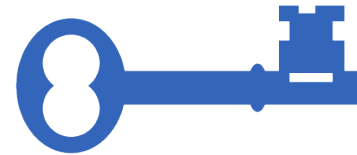
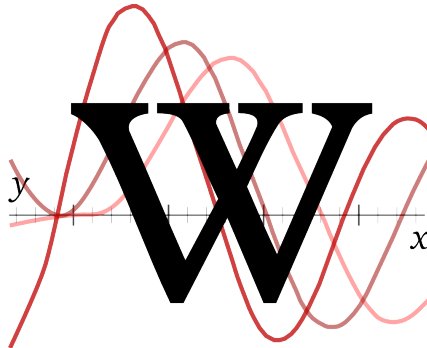
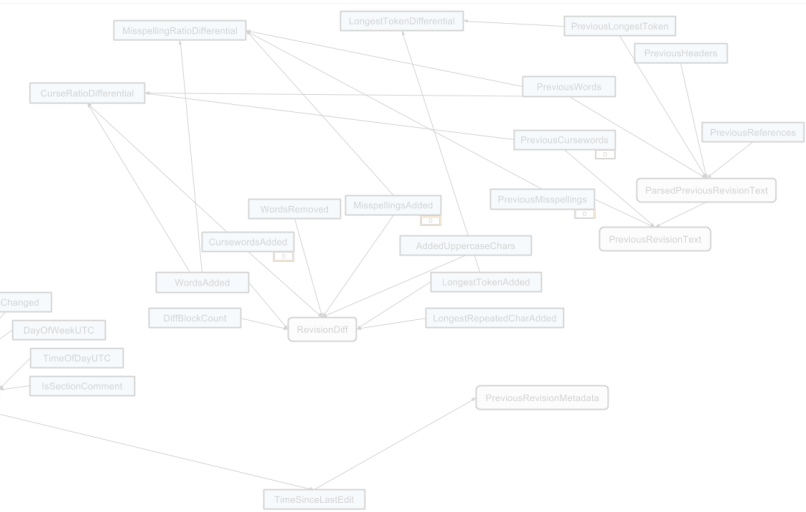


Revision Scoring as a Service

with Aaron Halfaker

[[User:EpochFail]]

[[User:Halfak (WMF)]]



IEG Funded



1. AI-based tools are critical for quality control in Wikipedia



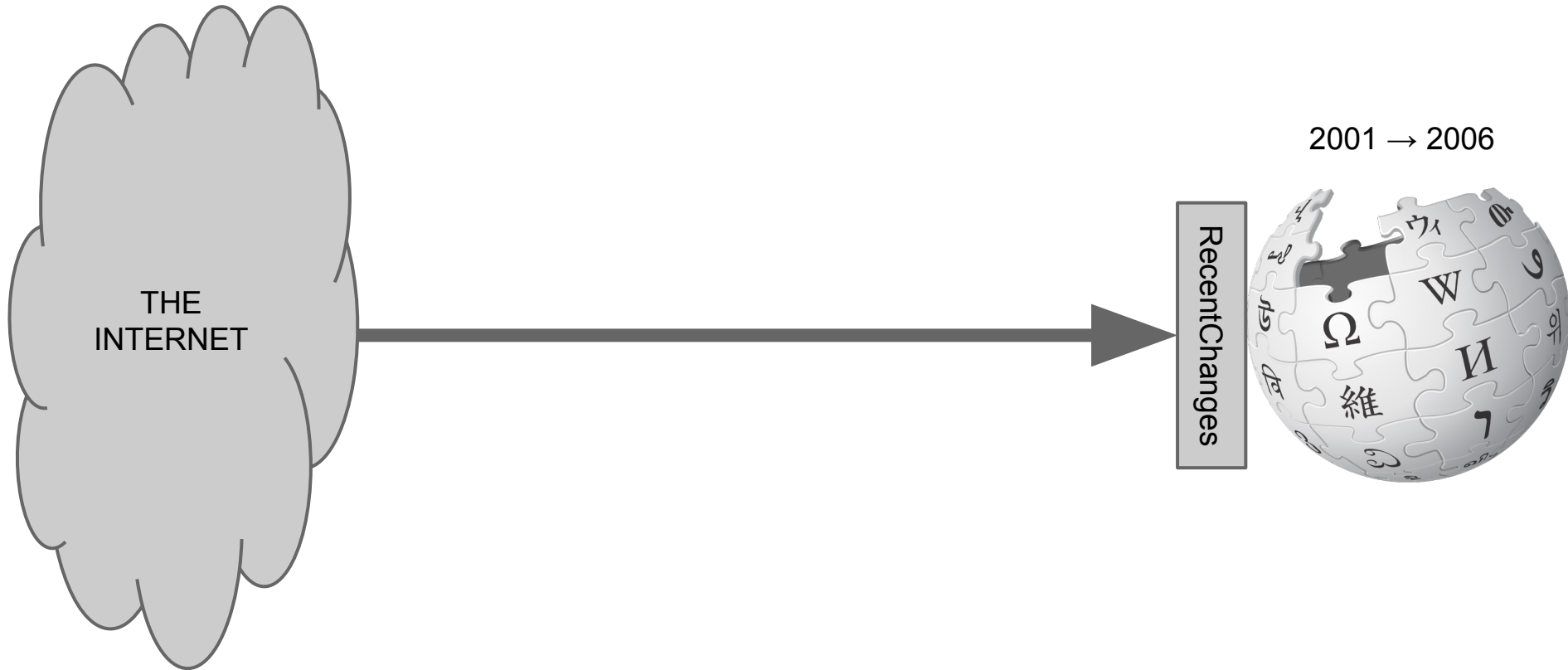
- 1. AI-based tools are critical for quality control in Wikipedia**
- 2. AI is hard -- requires expertise**



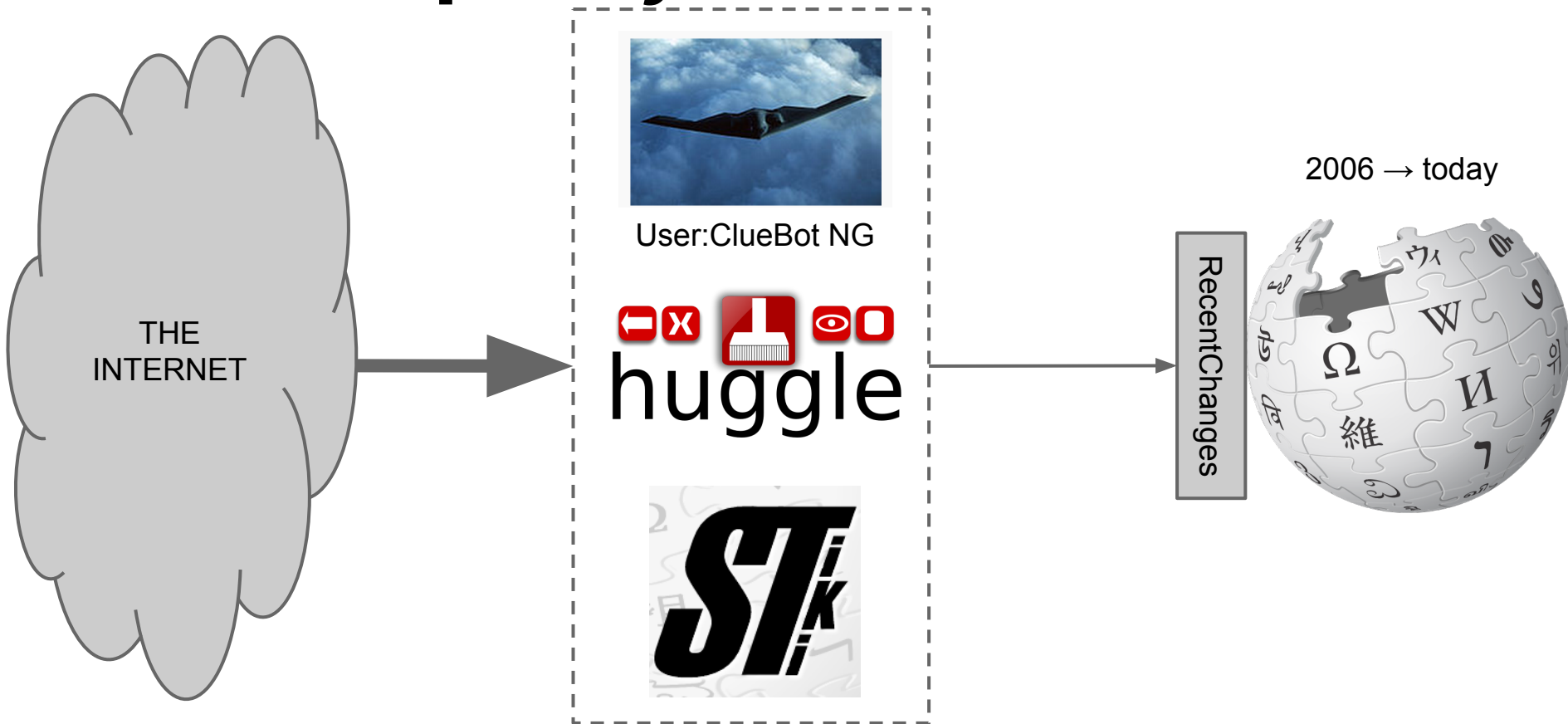
1. **AI-based tools are critical for quality control in Wikipedia**
2. **AI is hard -- requires expertise**
3. **Wrapping AI in libraries and services enables new tools**



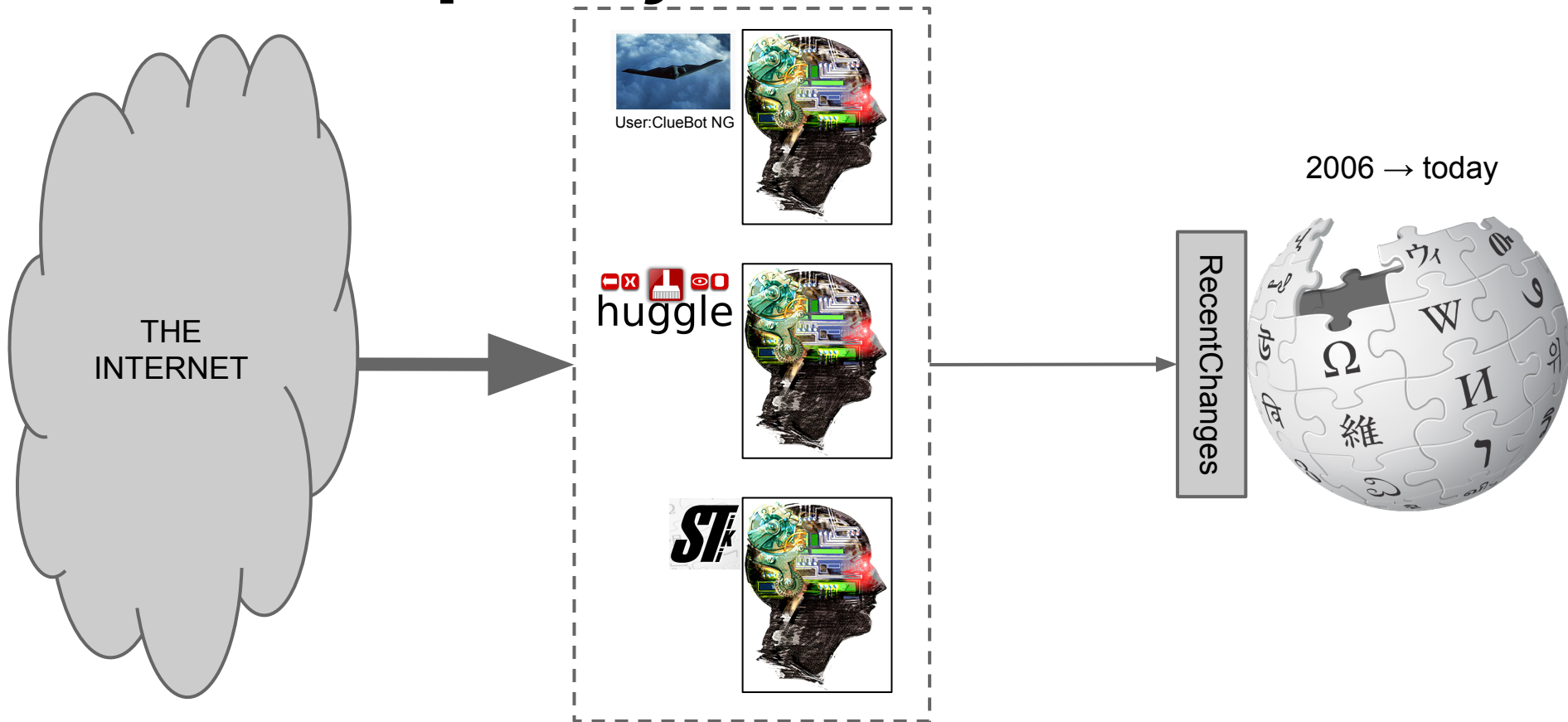
AI-based quality control



AI-based quality control



AI-based quality control



REVISION
#638307884



SCORE

Revision as of 23:14, 10 December 2014 (edit)

Blaise Mann (talk | contribs)

m (Adding 1 category)

← Previous edit

Revision as of 03:47, 16 December 2014 (edit) (undo) (thank)

Eloquence (talk | contribs)

m (← Theater: link to disambiguated title (currently a redlink) instead of disambig page)

Next edit ←

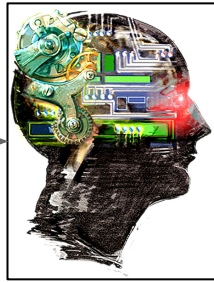
Line 35:

In 2013 Wilson, in collaboration with [[Mikhail Baryshnikov]] and co-starring [[Willem Dafoe]], developed "[[The Old Woman]]" an adaptation of the work by the Russian author [[Daniil Kharms]]. The play premiered at MIF13, Manchester International Festival.<ref>{{Cite web|url = http://www.mif.co.uk/event/the-old-woman|title = The Old Woman Robert Wilson, Mikhail Baryshnikov, Willem Dafoe|date = 2013|accessdate = 2014-11-23|website = Manchester International Festival|publisher = Manchester International Festival|last = Jansch|first = Lucie}}</ref> Wilson wrote that he and [[Mikhail Baryshnikov|Baryshnikov]] have discussed creating a play together for years, perhaps based on a Russian text.<ref>{{Cite web|url = http://calperformances.org/learn/program_notes/2014/pn_oldwoman.pdf|title = Cal Performances University of California, Berkeley Playbill|date = 2014|accessdate = 2014-11-23|website = Calperformances|publisher = UC Berkeley|last = Wilson|first = Robert}}</ref> The final production included dance, light, singing and bilingual monologue.

Line 35:

In 2013 Wilson, in collaboration with [[Mikhail Baryshnikov]] and co-starring [[Willem Dafoe]], developed "[[**The Old Woman (play)**]]The Old Woman]", an adaptation of the work by the Russian author [[Daniil Kharms]]. The play premiered at MIF13, Manchester International Festival.<ref>{{Cite web|url = http://www.mif.co.uk/event/the-old-woman|title = The Old Woman Robert Wilson, Mikhail Baryshnikov, Willem Dafoe|date = 2013|accessdate = 2014-11-23|website = Manchester International Festival|publisher = Manchester International Festival|last = Jansch|first = Lucie}}</ref> Wilson wrote that he and [[Mikhail Baryshnikov|Baryshnikov]] have discussed creating a play together for years, perhaps based on a Russian text.<ref>{{Cite web|url = http://calperformances.org/learn/program_notes/2014/pn_oldwoman.pdf|title = Cal Performances University of California, Berkeley Playbill|date = 2014|accessdate = 2014-11-23|website = Calperformances|publisher = UC Berkeley|last = Wilson|first = Robert}}</ref> The final production included dance, light, singing and bilingual monologue.

REVISION



SCORE

User:Eloquence fixes a link



armstrong.cis.upenn.edu/stiki_api.php?style=score&rid=638307884

0.00128848352474975

← Low score -- Probably good

User:Blank123456789 replaces a paragraph with “LLAMAS GROW ON TREES”



armstrong.cis.upenn.edu/stiki_api.php?style=score&rid=642215410

0.703287059009732

← High score -- Probably damaging

AI is important for quality control

Powerful

Fast

Minimizes human effort

When the Levee Breaks: Without Bots, What Happens to Wikipedia's Quality Control Processes? R. Stuart Geiger & Aaron Halfaker.
(2013). *WikiSym*.

But AI is hard

But AI is hard

- 20+ research papers on Wikipedia damage detection

But AI is hard

- 20+ research papers on Wikipedia damage detection
- “Machine classification” not part of standard CS degree

But AI is hard

- 20+ research papers on Wikipedia damage detection
- “Machine classification” not part of standard CS degree
 - Many tool devs don’t have a CS degree

But AI is hard

- 20+ research papers on Wikipedia damage detection
- “Machine classification” not part of standard CS degree
 - Many tool devs don’t have a CS degree
- Labor intensive, performance considerations, etc.

AI is hard

Each tool builds uses its own AI



User:ClueBot NG




huggle



ST



AI is hard

Each tool builds uses its own AI



User:ClueBot NG




huggle

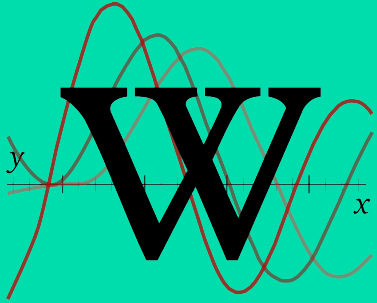


ST

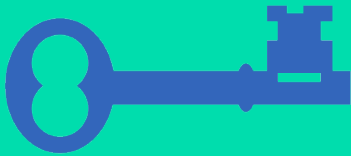


Mostly just English Wikipedia*

Revision Scoring Service



User:EpochFail



IEG Funded



User:He7d3r



User:とある白い猫

Volunteers

- User:Danilo.mac
- User:Jonas.agx
- User:West.g.andrew
- User:Yuvipanda

One AI to support them all

One AI to support them all



User:ClueBot NG




huggle



STK



One AI to support them all



User:ClueBot NG



One AI to support them all



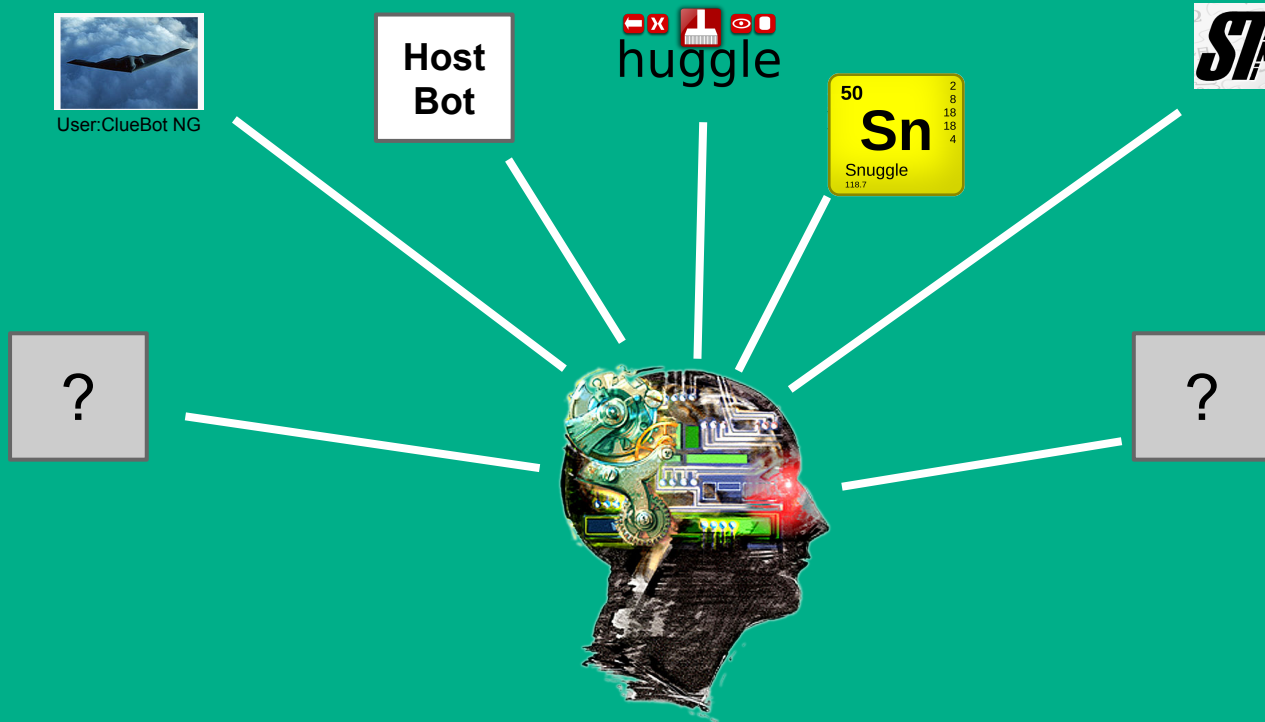
User:ClueBot NG

Host
Bot


huggle



One AI to support them all



One AI to rule them all

First release

- English
- Portuguese
- Turkish
- Azerbaijani*

* still working out some language feature issues

Web service

- ores.wmflabs.org/enwiki/
 - `scores=damaging&`
 - `revids=4567890|3456789`

```
{
  "4567890": {
    "damaging": {
      "probability": 0.234
    }
  },
  "3456789": {
    "damaging": {
      "probability": 0.952
    }
  }
}
```

Model files

- 100KB - 1MB
- Download, load and use

```
> from revscores.scorers import LinearSVCMoel
> model_file = open("enwiki.reverts.model", 'rb')
> model = LinearSVCMoel.load(model_file)
> model.score(...)
{"probability": 0.723}
```

Growing the wiki technology ecosystem

Growing the wiki technology ecosystem

- **IEG** -- Funding for distributed innovation

Growing the wiki technology ecosystem

- **IEG** -- Funding for distributed innovation
- **Wikimedia Labs** -- Raw materials and shared space

Growing the wiki technology ecosystem

- **IEG** -- Funding for distributed innovation
- **Wikimedia Labs** -- Raw materials and shared space
- **Services Team** -- Professional ecosystem builders



1. Wikipedia needs AI



1. Wikipedia needs AI

**2. We're making one important AI
easy**



- 1. Wikipedia needs AI**
- 2. We're making one important AI easy**
- 3. Growing the ecosystem of high tech wiki tools**



Thanks!

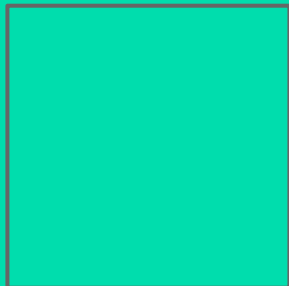
from the Revision Scoring team



User:EpochFail



User:He7d3r



User:とある白い猫



For more info:

- https://meta.wikimedia.org/wiki/Research:Revision_scoring_as_a_service
- https://meta.wikimedia.org/wiki/Grants:IEG/Revision_scoring_as_a_service