in data
Will Models be Future-proof?


DOING
Image Classification
Video + Picture classification
Support-Ticket Clustering
Zeitreihen-analyse
Time series analysis
Quality Control
Learning / Community
Kaggle
Churn Prediction
Forecasts
Forecasting
LIKE
Value (outcome)
Actionable Insights
Quality of Results
People
Team Sport
Democratization of Analytics
Nerds - Stuff
Mathe
get / show Pattern
Weiterentwicklung
advanced step of "Advanced" Analytics
driven decisions

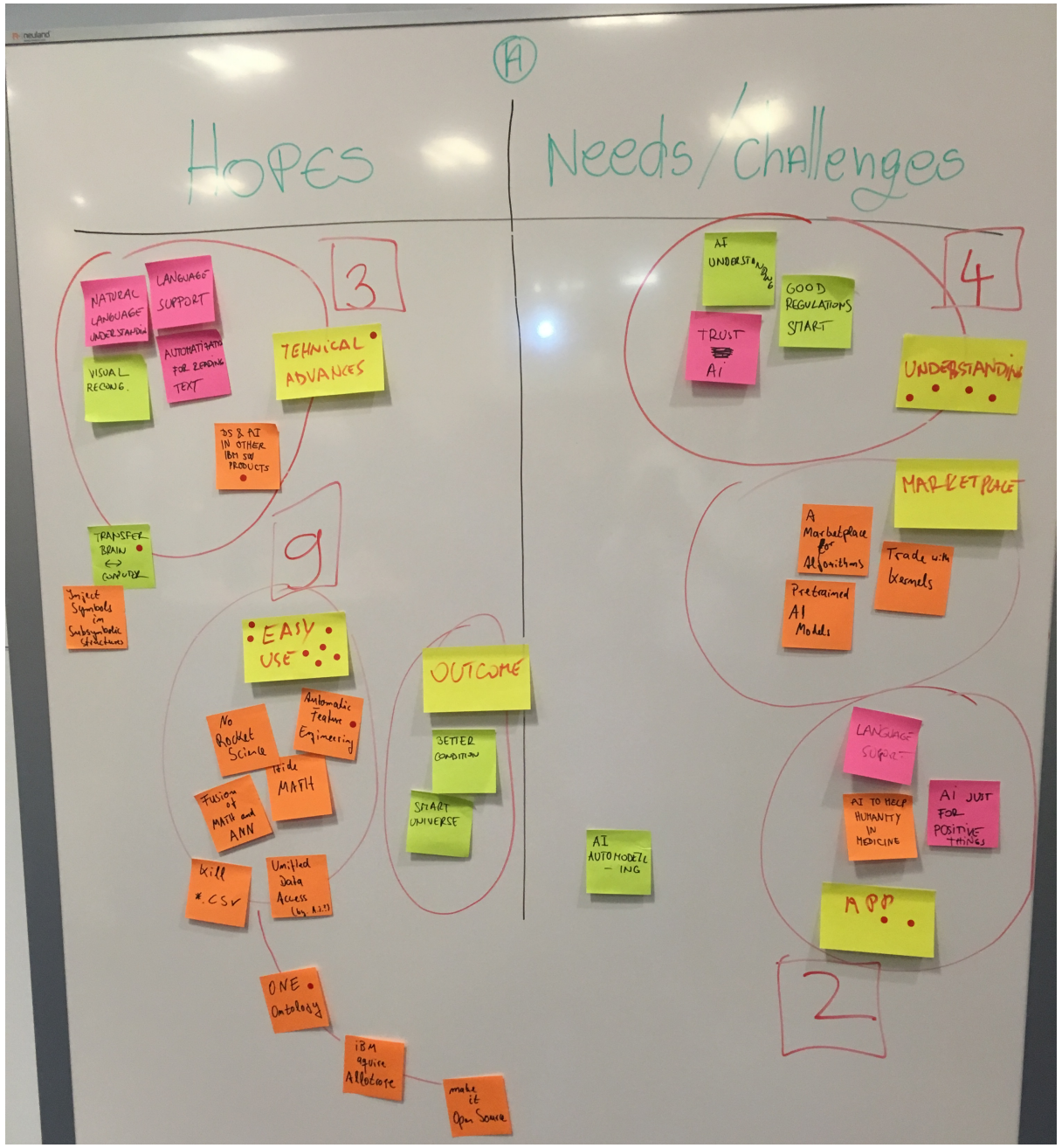

DON'T LIKE
Compute Power needed
Erwartung
Only 24 hours in a day
Data :3
Data Sources
Supervised Learning
=) Trainings material
Data cleaning
Skill / Knowledge
Knowledge transfer / Community
Questions & Comments
HUMAN BRAIN Project
=) Data Scaine
How to increase the succes rate of DS Project


DOING
METHODS
STRUCTURED AND UNSTRUCTURED DATA
self learning automatic
NEURAL NETWORKS
SIMPLE DATA ANALYSIS
SENTIMENT ANALYSIS
Improve Customer Experience
RECOMENDATION SYSTEMS
SUPPORT DEVELOPMENT OF DIGITAL SKILLS FOR NON-IT PEOPLE
USER SUPPORT
2
DEPLOYMENT
TESTING! TRYING! RESEARCHING
Optimize processes
Finding Use cases that actually make money
CUSTOMER
Customer Clustering
Segmenting Customers
DATA DOMAINS
BANKS, INSORANCE COMPANIES, ..
Predictions about behavior
SUPPORTING DATA DRIVEN DECISION MAKING
SALES FORECAST
LIKE
OPPORTUNITIES
EXPLORING NEW POSSIBILITIES
USING IN EVERY FIELD
All possible Domains
being open to new ideas
Try to be objective
DIFFERENT FROM DAY TO DAY
3
1
WIDENING SKILLS
LEARNING! DISCOVERING NEW APPROACHES NEW KNOWLEDGE
Being surprised by results.
COLLABORATION WITH USERS – NON IT PEOPLE
Getting in touch with lots of technics
WORKING DIFFERENT COMPANYS/ PEOPLE
SPREADING KNOWLEDGE TO ALL STAFF ABOUT POSSIBILITIES AI
DIFFERENT IDEA OF CUSTOMERS AND OURSELF
Help my own thoughts. Amazon knows what I want better than I do.

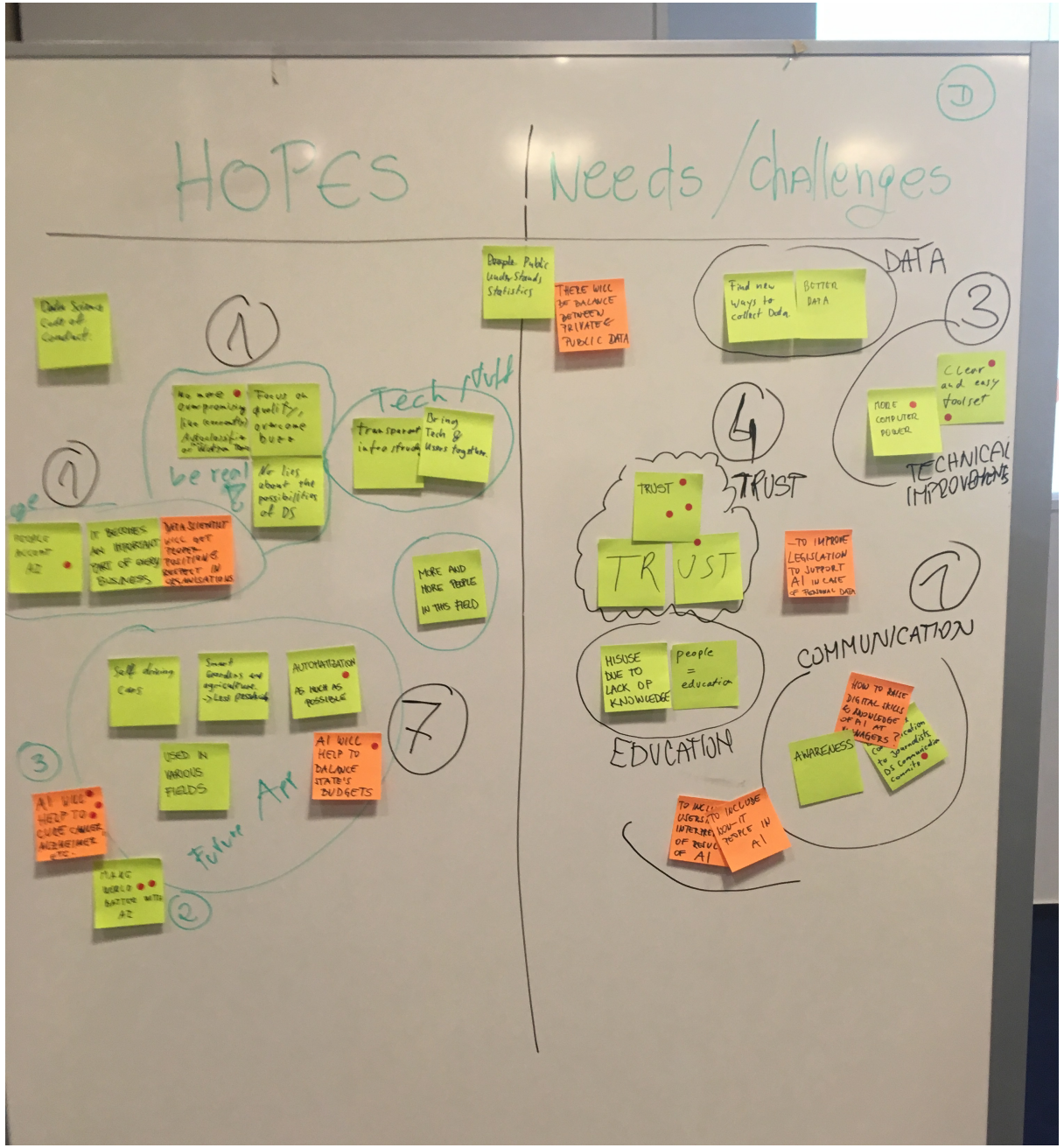

DON'T LIKE
BLACK - BOX METHOD
SO MANY METHODS
SOMETIMES IT IS HARD TO DECIDE
uncertainty in methods
COMMUNICATION CAN BE HARD
LOW LEVEL OF AWARENESS & KNOWLEDGE OF SOME MANAGERS
CLIENTS DON'T BELIEVE IN AI
trust and communication Issues w/ clients
GETTING DATA
IN CASE OF PERSONAL DATA — LIMITATIONS
LOOKING FOR SUITABLE SOLUTIONS FOR PERSONAL DATA WITH ie
Personal data Issues
PREPARE ENVIRONMENT
elated
Ignoring massiv biases in data
methods quality
getting stuck with strong data
Ethics
Fight for acceptance
Click bait studies
Harrasement to get data
Ignoring the people who are scaard by Data Science developments
Questions & Comments
A.I.
HOW TO GET ENOUGH POLITICAL SUPPORT FOR AI ?
where will AI go in the future
Trust PR.
HOW TO FIND RIGHT BALANCE FOR AI AND GDPR ?
There should be some way to rate Data Science for the public. Honest vs immoral
influance of politics?
trust in algorithms like for auton. driving
HOW TO GAIN TRUST IN AI?
HOW TO ORGANIZE YOURSELF IN THE MOST EFFICIENT WAY?
WHAT IS BEST MODEL
methods

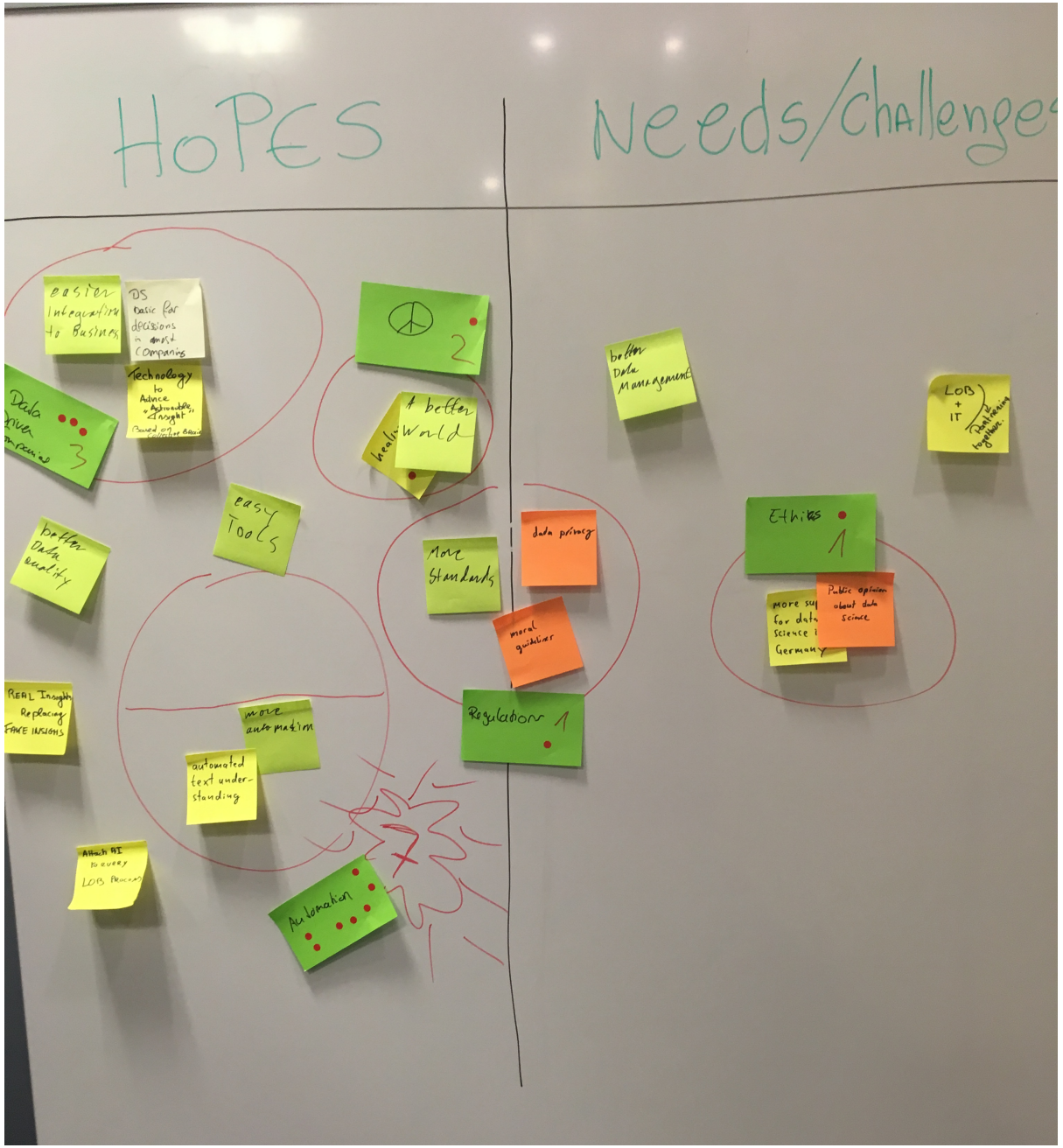
Data Science in 5 Years

Hopes / Needs & Challenges


A
Hopes
Needs/Challenges
3
NATURAL LANGUAGE UNDERSTANDING
LANGUAGE SUPPORT
AUTOMATIZATION FOR READING TEXT
VISUAL RECONG.
TEHNICAL ADVANCES
DS & AI IN OTHER IBM SW PRODUCTS
TRANSFER BRAIN ↔ COMPUTER
Inject Symbols in Subsymbolic Structures
9
EASY USE
No Rocket Science
Automatic Feature Engineering
Hide MATH
Fusion of MATH and ANN
Kill *.CSV
Unified Data Access (by A.I.?)
ONE Ontology
IBM aquire Allotrose
make it Open Source
OUTCOME
BETTER CONDITION
SMART UNIVERSE
AI AUTOMODELL-ING
4
AI UNDERSTANDING
GOOD REGULATIONS START
TRUST AI
UNDERSTANDING
MARKETPLACE
A Marketplace for Algorithms
Trade with kernels
Pretrained AI Models
LANGUAGE SUPPORT
AI TO HELP HUMANITY IN MEDICINE
AI JUST FOR POSITIVE THINGS
APP
2

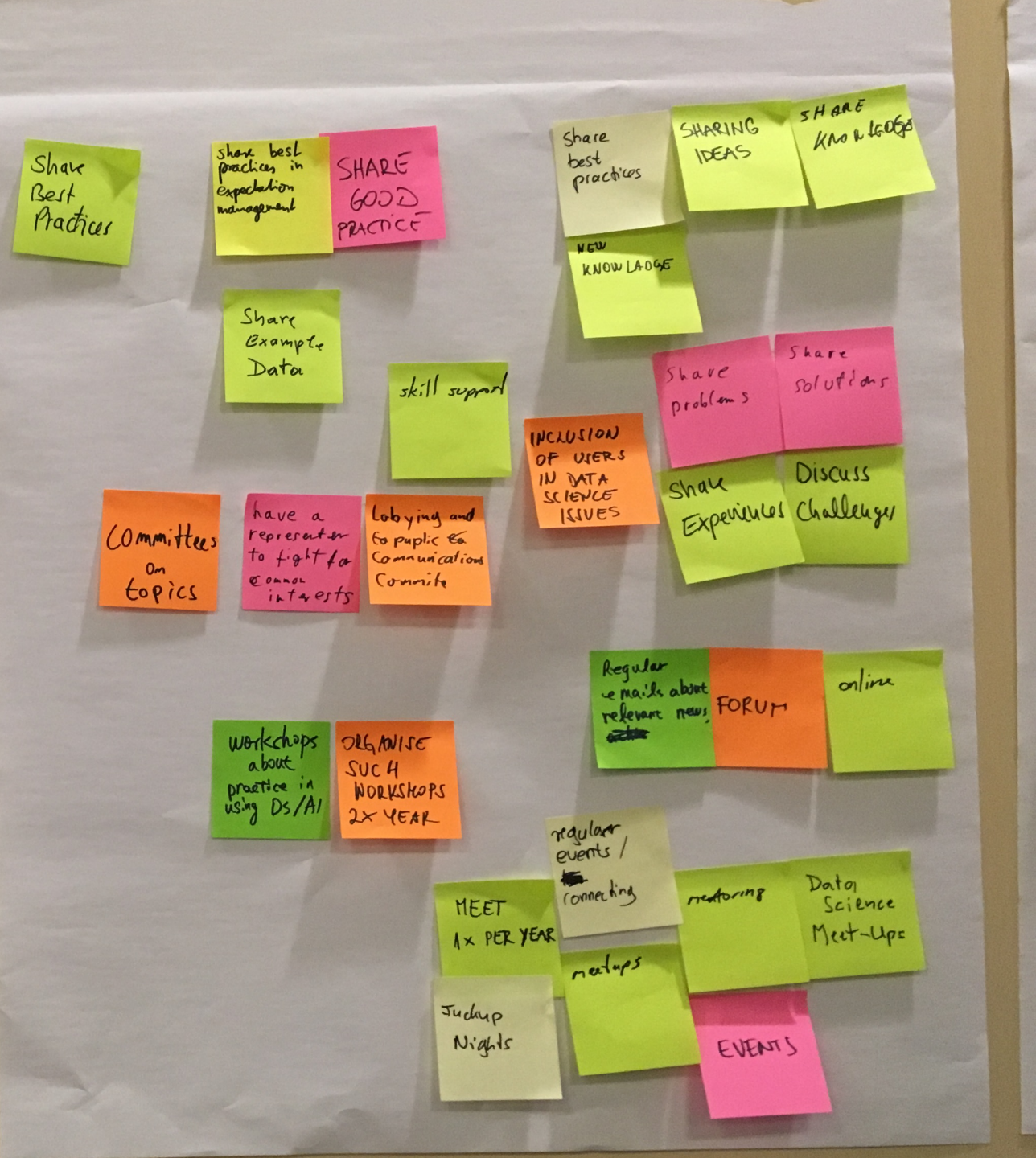
B
Hopes
Needs/Challenges
data access & sharing 5
Data Access (no Silos)
All data are perfect
easy data integration
easier access to data
Data Sharing
Data being shared between each other
business 3 + data science
Full Management Attention
better understanding of what is DS
more root cause analyses than high level analyses
business can deal with uncertainty in data science projects
Management has realistic expectations
Path of Enlightenment
technology
better/faster methods
bigger computing
More suitable candidates on the job market
Standards
worldwide commitment to data security/privacy standards
standards
Taming Tool Zoo
Data Science Engineering Practice
ethics 3
AI will not replace humans regarding DS
AI won't be used for bad things
public awareness of data privacy
Autonomous Cars
Spoken natural language conversations
applications 2
methods & improvement 1
allow randomness to avoid bias/local maxima issues
DS/AI products are safe & reliable
Interpretability (Deep Learning)
access frameworks
AI Maturity
Singularity
George Orwell Total Control
Take away Fear
Fake News
Social Net Bubbles
don't replace thinking by data science
society impact 4
skills 6
Empower Employees (Integration)
education of data scientists
better skills knowledge throughout companies
best practice 2
Quality Awareness
everyone documents data knowledge/assumptions properly
standards
regulations 1
Gov't intervention
Data security
GDPR
Compete with less restrictive societies


HOPES
Needs/challenges
Data Science Code of Conduct
People Public Understands Statistics
THERE WILL BE BALANCE BETWEEN PRIVATE & PUBLIC DATA
DATA
Find new ways to collect Data
BETTER DATA
Clear and easy toolset
MORE COMPUTER POWER
TECHNICAL IMPROVEMENTS
Tech stuff
Transparent infrastructure
Bring Tech & Users together
No more overpromising
Focus on quality, overcome buzz
be real
No lies about the possibilities of DS
IT BECOMES AN IMPORTANT PART OF EVERY BUSINESS
DATA SCIENTIST WILL GET PROPER POSITIONS RESPECT IN ORGANISATIONS
MORE AND MORE PEOPLE IN THIS FIELD
TRUST
TRUST
TO IMPROVE LEGISLATION TO SUPPORT AI IN CASE OF PERSONAL DATA
COMMUNICATION
HISUSE DUE TO LACK OF KNOWLEDGE
people = education
EDUCATION
HOW TO RAISE DIGITAL SKILLS & KNOWLEDGE OF AI AT MANAGERS
AWARENESS
Self driving cars
Smart Grandins and agriculture
AUTOMATIZATION AS MUCH AS POSSIBLE
USED IN VARIOUS FIELDS
AI WILL HELP TO BALANCE STATE'S BUDGETS
AI WILL HELP TO CURE CANCER, ALZHEIMER ETC.
MAKE WORLD BETTER WITH AI
Future Art


HOPES
Needs/Challenges
easier Integration to Business
DS basic for decisions in most Companies
Technology to Advice "Actionable Insight"
Data Driven Companies
better Data quality
easy Tools
A better World
better Data Management
LOB + IT
Ethics
Public opinion about data science
More Standards
data privacy
moral guidelines
Regulation
more automation
automated text understanding
Attach AI to every LOB Process
Automation

The Future of this Community

COMMUNITY ACTIVITIES
-Tech Guide
↳ Wiki
Guidelines & Best practic
INVEST AI IN HEALTH FIELD TO CURE BAD ILLNESSES
a event like this every Quarter
PRESENTATION OF OUR PROJECTS
Present new Tools / Flameworks ....
Present case studies
CONTACT DATA SCIENTISTS
visit other data science teams
CONTACT DIFFRENT COMPANYS
do something fun
Best Practices for Use case deployment -common agreement on topics to resolve
INFRA STRUCTURE
development support
Discuss new Technologies
attend the trainings in data science methods together
cross industry - Solution galleries - analytics notebook - sample data - ability to list own products / services
RAISE DIGITAL SKILLS OF NON-IT PEOPLE


Share Best Practices
share best practices in expectation management
SHARE GOOD PRACTICE
Share best practices
SHARING IDEAS
SHARE KNOWLEDGE
NEW KNOWLADGE
Share Example Data
skill support
INCLUSION OF USERS IN DATA SCIENCE ISSUES
Share problems
Share solutions
Share Experiences
Discuss Challenges
Committees on topics
have a representation to fight for common interests
Lobying and to puplic ea Communications Committe
Regular e mails about relevant news
FORUM
online
workshops about practice in using DS/AI
ORGANISE SUCH WORKSHOPS 2x YEAR
regular events / connecting
MEET 1x PER YEAR
mentoring
Data Science Meet-Ups
Judup Nights
meetups
EVENTS

Your Engagement

Time
money
locations

papers

TALK WHAT
WE DO

Working
Code in
Github

Open Source
vs.
Vendor Lock-in

Tooling
Webinars

Feedback

- Advocat
- PR work

3 x
per
year
Meet

KNOWLEDGE
DISSEMINATION
TRAINNINGS

documentation

Share
ideas /
solutions

talk about
what
failed

sharing
code

Present
Projects

Present
Technology
Topics

TO
PRESENT
DS PROJECTS
WE ARE
DOING

presentations/
talks

Other Ideas / Questions / Comments

drive ethics/
awareness of
risks & potential
of data science
in public

certification
?

Standards
between
Companies

„ONE
DS"
Initiative

Feedback
LOOP

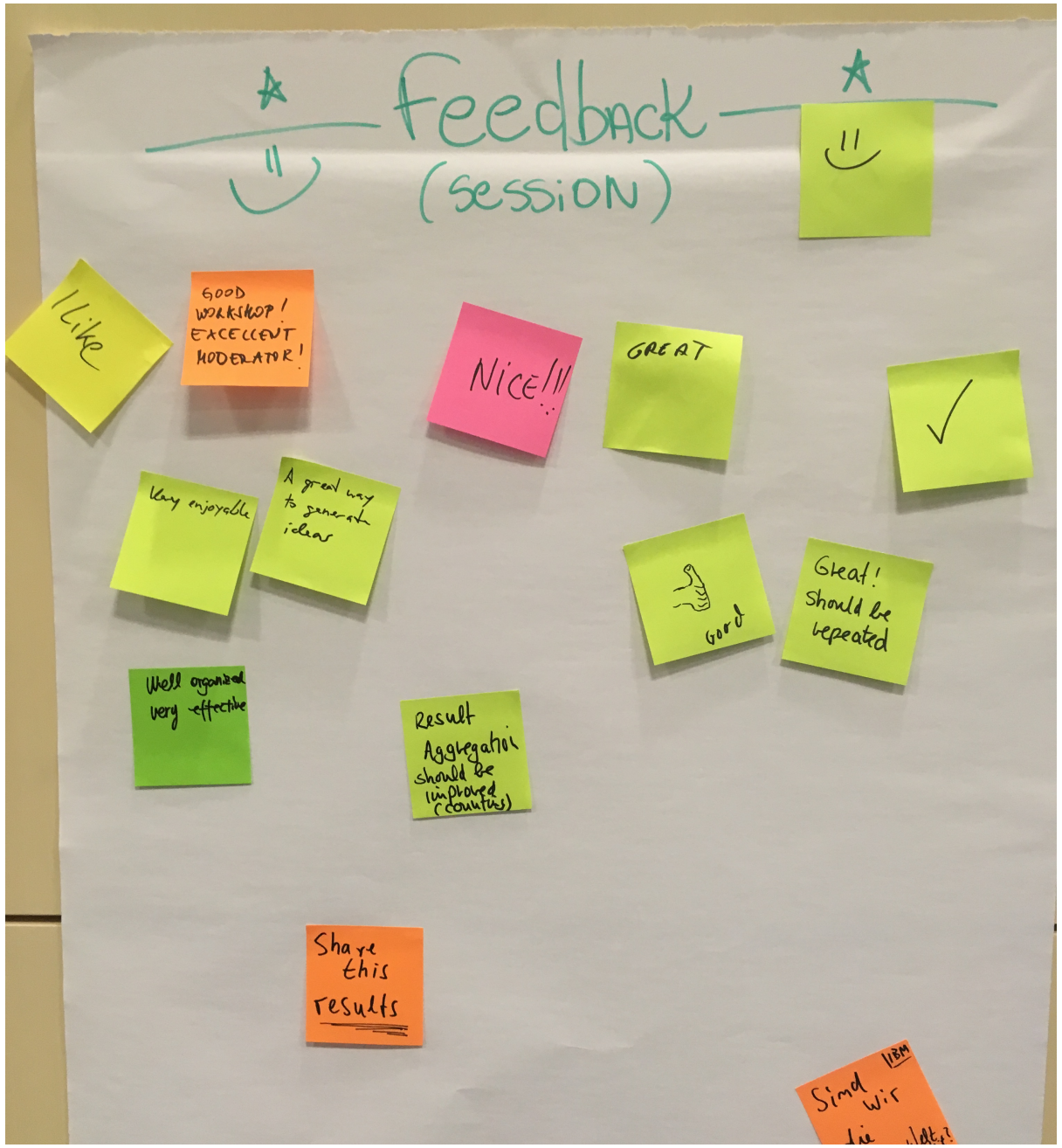
how to make
sure that
community
is active?

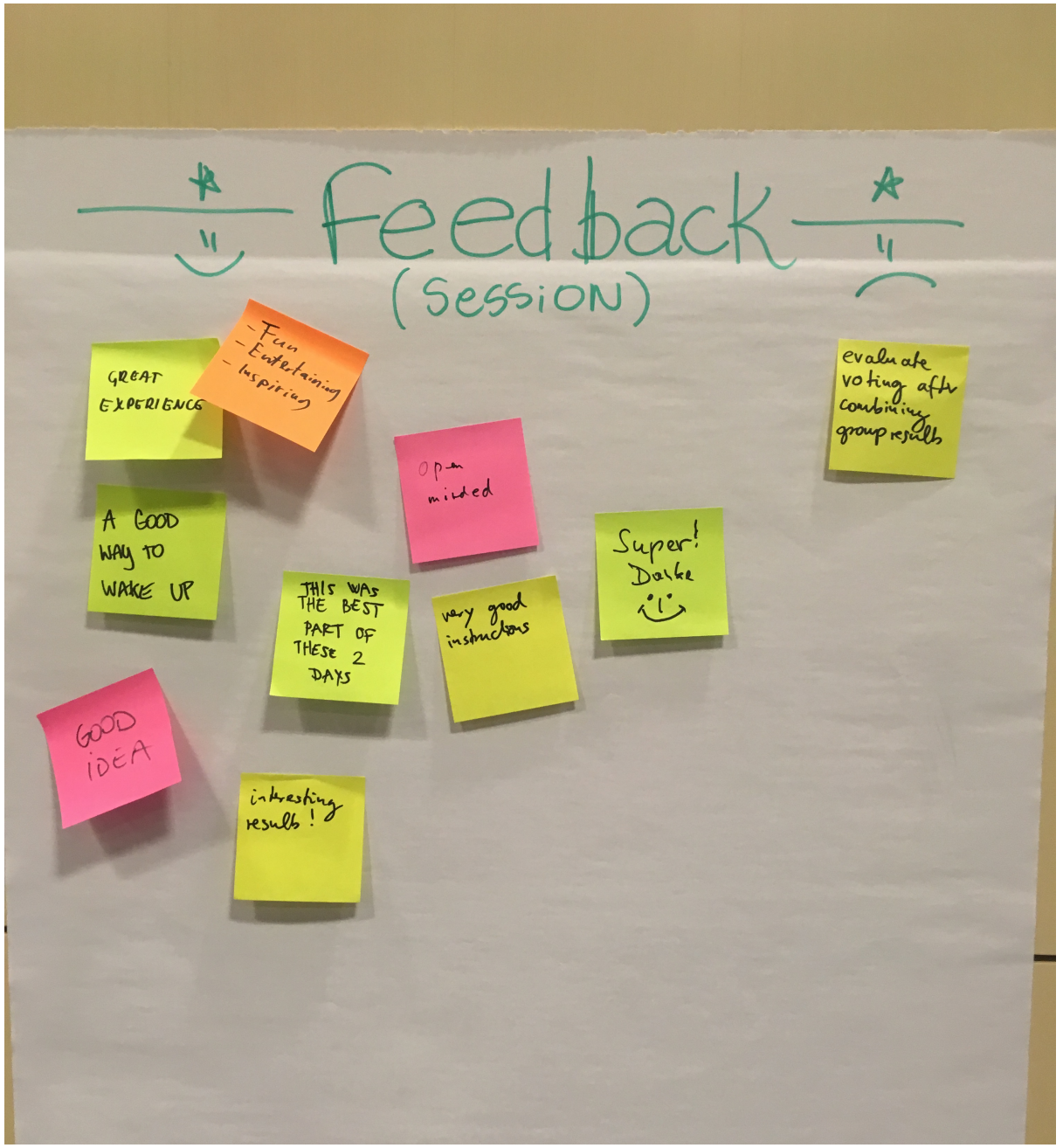
International
Partner
Groups

Chapters

Open
Challenges
$

Feedback


feedback
(session)
I like
GOOD WORKSHOP! EXCELLENT MODERATOR!
Nice!!!
GREAT
Very enjoyable
A great way to generate ideas
Gord
Great! Should be repeated
Well organized very effective
Result Aggregation should be improved (counters)
Share this results
Sind wir


feedback
(session)
GREAT EXPERIENCE
- Fun
- Entertaining
- Inspiring
evaluate voting after combining group results
Open minded
A GOOD WAY TO WAKE UP
Super! Danke
THIS WAS THE BEST PART OF THESE 2 DAYS
very good instructions
GOOD IDEA
interesting results!