

# Who recruits better, human or machine?



**Making recruitment an experience**  
The recruiting solution for every SME



AI MONDAY BERLIN



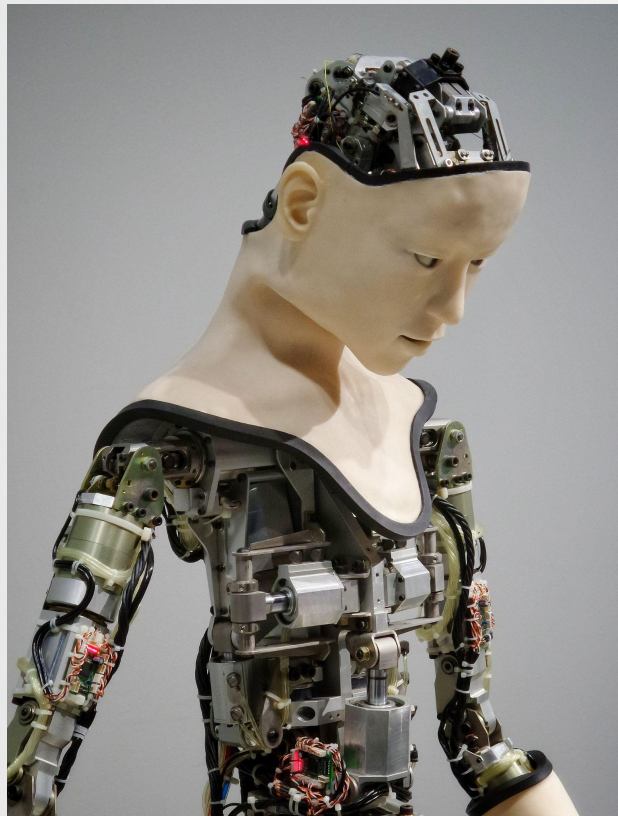
WHEN:  
NOV 8 - 2021

AI & TALENT  
MANAGEMENT

#AIMonday  
[ai-monday.de/berlin](https://ai-monday.de/berlin)



## THE RISE OF AI / THE STATE IN RECRUITING



<https://unsplash.com/photos/YKW0JJP7rIU>



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



# About Marcel

## **Taledo**

Co-founder & CTO

## **payleven (now sumup)**

Head of IT

## **Uni Lübeck**

BSc & MSc Informatics



## Situation of HR departments

---

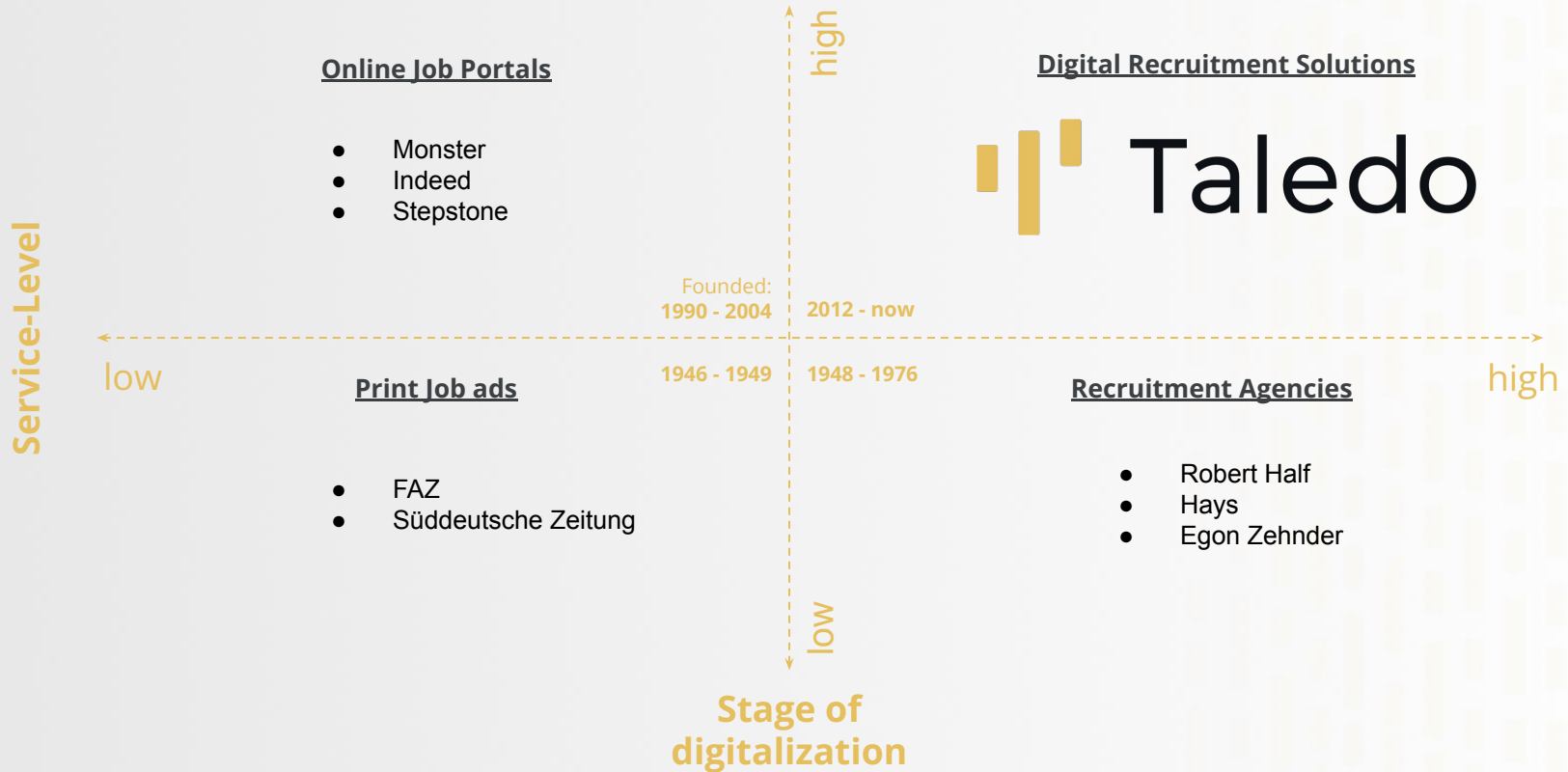
- Responsible for all phases of employees
- Chronically understaffed
- Cross-department requirements
- Time pressure
- Not trained for outreach / recruiting



<https://unsplash.com/photos/bmJAXAz6ads>



# DIGITALIZATION IN RECRUITING

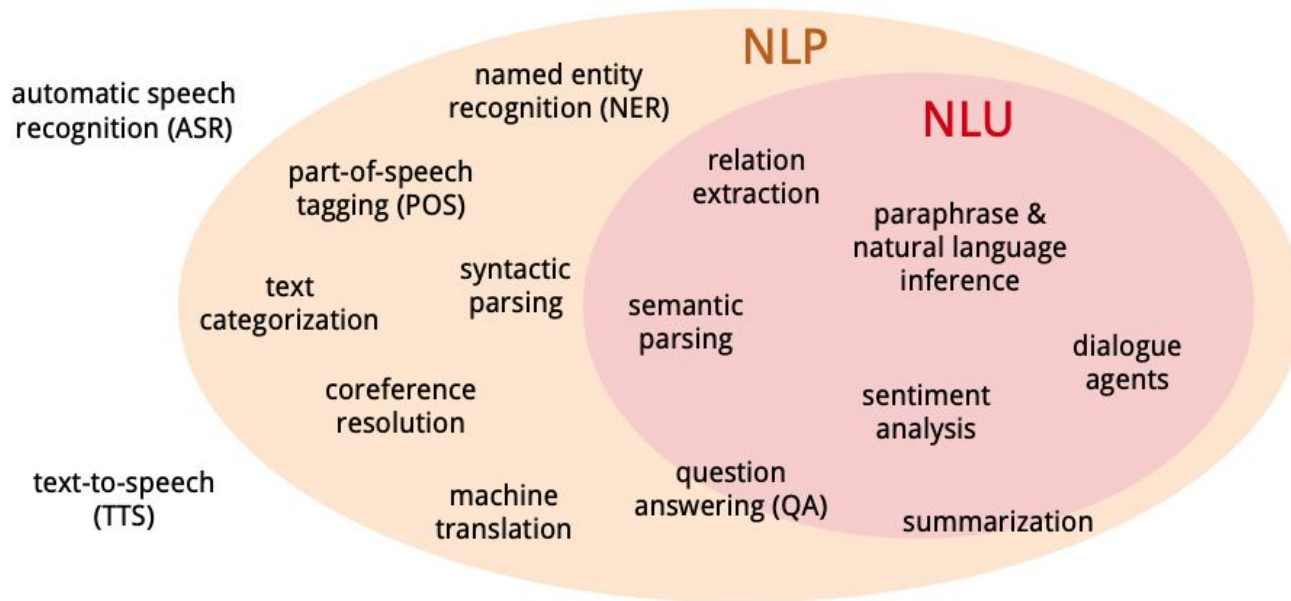


# MATCHING





## Terminology: NLU vs. NLP vs. ASR

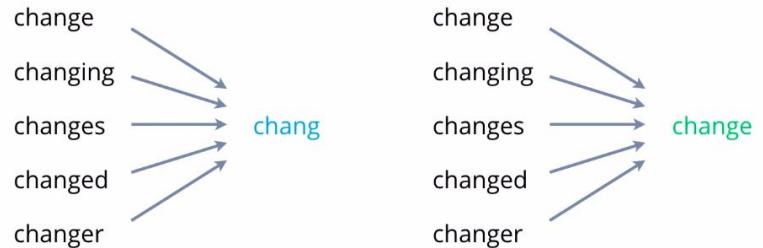


<https://nlp.stanford.edu/~wcmac/papers/20140716-UNLU.pdf>

## SYNTACTIC SEARCH: MATCHING SIMILAR WORDS

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

### Stemming vs Lemmatization



[https://github.com/Learn-Write-Repeat/Open-contributions/blob/master/B2-NLP/Amey\\_Nlp\\_Lemmatization\\_stemming.md](https://github.com/Learn-Write-Repeat/Open-contributions/blob/master/B2-NLP/Amey_Nlp_Lemmatization_stemming.md)





## LIMITS OF SYNTACTIC SEARCH

JAVA *is to*  
JAVASCRIPT

*as*

HAM *is to*  
HAMSTER



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



## SEMANTIC SEARCH: UNDERSTANDING THE SEARCH

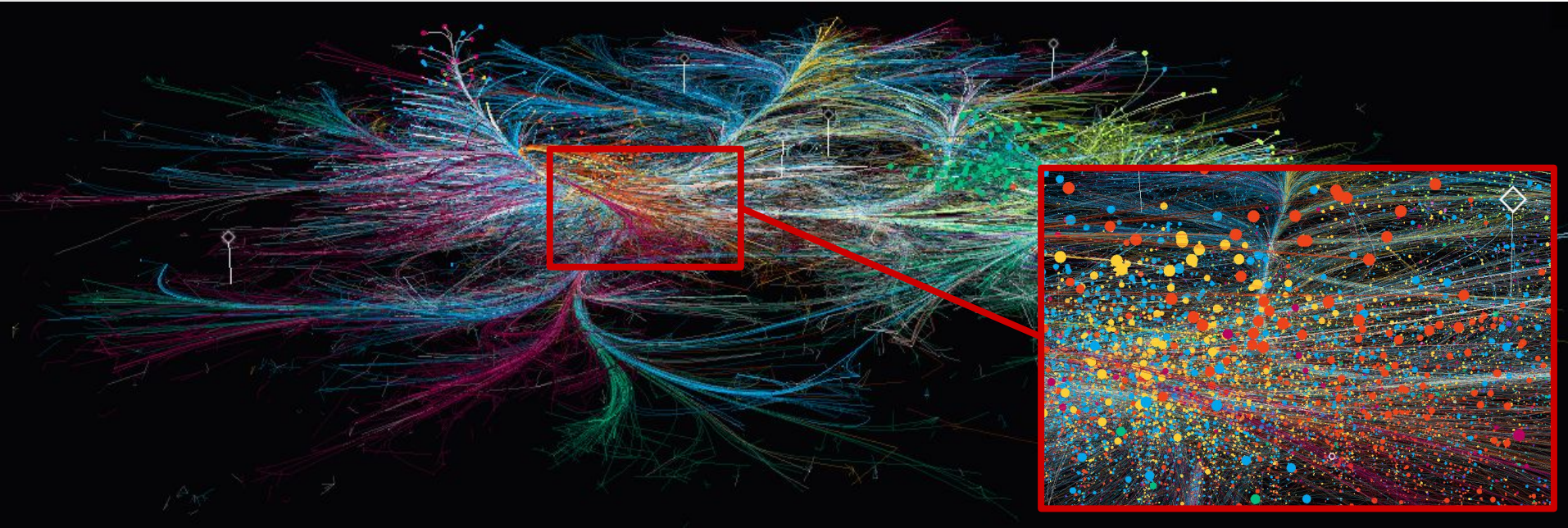
- Ontologies / Knowledge graphs (relations between concepts)



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)

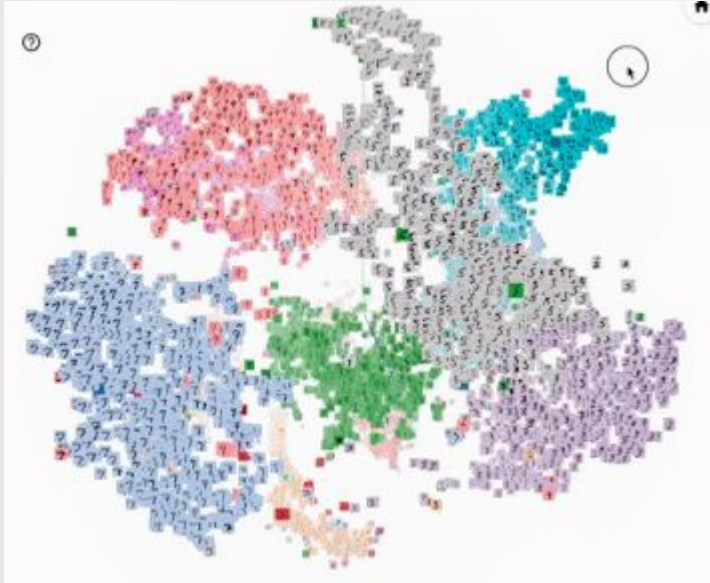


150 years of Nature papers  
<https://www.nature.com/immersive/d41586-019-03165-4/index.html>



## SEMANTIC SEARCH: UNDERSTANDING THE SEARCH

- Embeddings (map words into vector space)

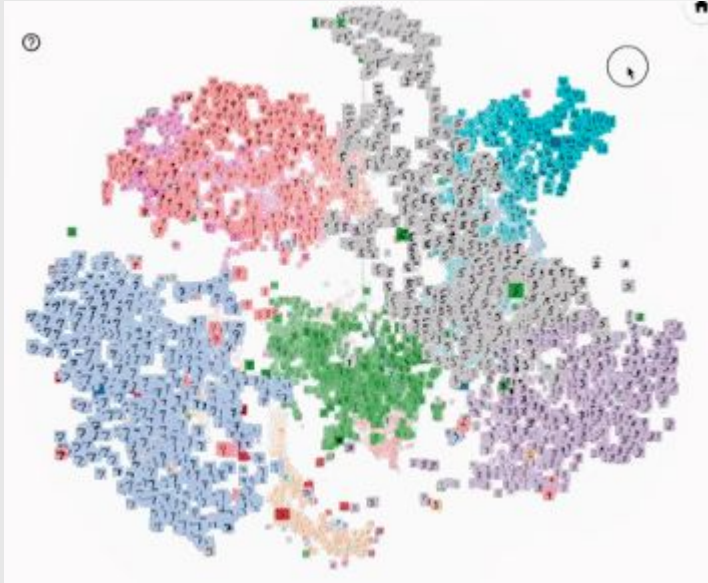


<https://ai.googleblog.com/2016/12/open-sourcing-embedding-projector-tool.html>

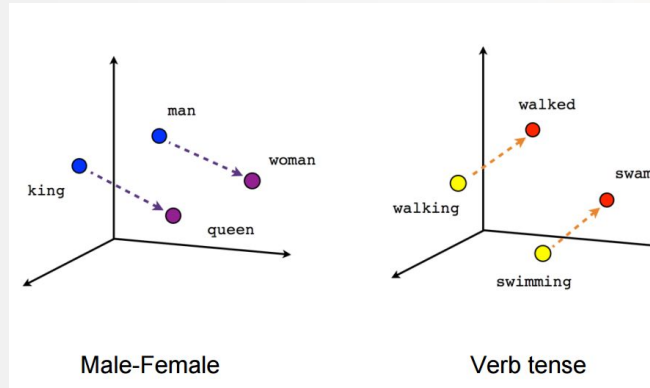


# SEMANTIC SEARCH: UNDERSTANDING THE SEARCH

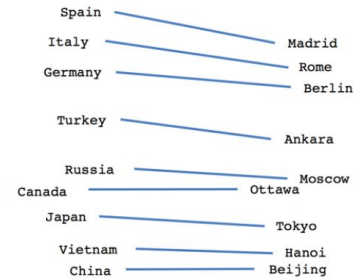
- Embeddings (map words into vector space)



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



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

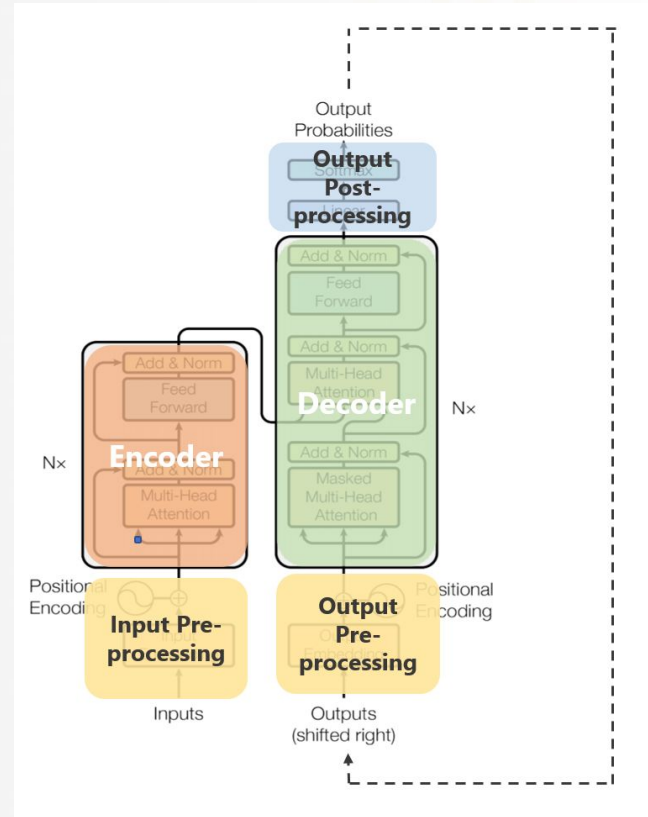


Country-Capital



# 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, Allaksel 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, Djamel Seddah and Benoit 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?
- ...

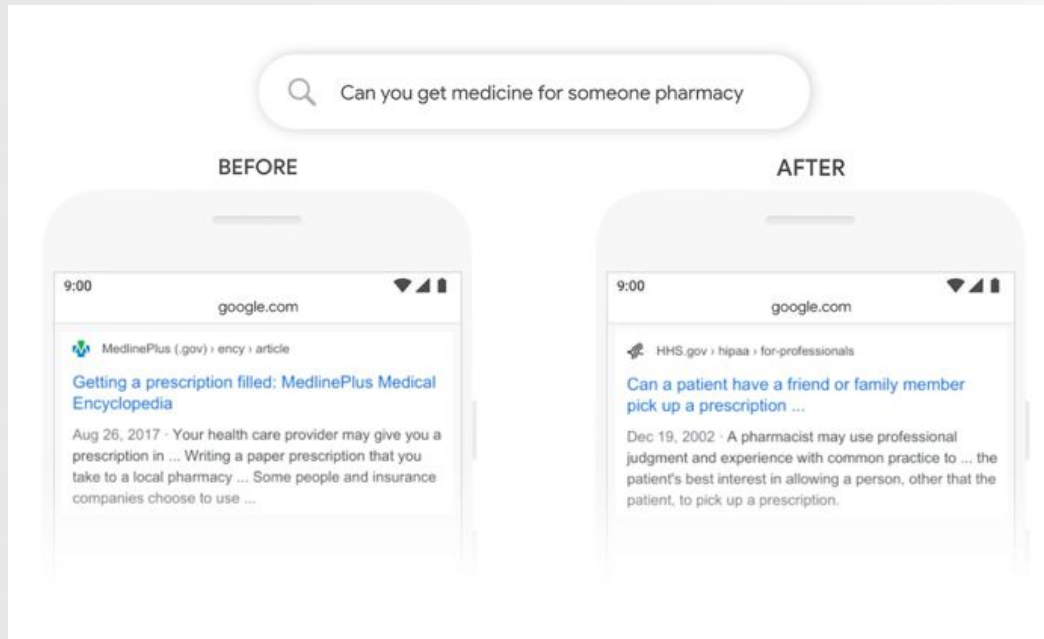
General advice:

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

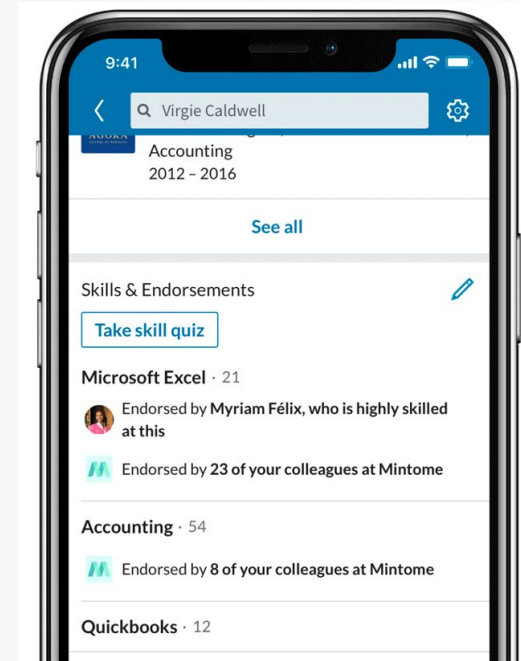


## SEARCHING/MATCHING REMAINS DIFFICULT PROBLEM

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



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



<https://blog.linkedin.com/2019/sepember/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

The screenshot displays the Taledo search interface. At the top, there are filter sections: 'Must-have skills (max 3)' with 'Java x' and 'Spring (Java framework) x' selected; 'Locations' with 'Berlin, Germany x' and a dropdown 'Already here or willing t...'; and 'Nice-to-have skills' with 'English-fluent x', '3 - 6 years x', and 'Salary x' selected. A 'CLEAR' button is also present. Below the filters, there are two tabs: 'All candidates' (selected) and 'Recommended'. The candidate list shows two entries: 'C. M.' with a score of 92 and a match description 'Found Java, Spring java framework.', and 'M. A.' with a score of 87. A 'SHORTLIST' button is visible next to the first candidate.

Must-have skills (max 3) Locations Display candidates within 100km radius

Java x Spring (Java framework) x Berlin, Germany x Already here or willing t... ▾

Nice-to-have skills English-fluent x 3 - 6 years x Salary x CLEAR

👁 All candidates ⌵ Recommended

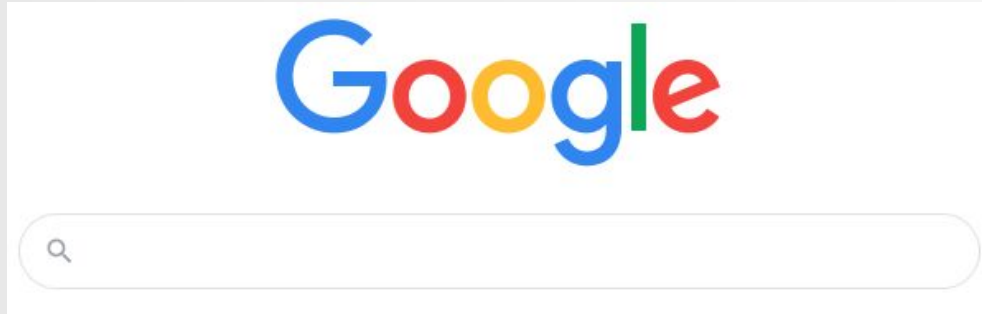
**C. M.** 92  
Found Java, Spring java framework.  
SHORTLIST ▾

**M. A.** 87

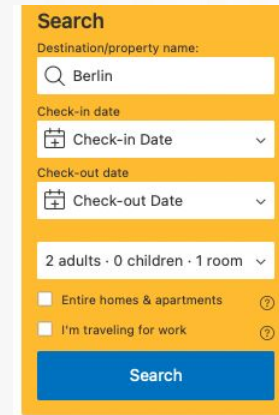
# BEYOND AI



# PRODUCT DESIGN MATTERS: SEARCH INPUTS



google.com

A screenshot of the Booking.com search form. It features a yellow header with the word "Search" in white. Below the header, there are several input fields: "Destination/property name:" with a search icon and the text "Berlin"; "Check-in date" with a calendar icon and a dropdown arrow; "Check-out date" with a calendar icon and a dropdown arrow; and "2 adults · 0 children · 1 room" with a dropdown arrow. There are also two checkboxes: "Entire homes & apartments" and "I'm traveling for work", each with a question mark icon. At the bottom is a blue "Search" button.

## Filter by:

### Health & safety

- Properties that take health & safety measures 942

### Popular Filters

- Hotels 537
- Indoor pool 28
- 5 stars/other ratings 39
- Less than 1 km 29  
Distance from center of Berlin
- Sauna 134
- Wonderful: 9+ 163  
Based on guest reviews
- Vacation Homes 47
- Guesthouses 75

### Stars and other ratings

- 1 star/other ratings 16

booking.com



Taledo.com



## TALEDO: PRODUCT AND AI COOPERATE

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

The screenshot displays a job search interface with the following elements:

- Must-have skills (max 3):** Java x, Spring (Java framework) x
- Locations:** Berlin, Germany x. Display candidates within 100km radius. A dropdown menu is open with options: "Already here or willing to...", "Recommended" (highlighted in a red box), "Best Fit", and "Recency".
- Nice-to-have skills:** English-fluent x, 3 - 6 years x, Salary x. A "CLEAR" button is present.
- Sort options:** "All candidates" (eye icon), "Recommended" (dropdown arrow icon).
- Candidate C. M.:** Found Java, Spring java framework. A "SHORTLIST" button with a dropdown arrow is visible.
- Candidate M. A.:** A "87" badge is visible in the bottom right corner.

## TALEDO: PRODUCT AND AI COOPERATE

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

The screenshot displays a job search interface. At the top left, a red box highlights the 'Reduce bias' toggle switch, which is currently turned on. Below this, the search filters are visible: 'Must-have skills (max 3)' includes 'Java x' and 'Spring (Java framework) x'; 'Locations' is set to 'Berlin, Germany x' with a '100km radius' filter; and 'Nice-to-have skills' includes 'English-fluent x', '3 - 6 years x', and 'Salary x'. A 'CLEAR' button is also present. On the right, there are options for 'All candidates' and 'Recommended', with a dropdown menu showing 'Recommended', 'Best Fit', and 'Recency'. The search results list a candidate 'C. M.' with a red box around their profile picture and name. Below their name, it says 'Found Java, Spring java framework.' and a 'SHORTLIST' button is visible. At the bottom right, a blue box contains the number '87'.

## TALEDO: THE NATURE OF RECRUITING IS HUMAN

- High service level
- Important decision
- Trust

**People business**



[https://unsplash.com/photos/376KN\\_ISpIE](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**

