Who recruits better, human or machine?

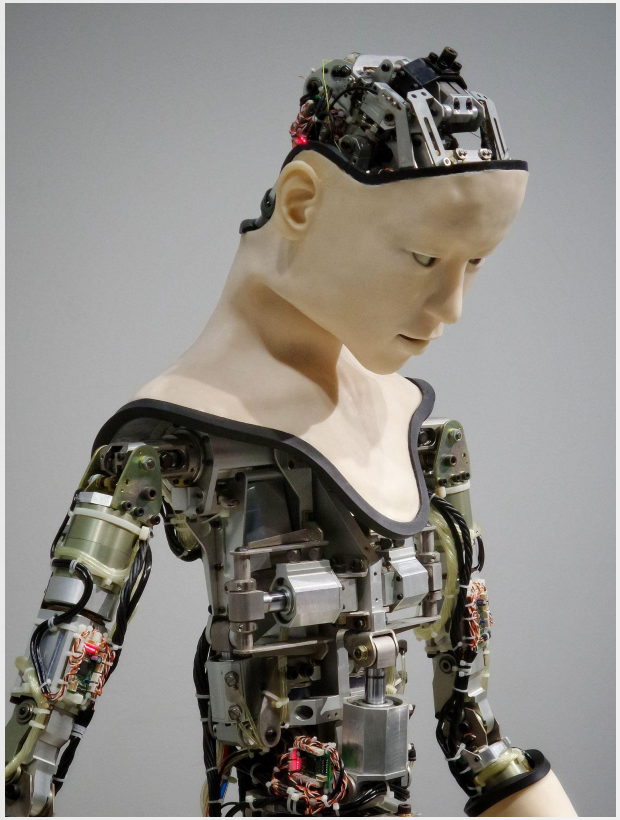
Taledo

Making recruitment an experience
The recruiting solution for every SME

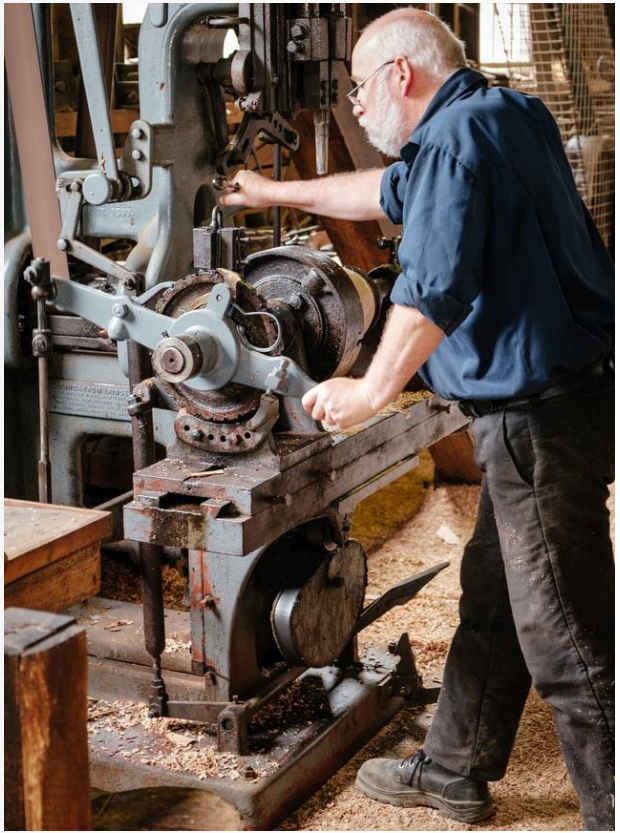
ZUKUNFT PERSONAL 2021 HR INNOVATION AWARD

EUROPÄISCHE UNION
Europäischer Fonds für Regionale Entwicklung

THE RISE OF AI / THE STATE IN RECRUITING



https://unsplash.com/photos/YKW0JjP7rlU



https://unsplash.com/photos/1qFmB3sZPSo

About Marcel

Taledo
Co-founder & CTO

payleven (now sumup)
Head of IT

Uni Lübeck
BSc & MSc Informatics

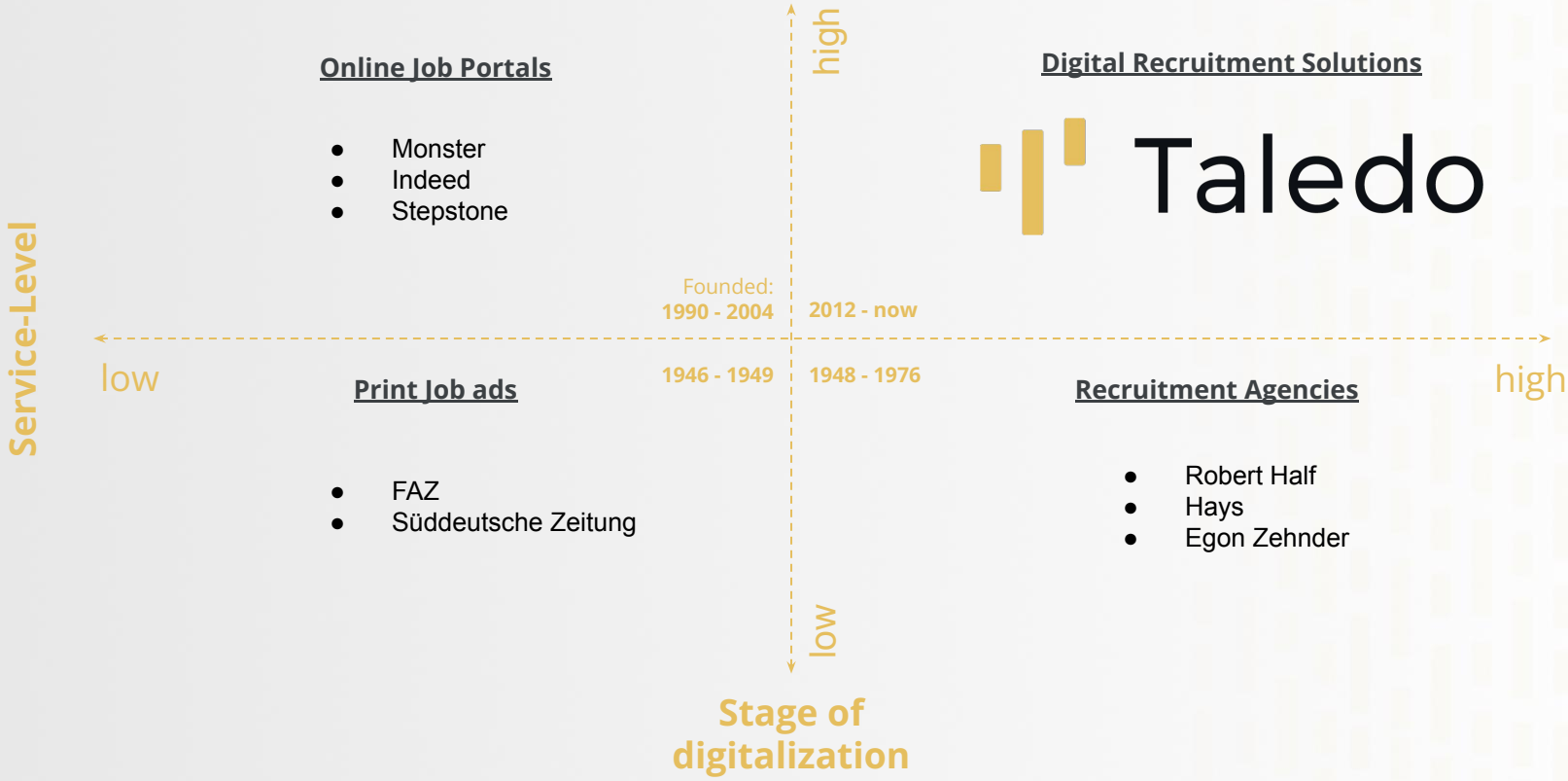
Marcel Poelker (personal picture)

Situation of HR departments

- Responsible for all phases of employees
- Chronically understaffed
- Cross-department requirements
- Time pressure
- Not trained for outreach / recruiting

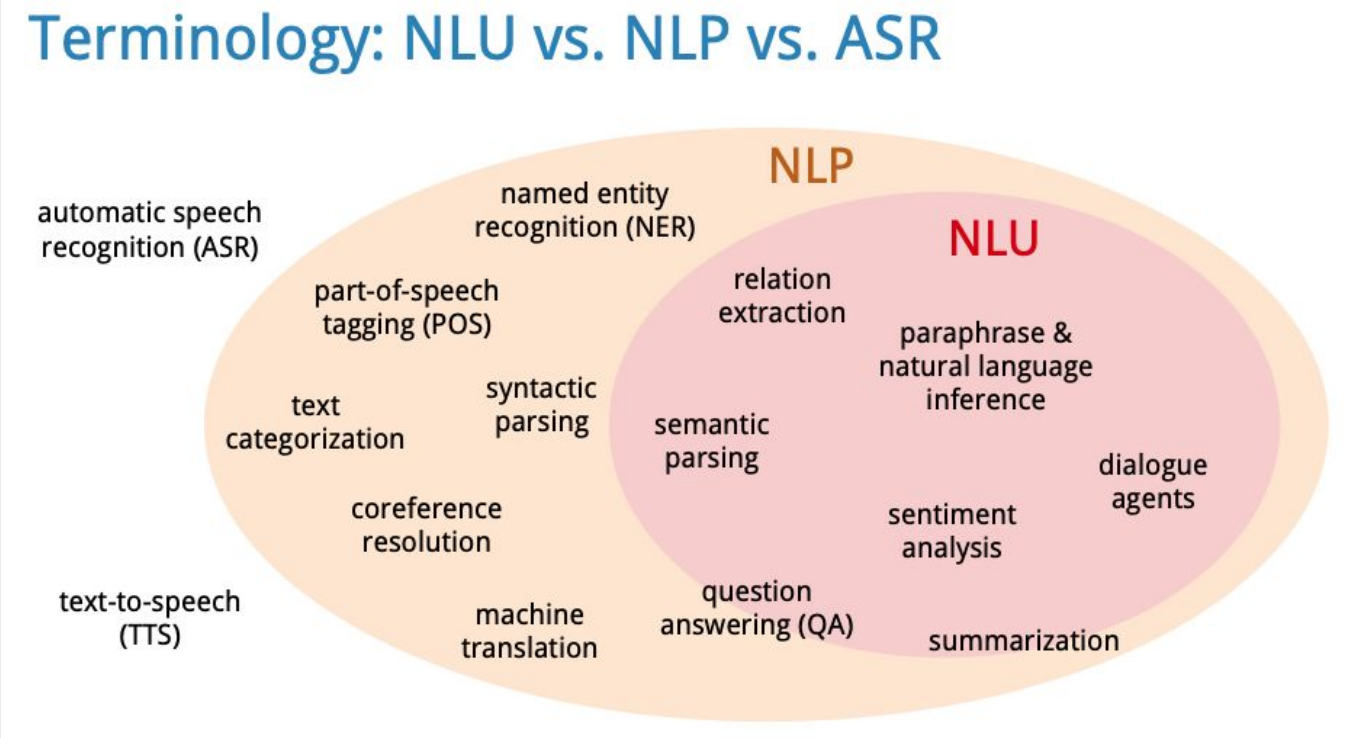
https://unsplash.com/photos/bmJAXAz6ads

DIGITALIZATION IN RECRUITING



MATCHING

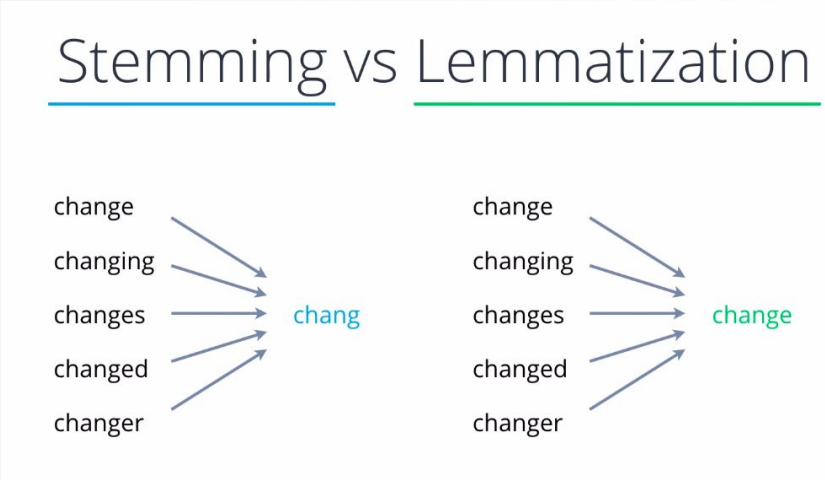
NLP: NATURAL LANGUAGE PROCESSING



https://nlp.stanford.edu/~wcmac/papers/20140716-UNLU.pdf

SYNTACTIC SEARCH: MATCHING SIMILAR WORDS

- Simple approach, but too strict
- Smoothing / Fuzzy search:
  - Normalization (Stemming / Lemmatization)
  - Typos (Levenshtein)
  - Phonetic similarity (Metaphone)
  - Synonyms
  - Boolean search



https://github.com/Learn-Write-Repeat/Open-contributions/blob/master/B2-NLP/Amey_Nlp_Lemmatization_stemming.md

LIMITS OF SYNTACTIC SEARCH

https://www.wildcodeschool.com/en-GB/blog/difference-between-java-and-javascript

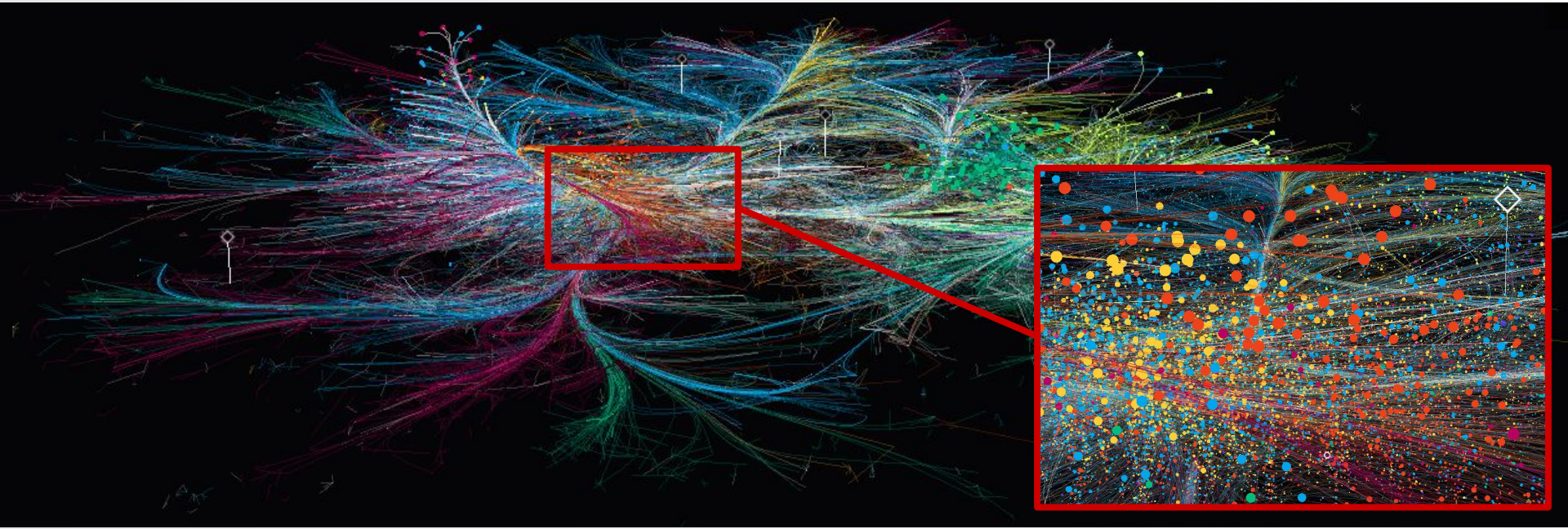
SEMANTIC SEARCH: UNDERSTANDING THE SEARCH

- Ontologies / Knowledge graphs (relations between concepts)

150 years of Nature papers
https://www.nature.com/immersive/d41586-019-03165-4/index.html

SEMANTIC SEARCH: UNDERSTANDING THE SEARCH

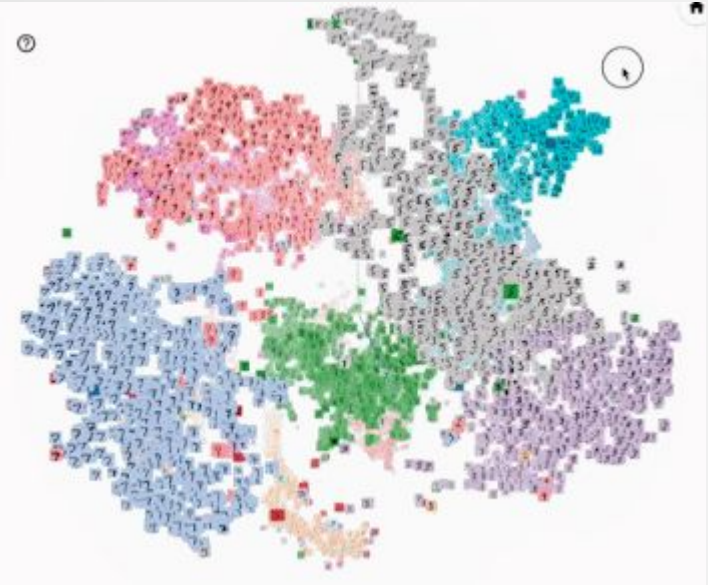
- Ontologies / Knowledge graphs (relations between concepts)



150 years of Nature papers
https://www.nature.com/immersive/d41586-019-03165-4/index.html

SEMANTIC SEARCH: UNDERSTANDING THE SEARCH

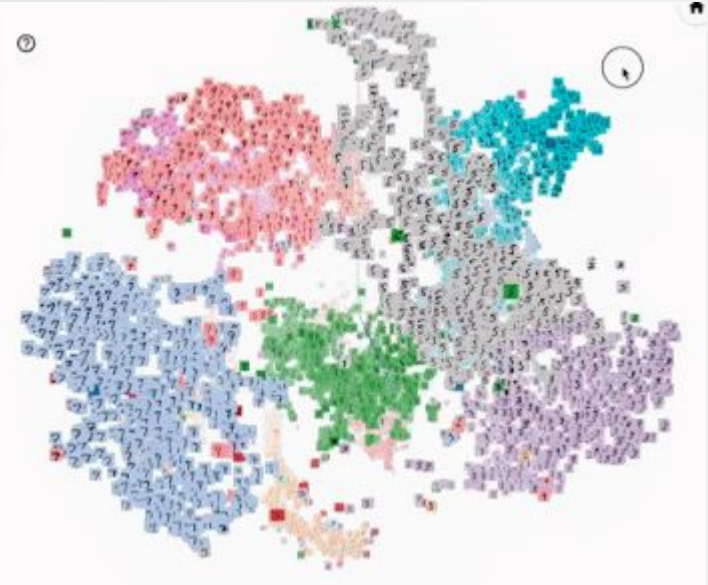
- Embeddings (map words into vector space)



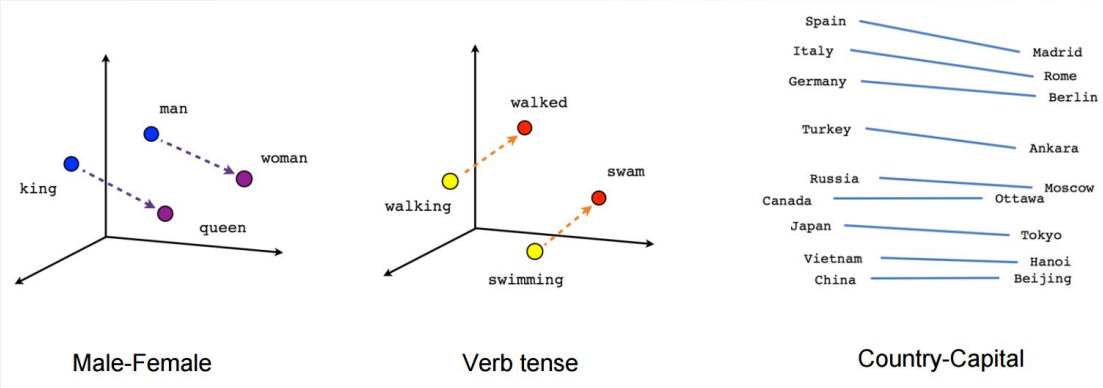
https://ai.googleblog.com/2016/12/open-sourcing-embedding-projector-tool.html

SEMANTIC SEARCH: UNDERSTANDING THE SEARCH

- Embeddings (map words into vector space)



https://ai.googleblog.com/2016/12/open-sourcing-embedding-projector-tool.html

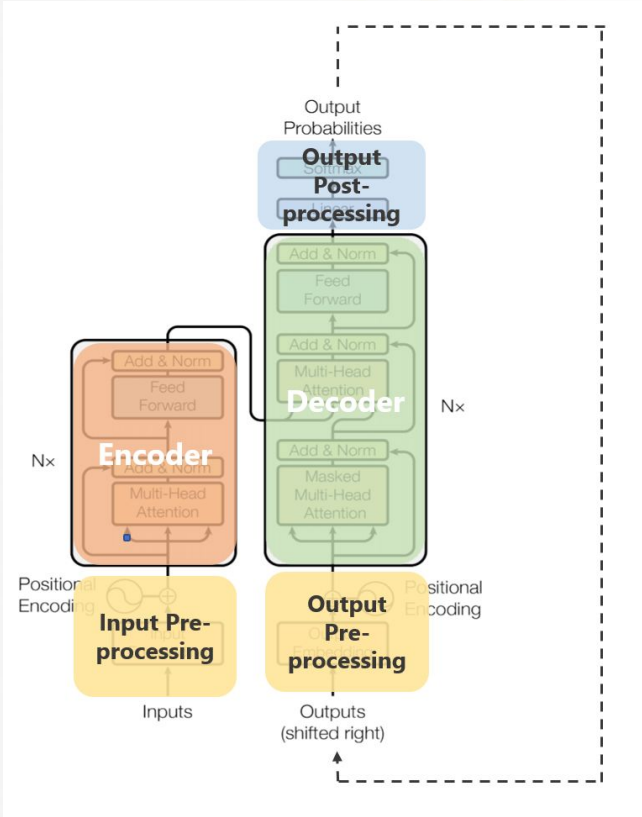


https://towardsdatascience.com/deep-learning-4-embedding-layers-f9a02d55ac12

TRANSFORMER MODELS

- Good DL architecture for NLP tasks
  - Designed to process sequential data
  - Works in parallel
  - Able to find out the relevant parts of a text

- Famous models:
  - BERT
  - GPT-3

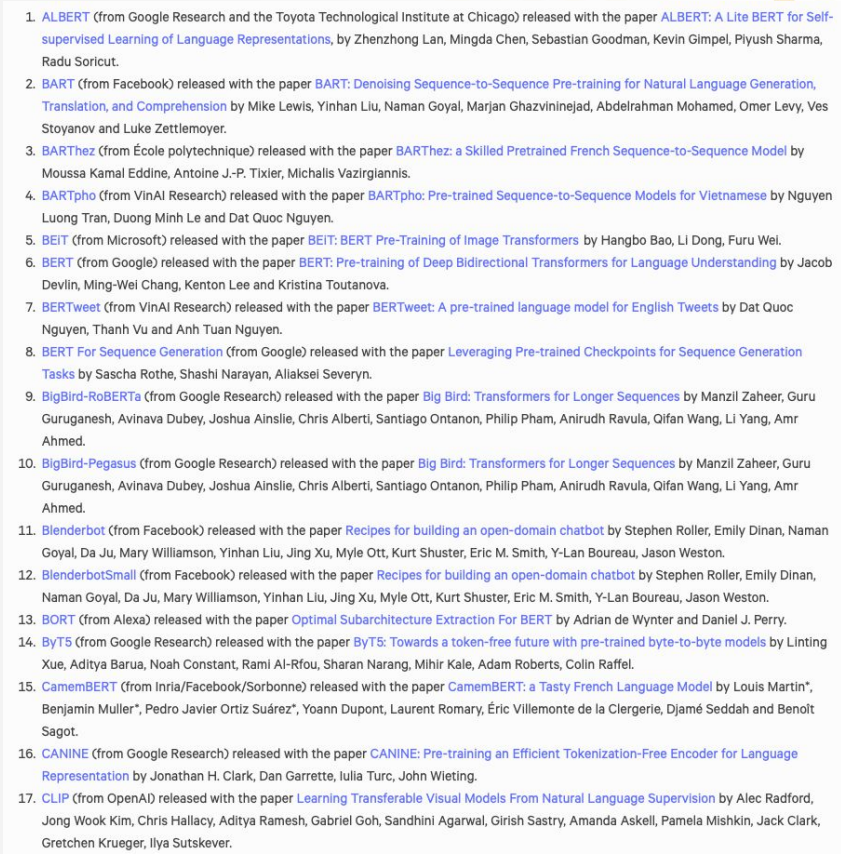


https://towardsdatascience.com/transformers-89034557de14

TRANSFORMER MODELS

- Many models available (huggingface)
  - Listed: 82 * variants
  - Search page: ~20k models

- If you build upon that, know your criteria:
  - What's the task solved by the model?
  - Multi-lingual?
  - How context-aware?
  - How easy to fine-tune / re-train?
  - Production-ready?



1. ALBERT (from Google Research and the Toyota Technological Institute at Chicago) released with the paper ALBERT: A Lite BERT for Self-supervised Learning of Language Representations, by Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, Radu Soricut.
2. BART (from Facebook) released with the paper BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension by Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov and Luke Zettlemoyer.
3. BARThez (from École polytechnique) released with the paper BARThez: a Skilled Pretrained French Sequence-to-Sequence Model by Moussa Kamal Eddine, Antoine J.-P. Tixier, Michalis Vazirgiannis.
4. BARTpho (from VinAI Research) released with the paper BARTpho: Pre-trained Sequence-to-Sequence Models for Vietnamese by Nguyen Luong Tran, Duong Minh Le and Dat Quoc Nguyen.
5. BEiT (from Microsoft) released with the paper BEiT: BERT Pre-Training of Image Transformers by Hangbo Bao, Li Dong, Furu Wei.
6. BERT (from Google) released with the paper BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding by Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova.
7. BERTweet (from VinAI Research) released with the paper BERTweet: A pre-trained language model for English Tweets by Dat Quoc Nguyen, Thanh Vu and Anh Tuan Nguyen.
8. BERT For Sequence Generation (from Google) released with the paper Leveraging Pre-trained Checkpoints for Sequence Generation Tasks by Sascha Rothe, Shashi Narayan, Aliaksei Severyn.
9. BigBird-RoBERTa (from Google Research) released with the paper Big Bird: Transformers for Longer Sequences by Manzil Zaheer, Guru Guruganesh, Avinava Dubey, Joshua Ainslie, Chris Alberti, Santiago Ontanon, Philip Pham, Anirudh Ravula, Qifan Wang, Li Yang, Amr Ahmed.
10. BigBird-Pegasus (from Google Research) released with the paper Big Bird: Transformers for Longer Sequences by Manzil Zaheer, Guru Guruganesh, Avinava Dubey, Joshua Ainslie, Chris Alberti, Santiago Ontanon, Philip Pham, Anirudh Ravula, Qifan Wang, Li Yang, Amr Ahmed.
11. Blenderbot (from Facebook) released with the paper Recipes for building an open-domain chatbot by Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Eric M. Smith, Y-Lan Boureau, Jason Weston.
12. BlenderbotSmall (from Facebook) released with the paper Recipes for building an open-domain chatbot by Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Eric M. Smith, Y-Lan Boureau, Jason Weston.
13. BORT (from Alexa) released with the paper Optimal Subarchitecture Extraction For BERT by Adrian de Wynter and Daniel J. Perry.
14. ByT5 (from Google Research) released with the paper ByT5: Towards a token-free future with pre-trained byte-to-byte models by Linting Xue, Aditya Barua, Noah Constant, Rami Al-Rfou, Sharan Narang, Mihir Kale, Adam Roberts, Colin Raffel.
15. CamemBERT (from Inria/Facebook/Sorbonne) released with the paper CamemBERT: a Tasty French Language Model by Louis Martin*, Benjamin Muller*, Pedro Javier Ortiz Suárez*, Yoann Dupont, Laurent Romary, Éric Villemonte de la Clergerie, Djamé Seddah and Benoît Sagot.
16. CANINE (from Google Research) released with the paper CANINE: Pre-training an Efficient Tokenization-Free Encoder for Language Representation by Jonathan H. Clark, Dan Garrette, Iulia Turc, John Wieting.
17. CLIP (from OpenAI) released with the paper Learning Transferable Visual Models From Natural Language Supervision by Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, Ilya Sutskever.

https://huggingface.co/transformers/

FURTHER DESIGN DECISIONS

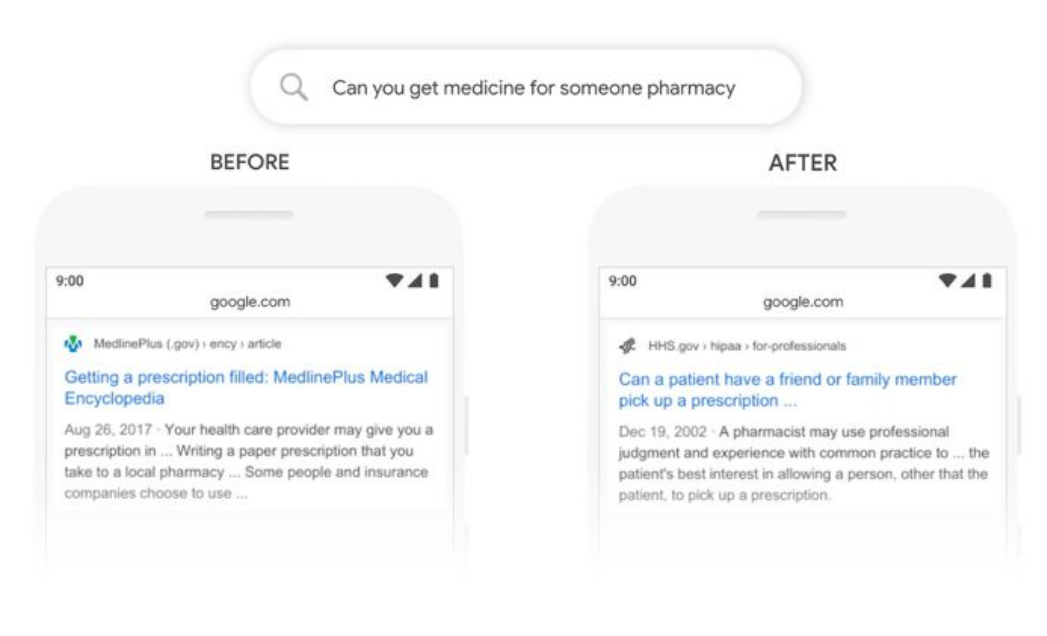
- What to search for (search space)? What to embed?
- How to rank? Is a visual score required?
- What to train for (target variable)? How to measure success?
- What model configuration works best?
- How to validate and compare results?
- ...

General advice:

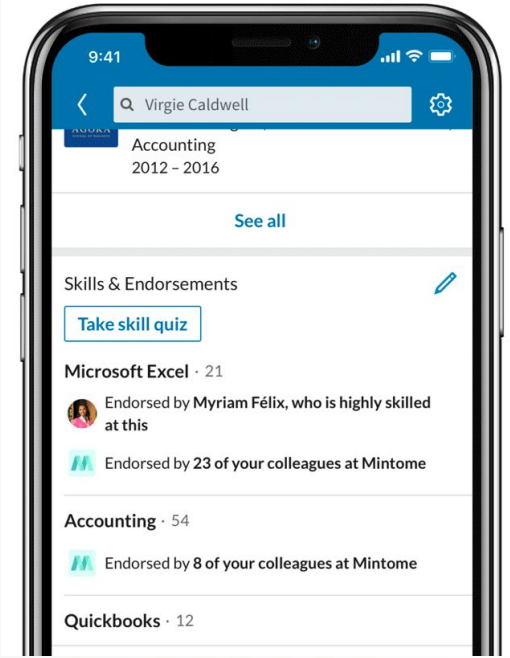
- Understand the problem & the domain
- Have a baseline
- Improve through exploration and iteration

SEARCHING/MATCHING REMAINS DIFFICULT PROBLEM

Big companies are actively iterating, e.g. Google and LinkedIn



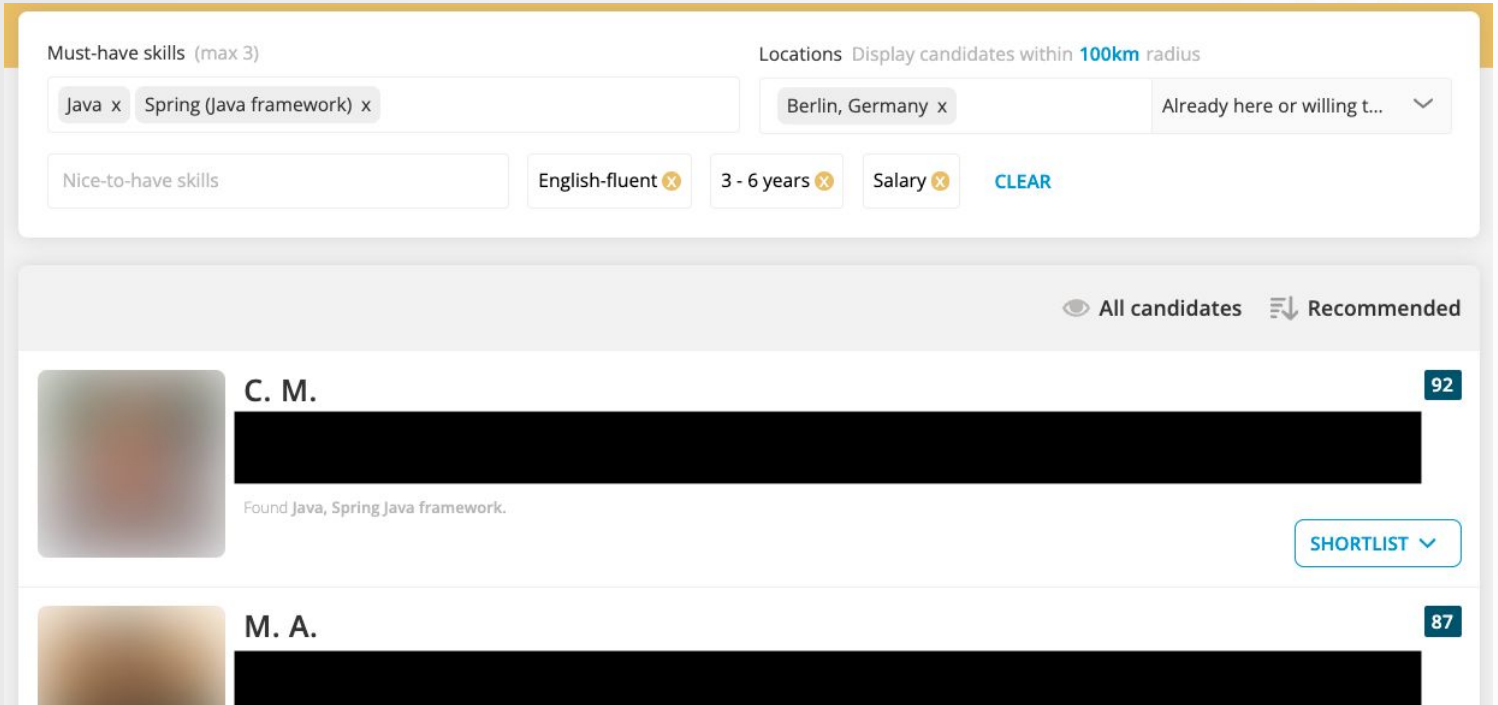
https://blog.google/products/search/search-language-understanding-bert/



https://blog.linkedin.com/2019/september/17/announcing-skill-assessments-to-help-you-showcase-your-skills

TALEDO'S MATCHING

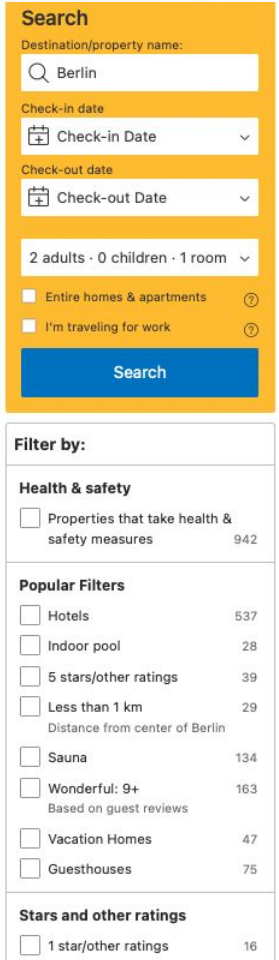
- Semantic Search
- AI for NLP + Search Algorithm + Scoring
- Skill based approach + behavior on the Taledo software



BEYOND AI

PRODUCT DESIGN MATTERS: SEARCH INPUTS

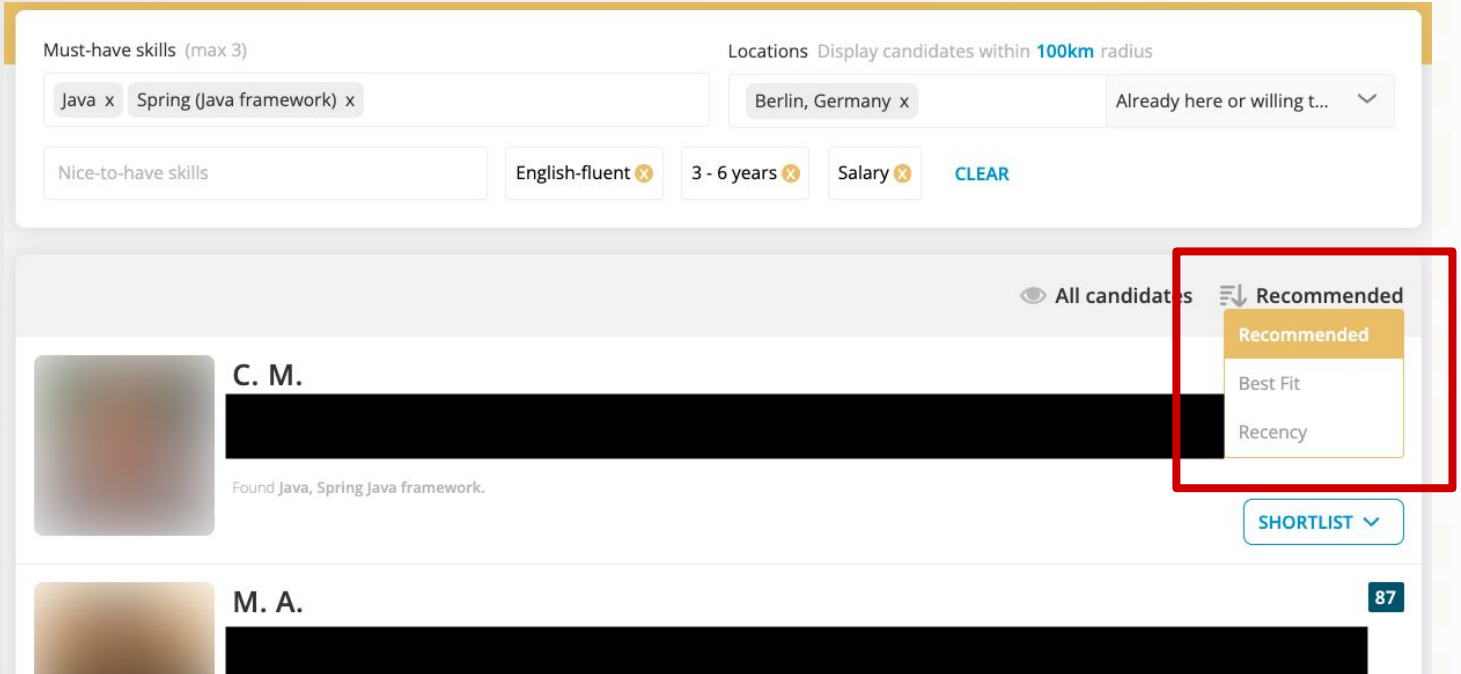
google.com



booking.com

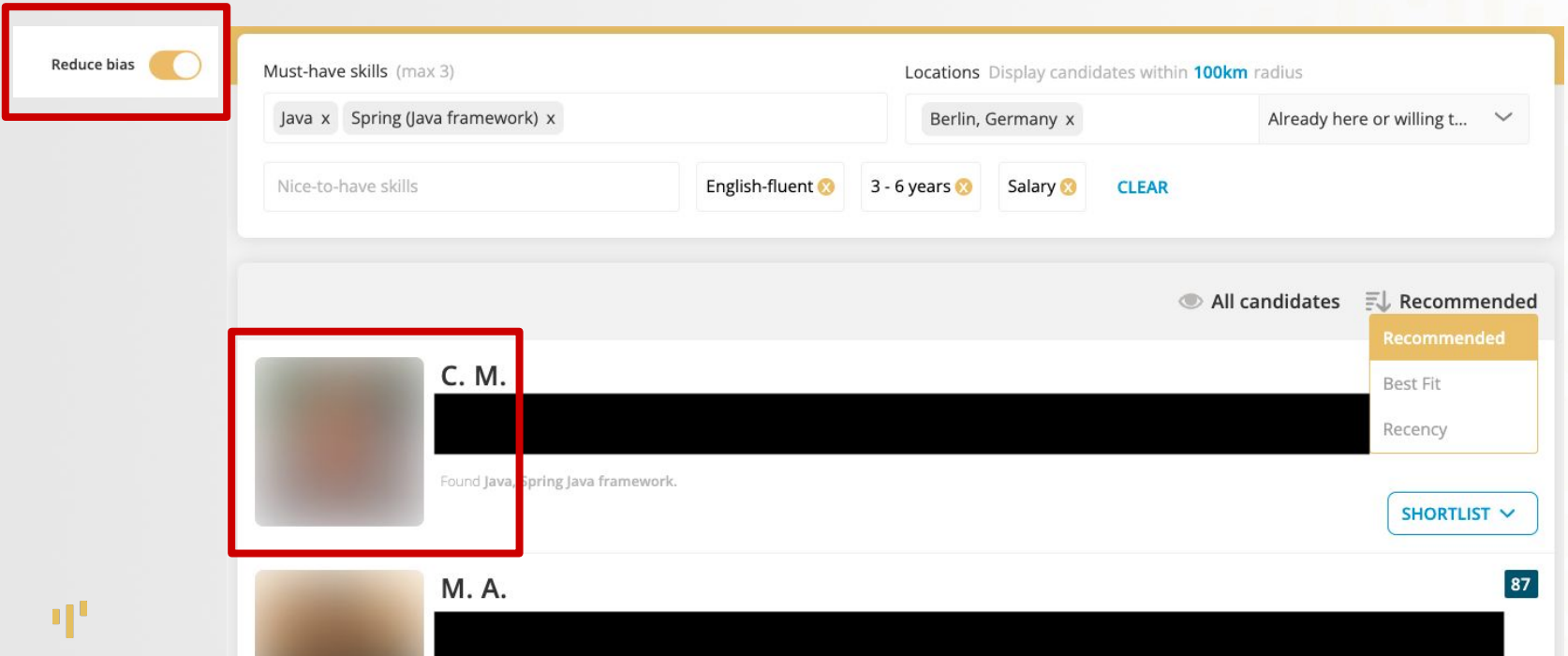
TALEDO: PRODUCT AND AI COOPERATE

- Best match vs availability?
  - Group 1: Best match only
  - Group 2: Availability first (if qualifies)
  - Recommendation (“Taledo Score”)



TALEDO: PRODUCT AND AI COOPERATE

- Avoid bias
  - Skill based approach for AI
  - “Reduce Bias” feature to adjust presentation



TALEDO: THE NATURE OF RECRUITING IS HUMAN

- High service level
- Important decision
- Trust

People business

https://unsplash.com/photos/376KN_ISpIE

WHO RECRUITS BETTER?

Taledo's answer is:

The combination does.

PRODUCT DESIGN

+

AI

+

HUMAN TOUCH

THANK YOU!

Marcel Poelker
marcel@taledo.com