



The State of AI 2016

Ines Montani
Matthew Honnibal
Explosion AI



Explosion AI is a digital studio specialising in Artificial Intelligence and Natural Language Processing.

- custom **algorithms**, applications and **data assets**
- makers of **spaCy**, the fastest-growing open-source **NLP** library
- our backgrounds are **computer science**, **linguistics** and **front-end development**

We are publishing...



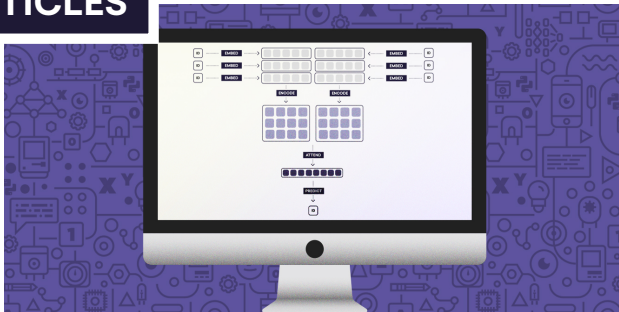
OPEN SOURCE



spaCy

Industrial-Strength Natural Language Processing in Python

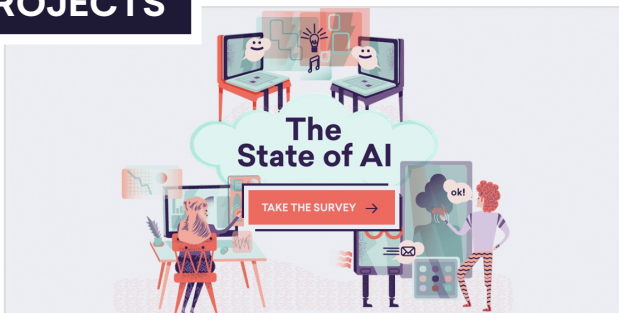
ARTICLES



Embed, encode, attend, predict

The new deep learning formula for state-of-the-art NLP models

PROJECTS



The State of AI

An open-data industry survey on the state of Artificial Intelligence in 2016

434 practitioners told us about the state of AI.



- industry survey launched on **October 10, 2016**
- collaboration with **Nathan Benaich** (Playfair Capital)
- results will be **100% open-source**
- you can still take part: **thestateofai.com**

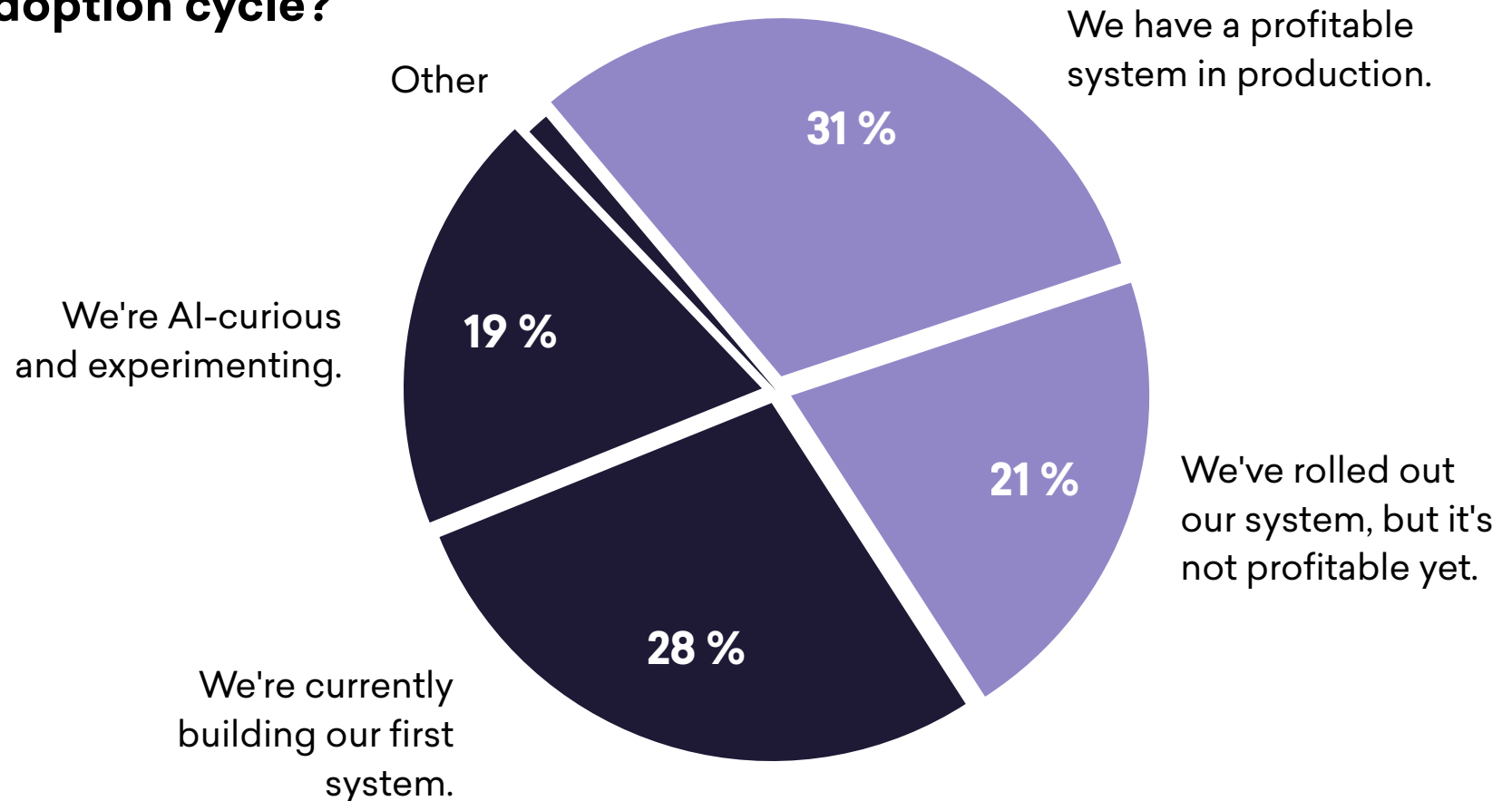


**The AI hype: about
90% justified.**

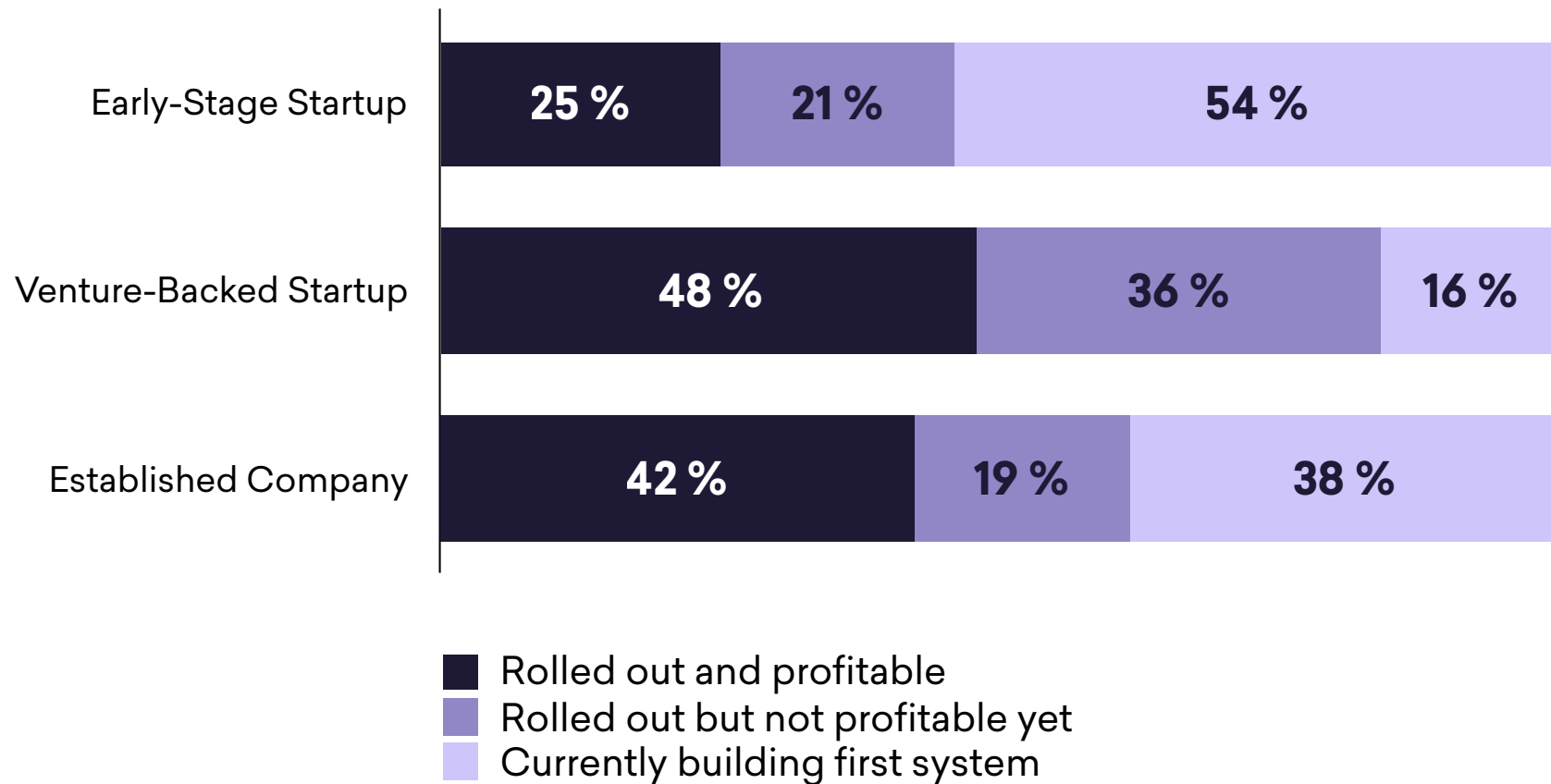
AI systems are already going into production.



**Where is your company
in the AI adoption cycle?**



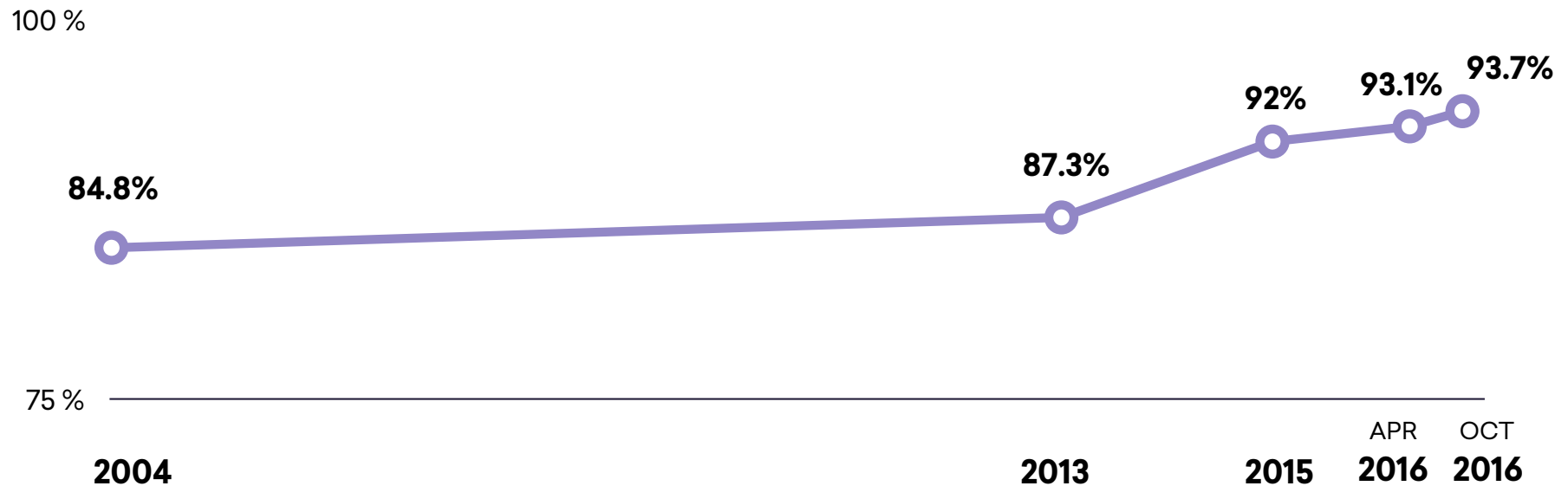
Young companies already have profitable systems.



NLP is improving faster than ever.



Example: Speech recognition accuracy





Researchers deliver blueprints, not products.

- **neural networks** really took off when **ImageNet** data set was created
- research runs on **benchmark tasks**
- benchmarks demonstrate solution **in principle**
- solutions in practice require **different data**

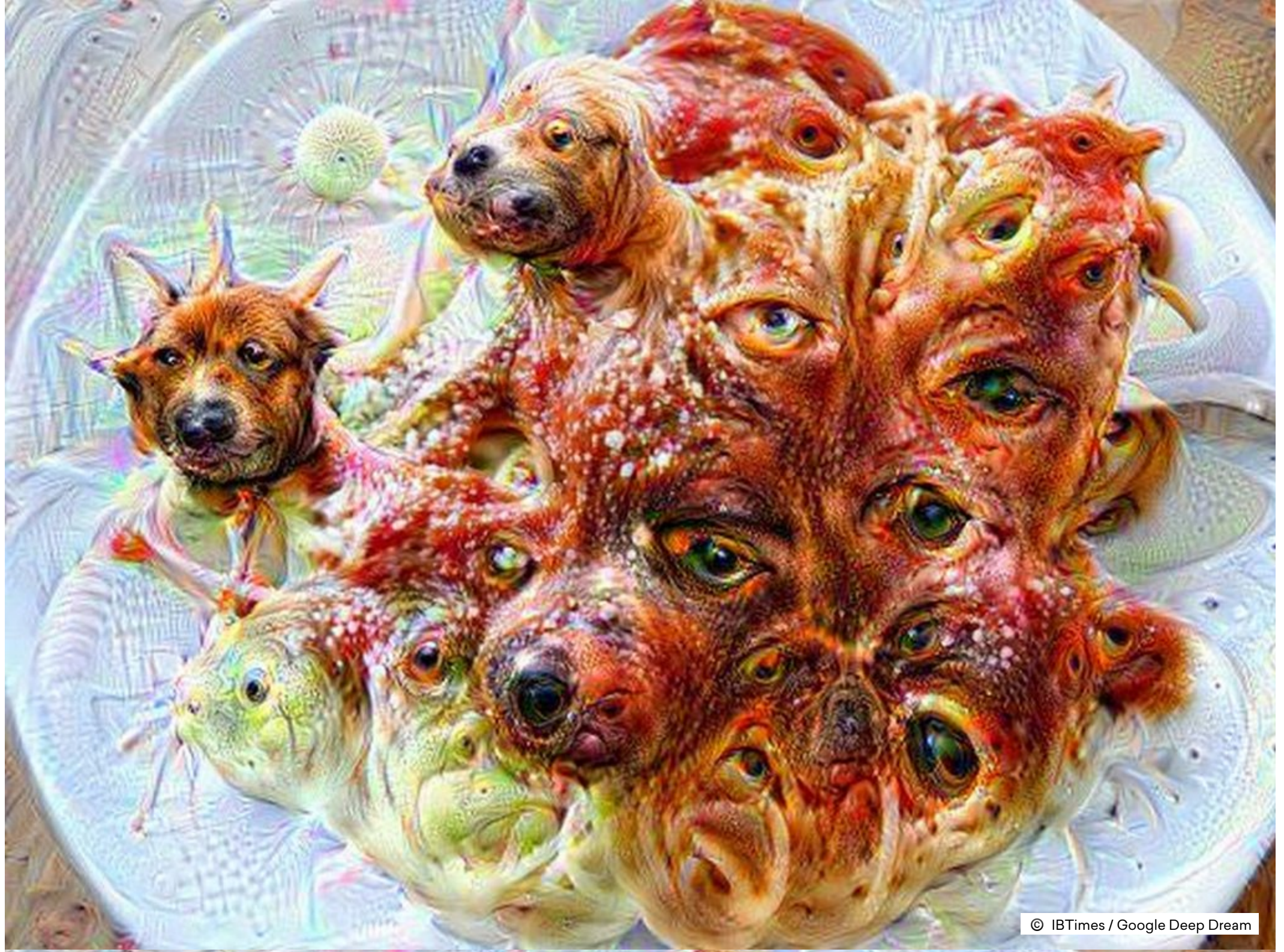
ImageNet: mostly dogs



Curly-Coated Retriever



Chesapeake Bay Retriever



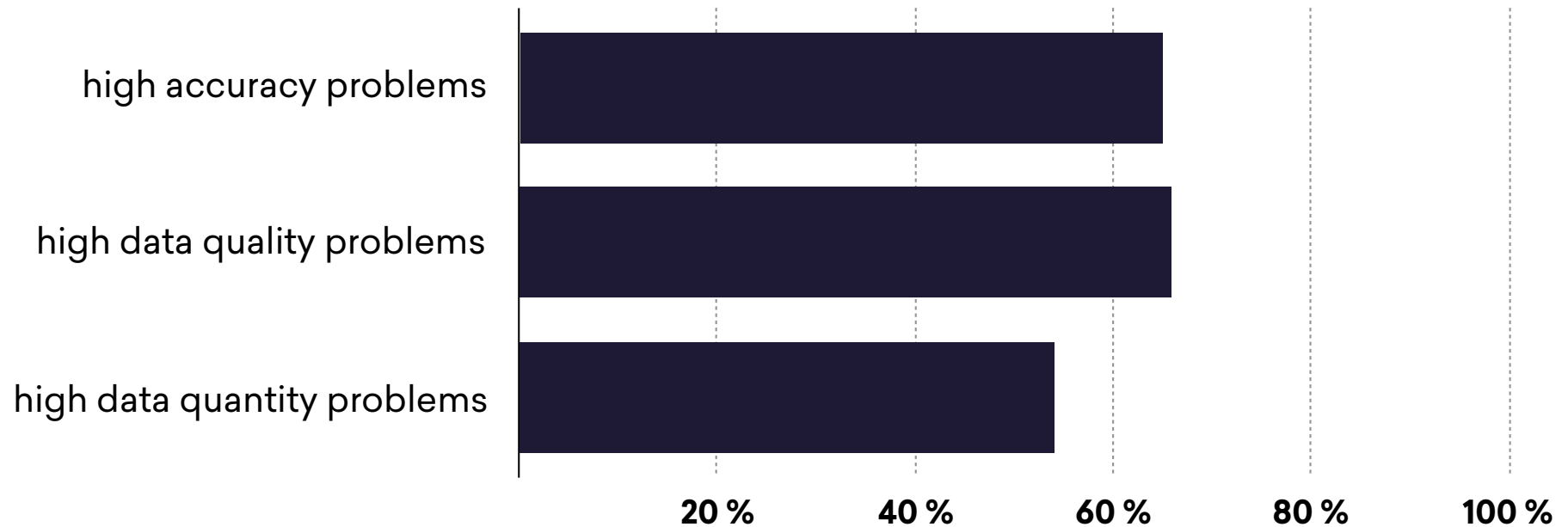


**The bottleneck in AI
is data, not algorithms.**

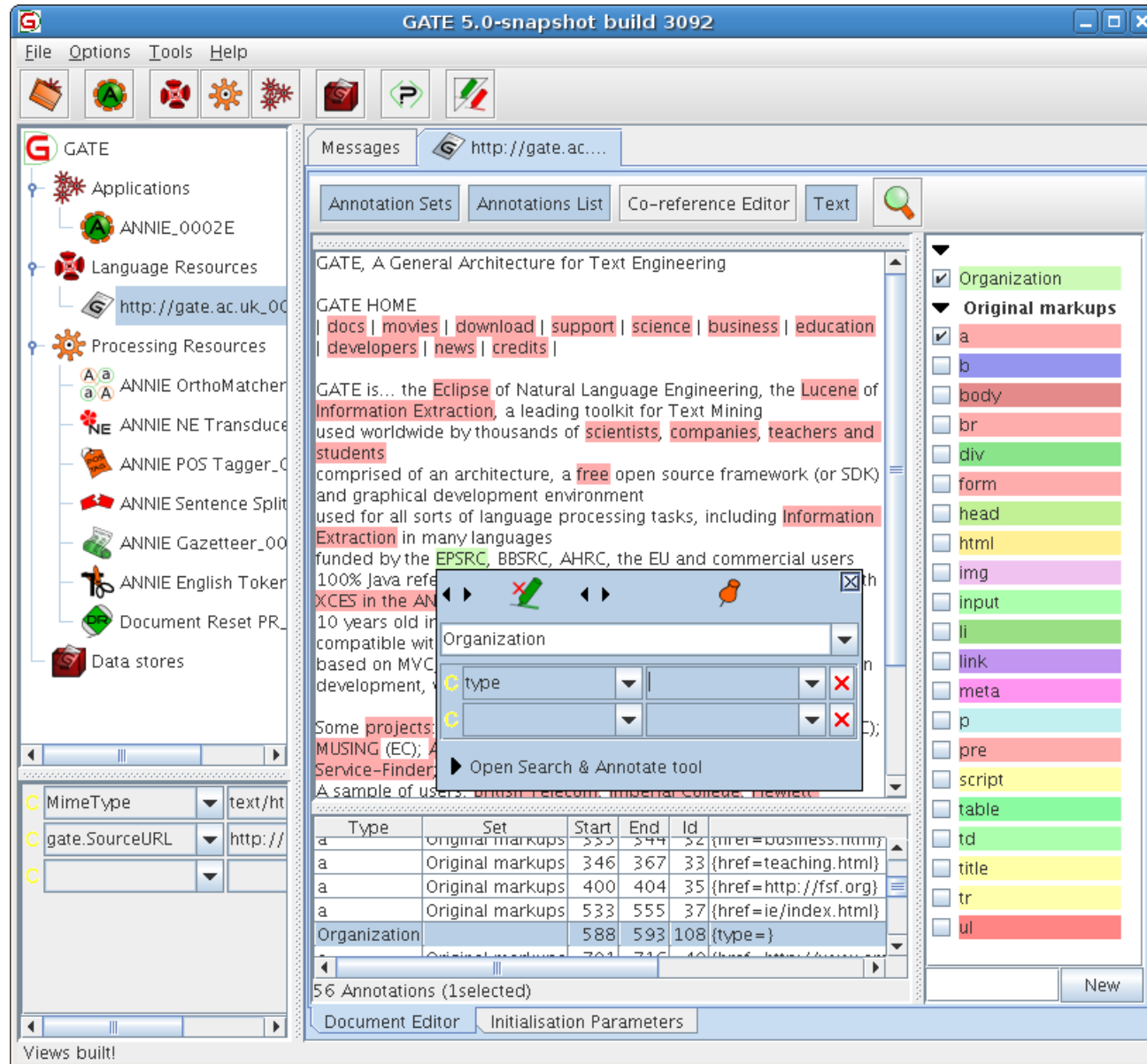
Data problems are still the biggest problems.



How much do the following problems keep you up at night? (4/5 and 5/5)



Also the state of AI 2016...



You can't get quality data by boring the shit out of underpaid people.



- machine learners need **machine teachers**
- Amazon **Mechanical Turk** is the most popular way to create **annotated data**
- **bad incentives**: below minimum-wage payment, outdated tools, **zero investment, zero satisfaction**



**We don't just
need innovation
on the back-end.**

It's time to apply what we know about UX to AI.



- the most valuable knowledge is in decisions humans **don't have to think about**
- **intuitive** interfaces, **fast** decisions
- **gamification**: take inspiration from what humans **enjoy** doing (and even pay for!)

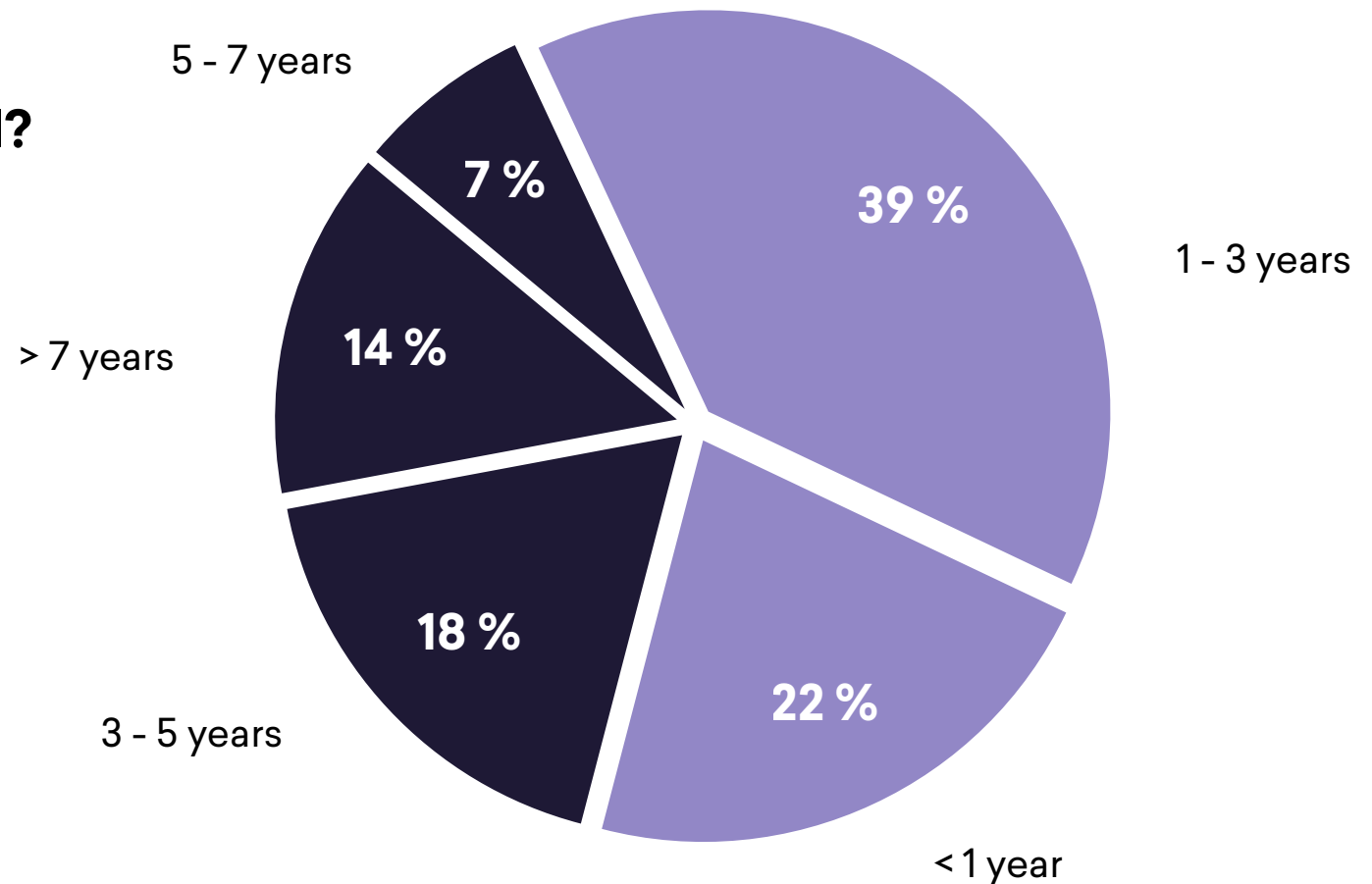
Don't assume your task can't be fun.



AI isn't a closed shop – there's a lot to contribute.




**How long have you
been working with AI?**



Thanks!



 **Take part in the survey**
thestateofai.com

 **Follow us on Twitter**
[@explosion_ai](https://twitter.com/explosion_ai)
[@_inesmontani](https://twitter.com/_inesmontani)
[@honnibal](https://twitter.com/honnibal)

 **Explosion AI**
explosion.ai