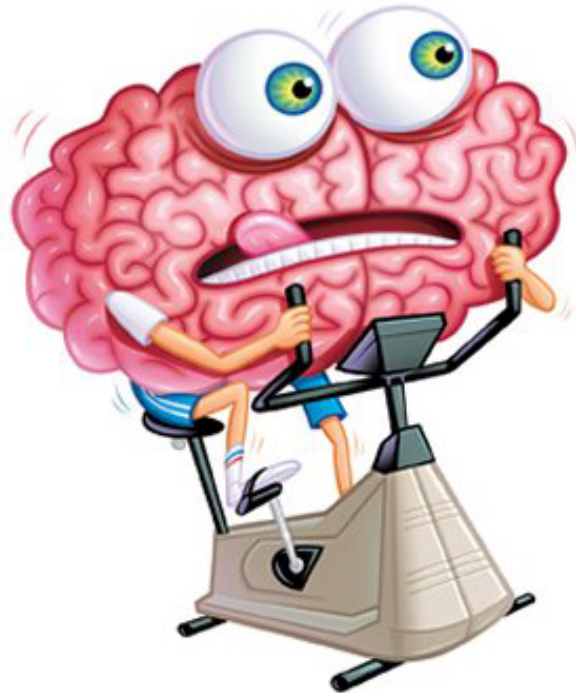


Priors, Deep Architectures and NLP :

YOU ARE DOING EVERYTHING WRONG!!

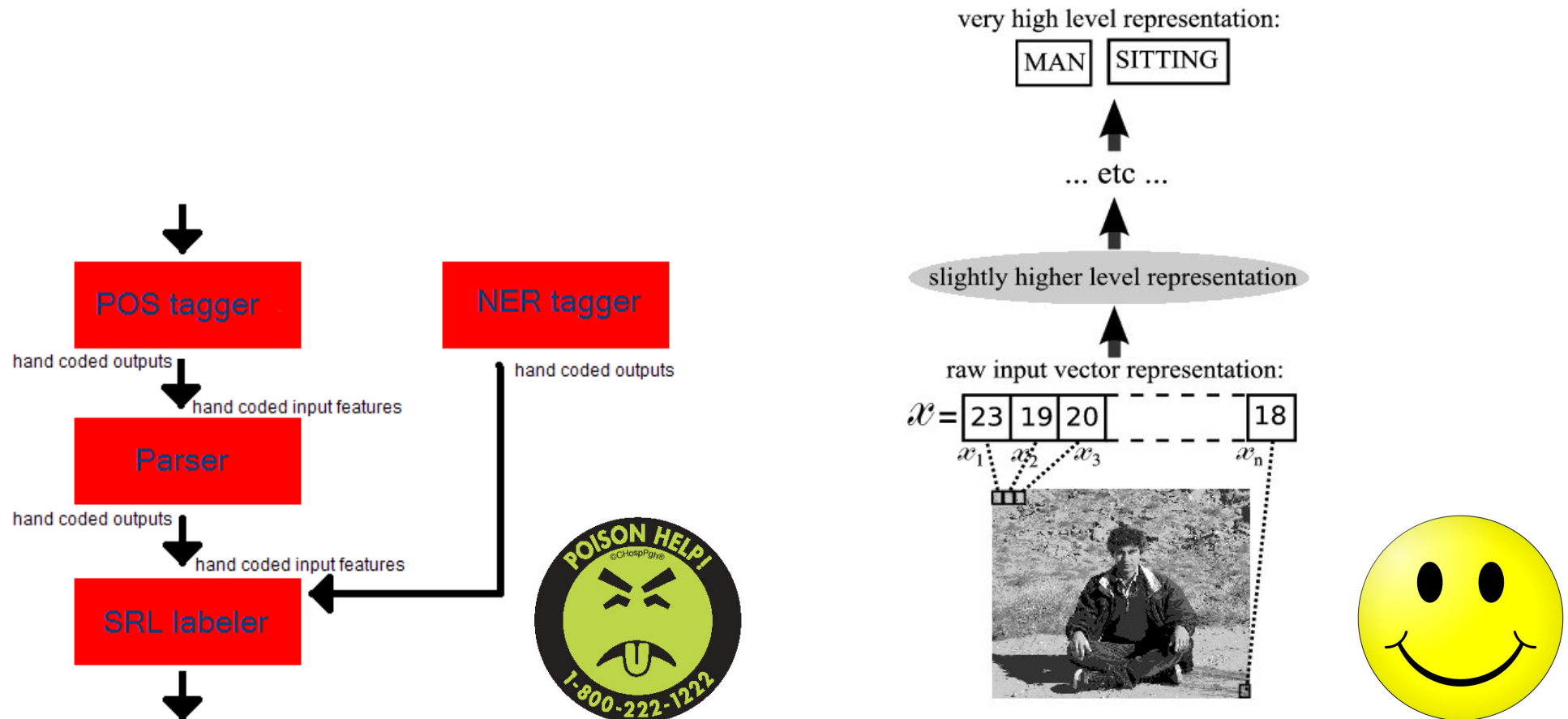


Jason Weston, NEC Labs America, Princeton, USA

Shallow System Cascades Vs Deep Architectures

Two types of prior:

- **Brittle** – system must use chosen input representation/ output labels
- **Robust** – priors injected into system, but it can ignore them/adapt
E.g: multi-tasking, choice of architecture ...



Shallow features + Shallow systems → Shallow results

I'm new to NLP – but because I wanted to publish something I feel like I'm already trapped in a poor direction of research!

Hand-made tags: POS, Parsing, SRL, ...

- brittle encodings of prior knowledge
- wrong prior for good research

Hand-made inputs:

- lots of time spent on making these for the wrong outputs!

ϵ % improvements in cascades won't lead to NLU.

Better: learn language to perform in an environment (NLU \neq NLP?)



Train a Brain !!!!!!!!!!!