Some Biases in Data Science

Allan Hanbury

TU WIEN

COMPLEXITY SCIENCE HUB VIENNA


Data
Algorithm
Results
BIASED


Creation
Data
Algorithm
Results

BUSINESS NEWS OCTOBER 10, 2018 / 5:12 AM / 7 MONTHS AGO

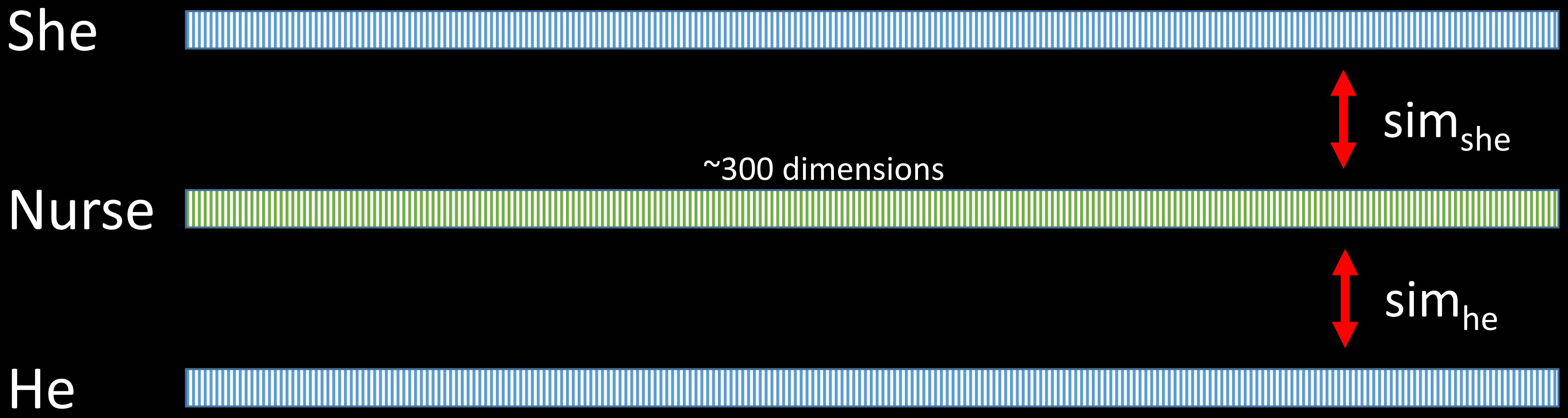
Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin 8 MIN READ

SAN FRANCISCO (Reuters) - Amazon.com Inc's (AMZN.O) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

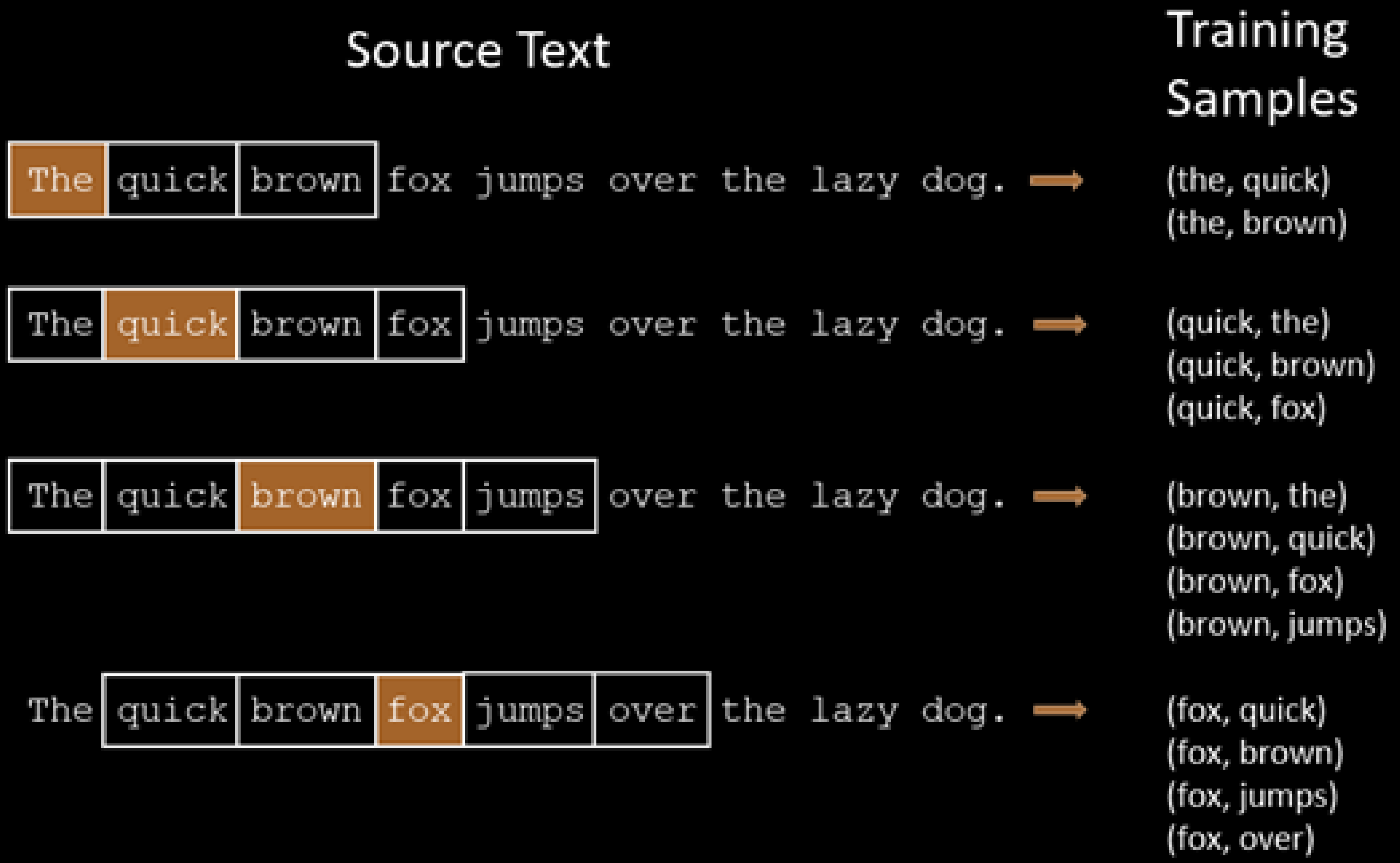
https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

Word Embedding



$$\text{sim}_{she} > \text{sim}_{he}$$

Creating Word Embedding



http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/

She	is	a	nurse	in	a	clinic

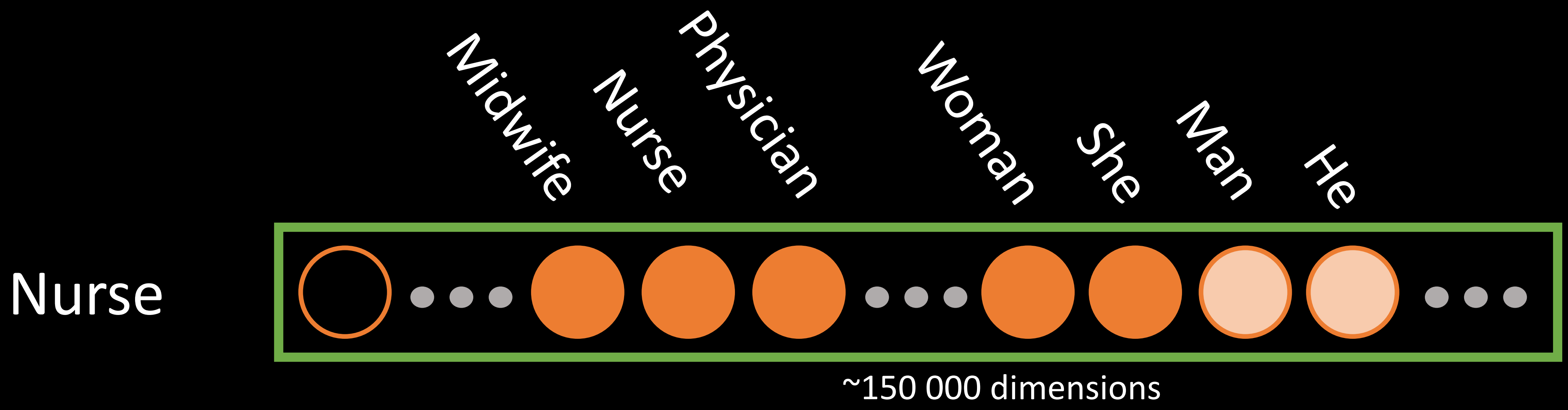
(nurse, she)

The vector for “nurse” contains an element of “she”-ness

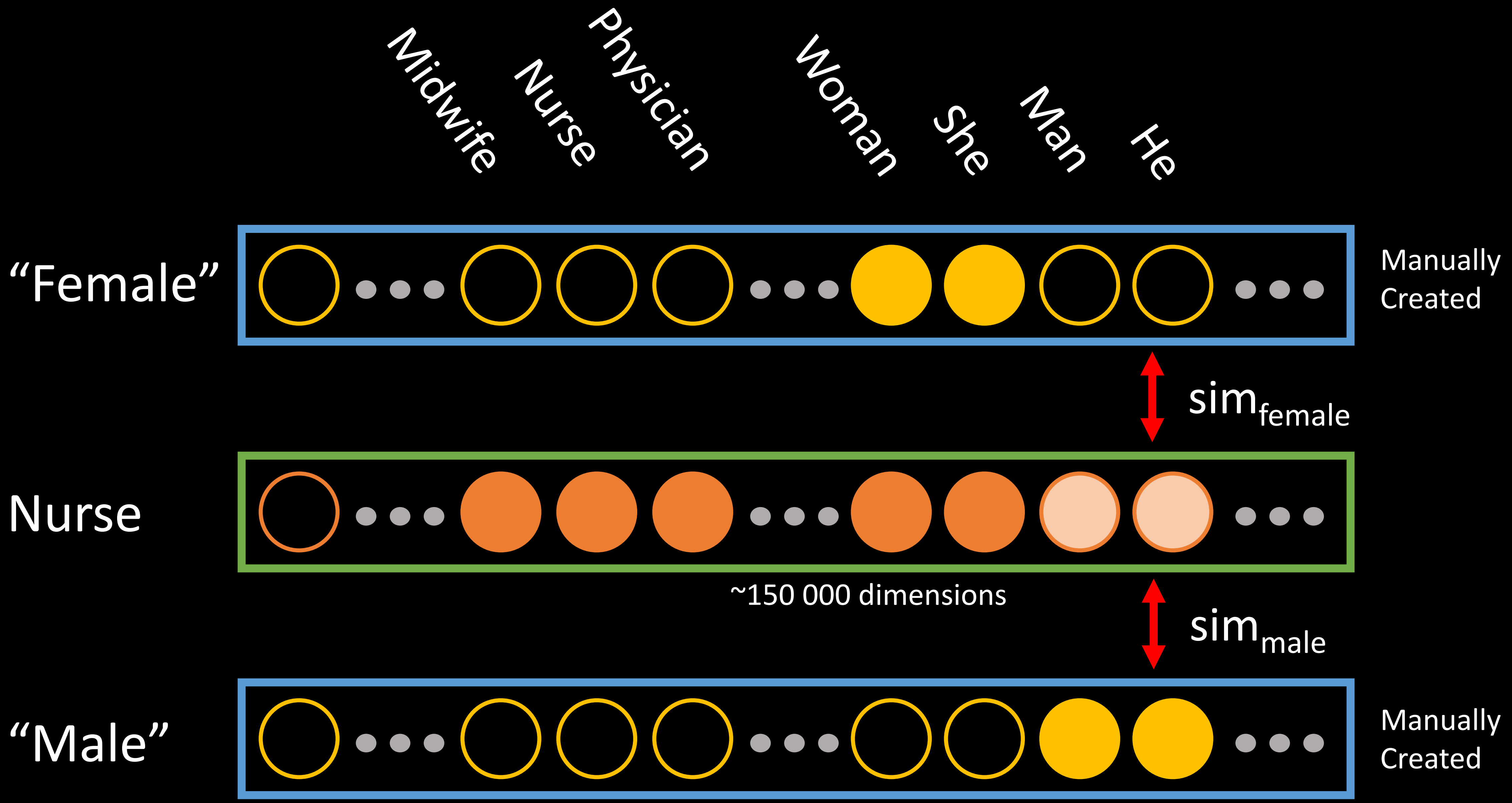
(she, nurse)

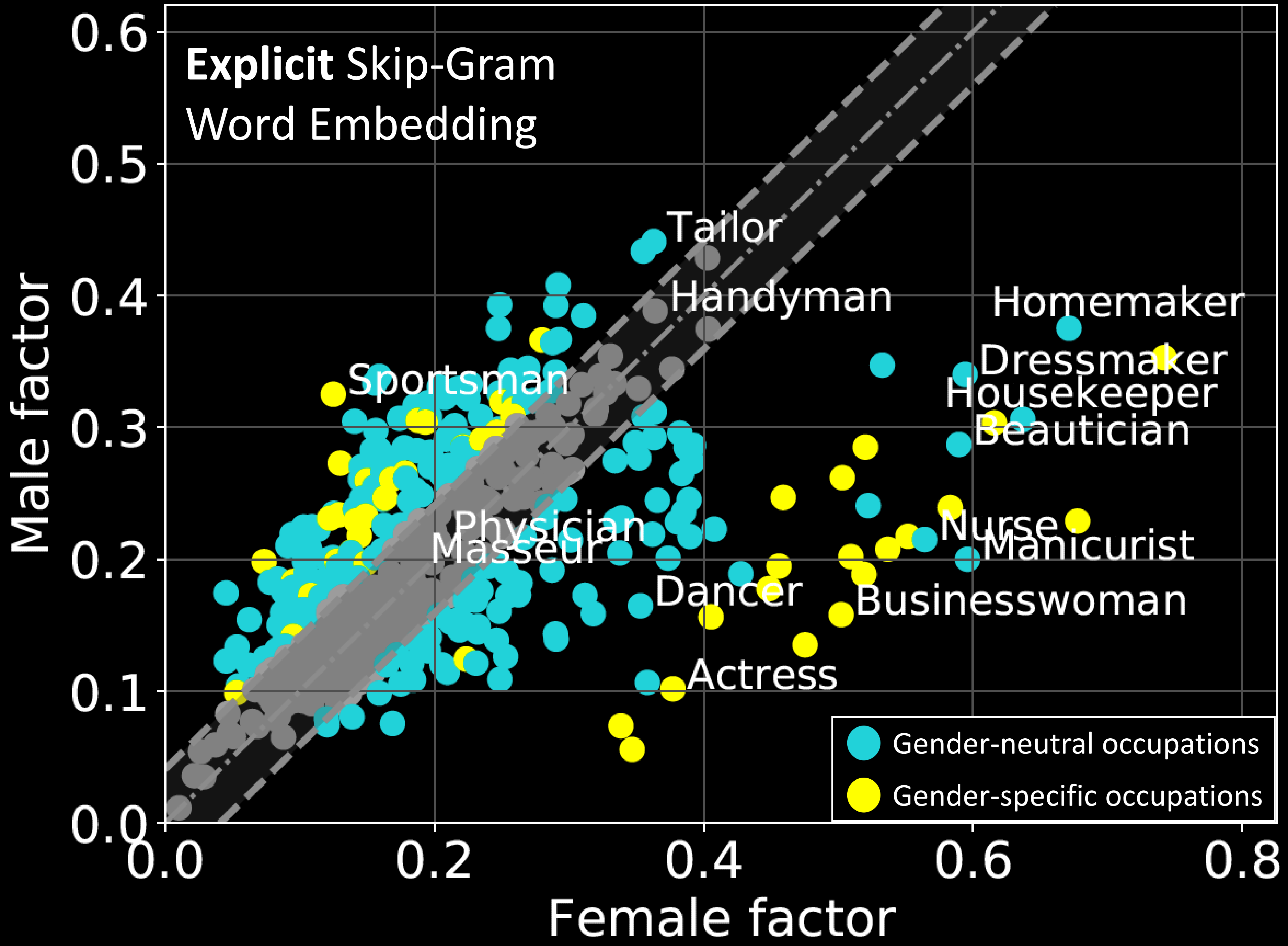
The vector for “she” contains an element of “nurse”-ness

Explicit Word Embedding Vectors



Explicit Word Embedding Vectors





Explicit Skip-Gram
Word Embedding
Male factor
Female factor
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.0
0.2
0.4
0.6
0.8
Tailor
Handyman
Homemaker
Dressmaker
Housekeeper
Beautician
Sportsman
Physician
Masseur
Nurse
Manicurist
Dancer
Businesswoman
Actress
Gender-neutral occupations
Gender-specific occupations

Correlation to Bias in the Real World

US Labour Statistics

1. Secretary	95%
2. Hairdresser	92%
3. Receptionist	90%
4. Nurse	90%
5. Housekeeper	89%
6. Cleaner	89%
7. Assistant	85%
8. Librarian	84%
⋮	

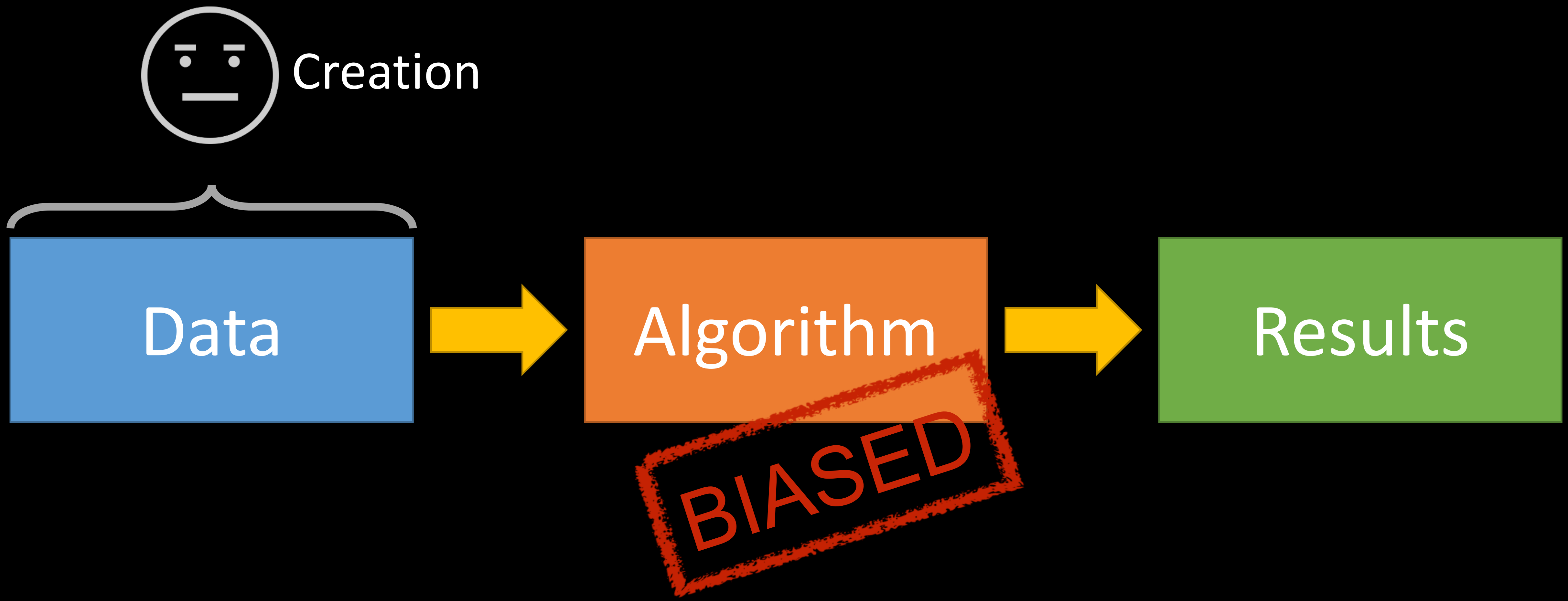
Spearman ρ correlation

Word Embedding Female Factor

1. Nanny	0.74
2. Midwife	0.68
3. Housekeeper	0.64
4. Manicurist	0.60
5. Dressmaker	0.59
6. Beautician	0.59
7. Maid	0.58
8. Nurse	0.56
⋮	

Spearman ρ correlation coefficient between Wikipedia embedding and US labour data

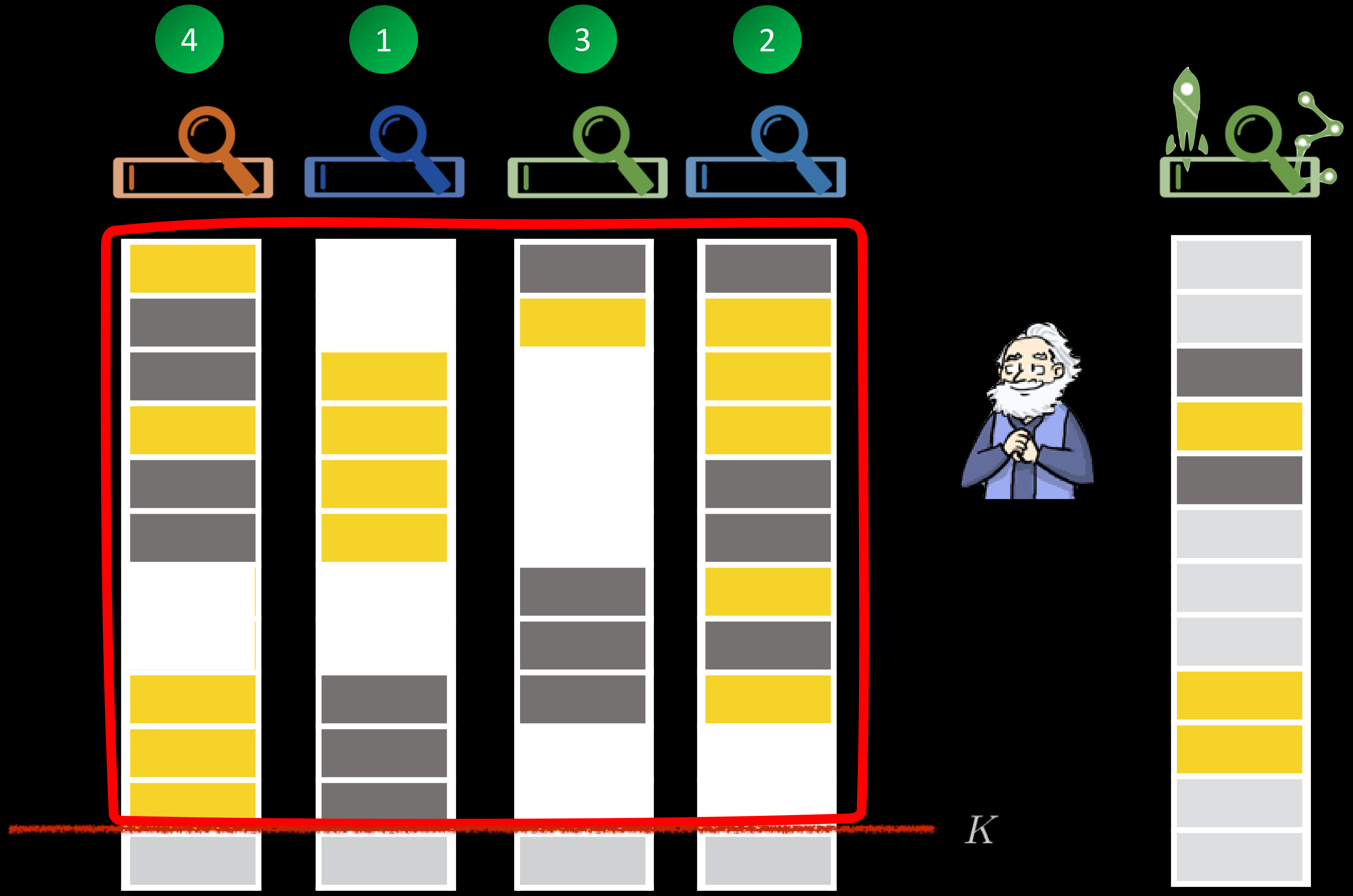
Standard Skip-Gram Word Embedding	0.53
Explicit Skip-Gram Word Embedding	0.64

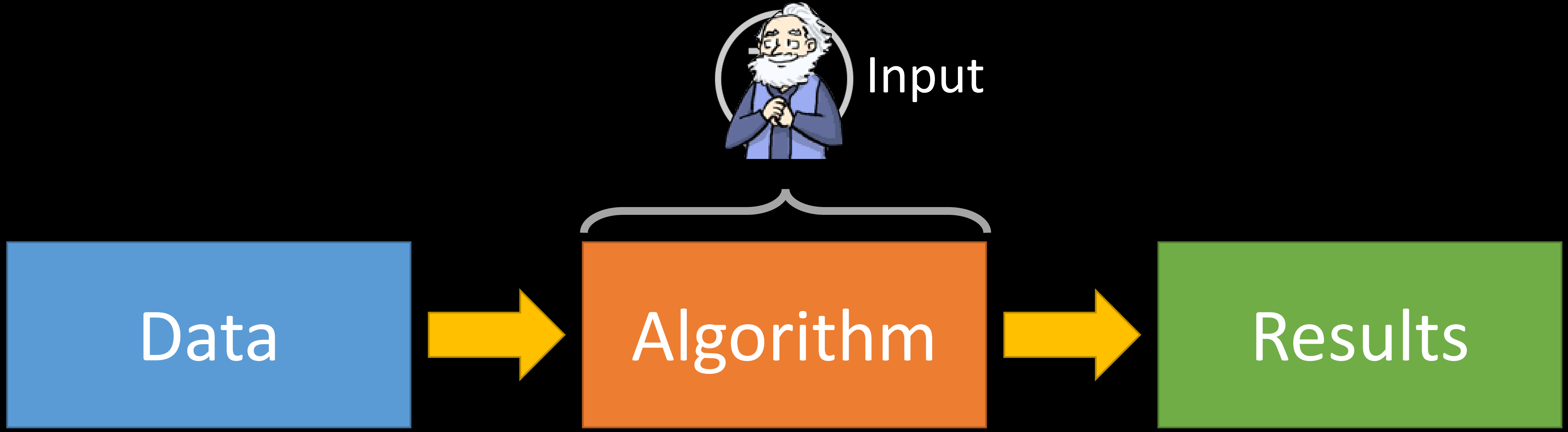

Creation
Data
Algorithm
BIASED
Results


Search


Search
An interesting query

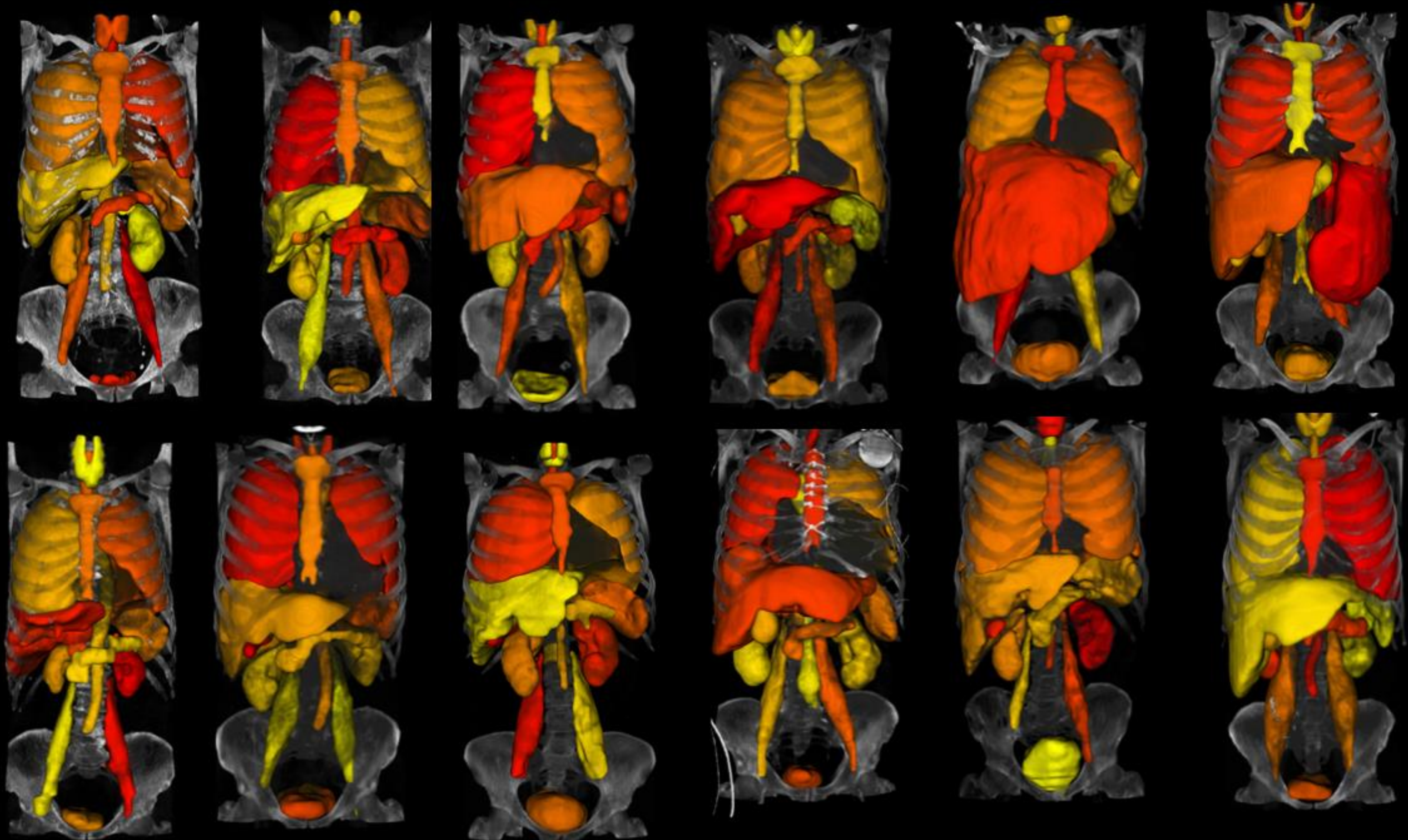
Pooling

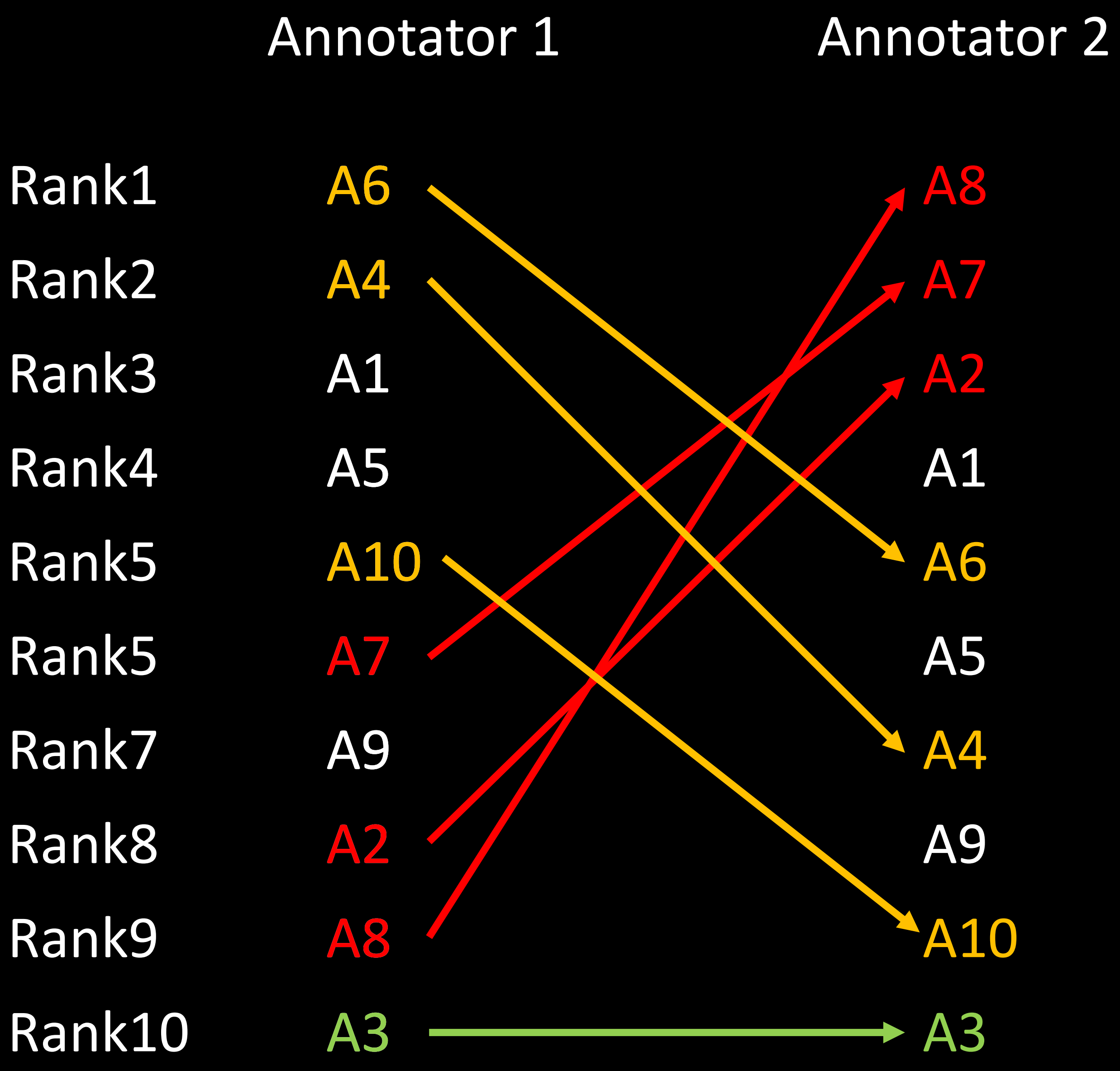



Input
Data
Algorithm
Results

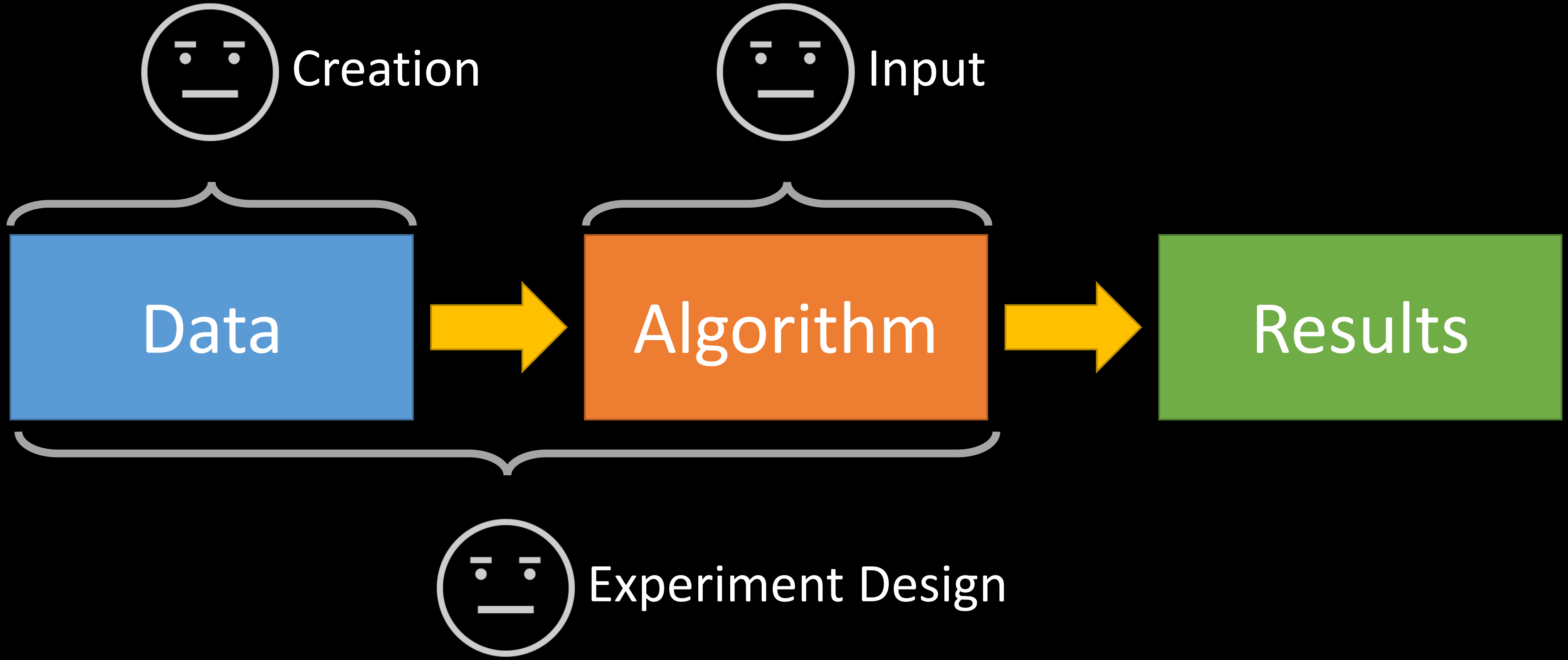
“Unpaid assessors largely disagree with paid assessors with respect to relevance labels [...] These differences have a noticeable impact on system ranking.”

Joao Palotti, Guido Zuccon, Johannes Bernhardt, Allan Hanbury, Lorraine Goeuriot, Assessors Agreement: A Case Study Across Assessor Type, Payment Levels, Query Variations and Relevance Dimensions, Proc. CLEF 2016, pp. 40-53



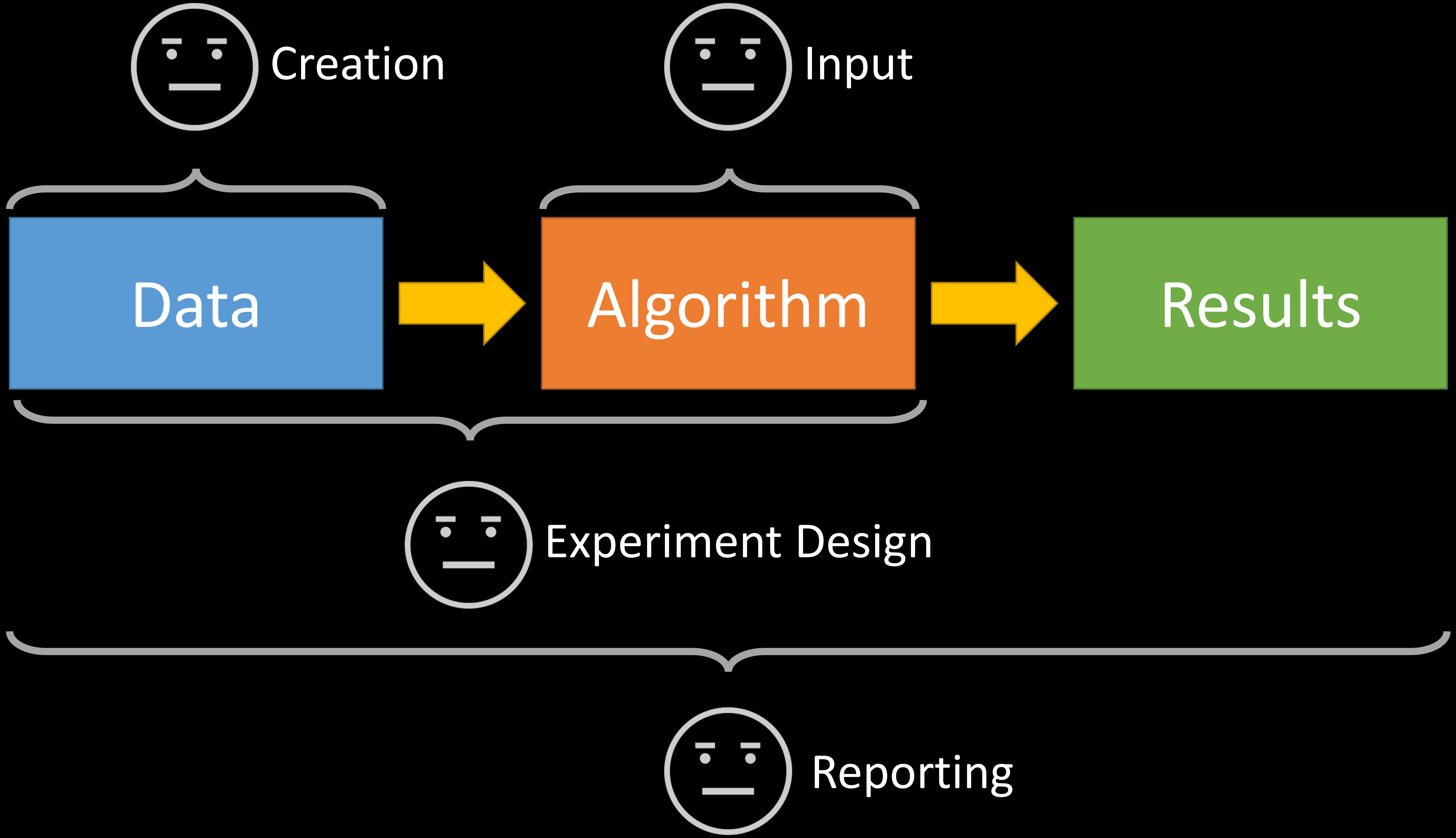


Maier-Hain et al., Why rankings of biomedical image analysis competitions should be interpreted with care, Nature Communications, volume 9, Article number: 5217 (2018)


Creation
Input
Data
Algorithm
Results
Experiment Design

Aggregation with mean	Aggregation with median
A1	A5
A2	A2
A3	A3
A4	A4
A5	A1
A6	A8
A7	A7
A8	A6
A9	A9
A10	A10
A11	A11
A12	A12
A13	A13

Maier-Hain et al., Why rankings of biomedical image analysis competitions should be interpreted with care, Nature Communications, volume 9, Article number: 5217 (2018)


Creation
Input
Data
Algorithm
Results
Experiment Design
Reporting

“Half of the relevant information is not reported”

Maier-Hain et al., Why rankings of biomedical image analysis competitions should be interpreted with care, Nature Communications, volume 9, Article number: 5217 (2018)

Parameter name	Coverage [%]
Challenge name[a]	100
Challenge website[a]	99
Organizing institutions and contact person[a]	97
Life cycle type[a]	100
Challenge venue or platform	99
Challenge schedule[a]	81
Ethical approval[a]	32
Data usage agreement	60
Interaction level policy[a]	62
Organizer participation policy[a]	6
Training data policy[a]	16
Pre-evaluation method	5

Maier-Hain et al., Why rankings of biomedical image analysis competitions should be interpreted with care, Nature Communications, volume 9, Article number: 5217 (2018)

"In 66% of all tasks, there was no description of how the reference annotation was performed."

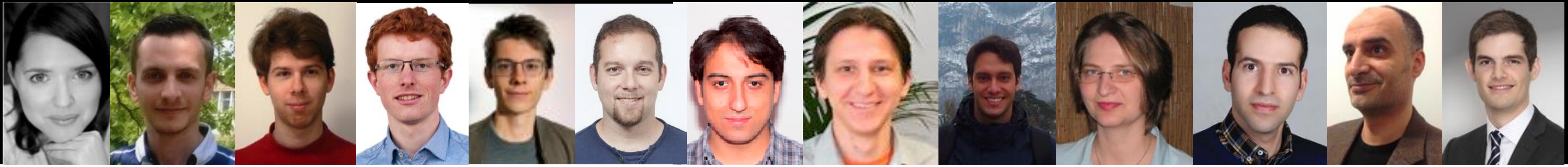
Maier-Hain et al., Why rankings of biomedical image analysis competitions should be interpreted with care, Nature Communications, volume 9, Article number: 5217 (2018)

Closing Questions

1. How can the reliability of published results be improved?
2. What is bias and who defines it?

Acknowledgements

Linda Andersson, Alexandros Bampoulidis, Tobias Fink, Georg Heiler, Sebastian Hofstätter, Florian Kromp, Aldo Lipani, Mihai Lupu, João Palotti, Florina Piroi, Navid Rekabsaz, Abdel Aziz Taha, Markus Zlabinger



FWF

FFG

ADMIRE

SELF OPTIMIZER

HISNI

KCONNECT

VISUAL HISTORY OF THE HOLOCAUST

TU WIEN
FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

COMPLEXITY SCIENCE HUB VIENNA

Univ.-Prof. Dr. Allan Hanbury

Institute for Information Systems Engineering
TU Wien
Favoritenstraße 9-11/194-04
1040 Vienna
Austria

Telephone: +43 1 58801 188310
Mobile: +43 676 978 0991

e-Mail: allan.hanbury@tuwien.ac.at