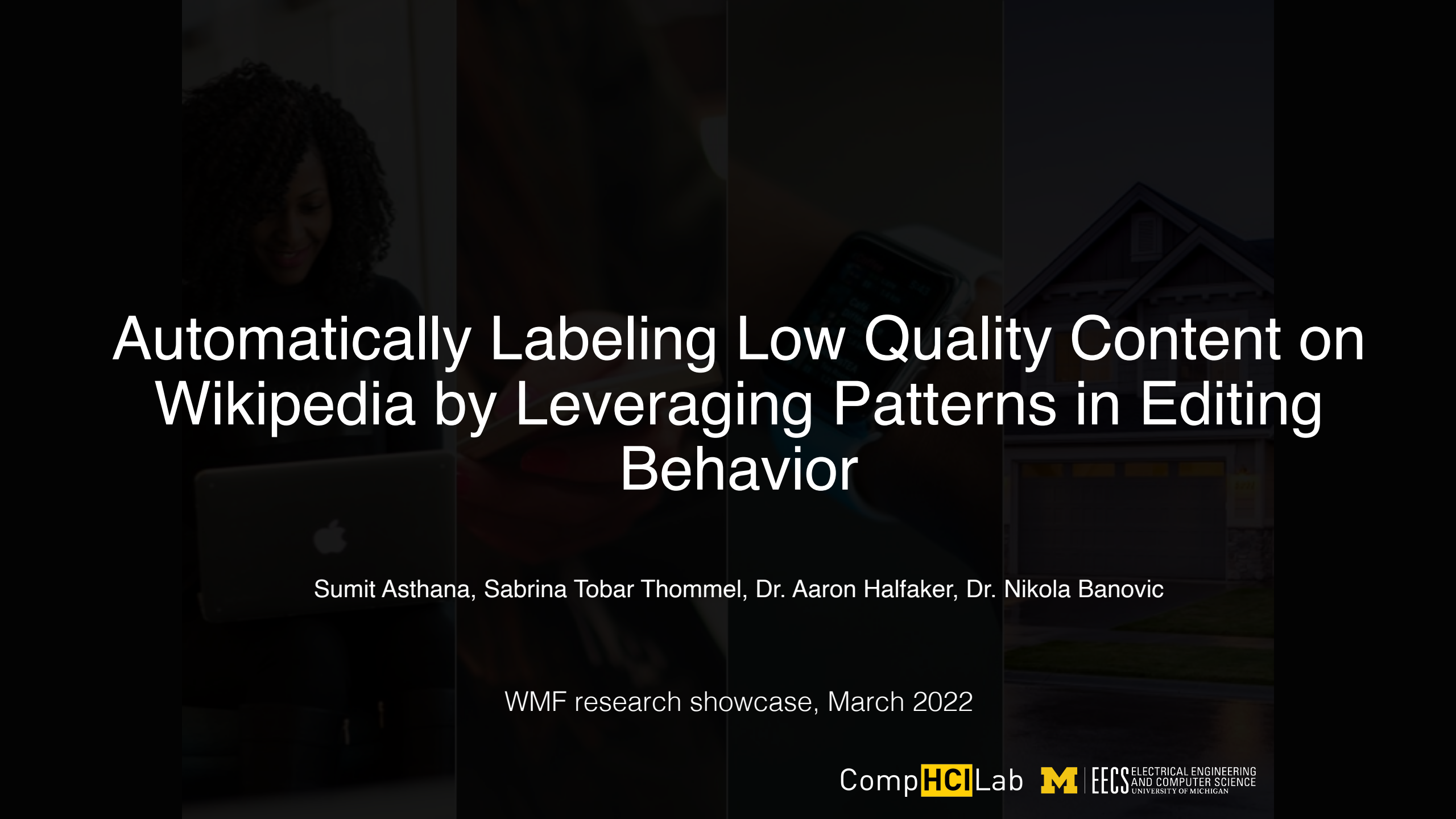


Automatically Labeling Low Quality Content on Wikipedia by Leveraging Patterns in Editing Behavior

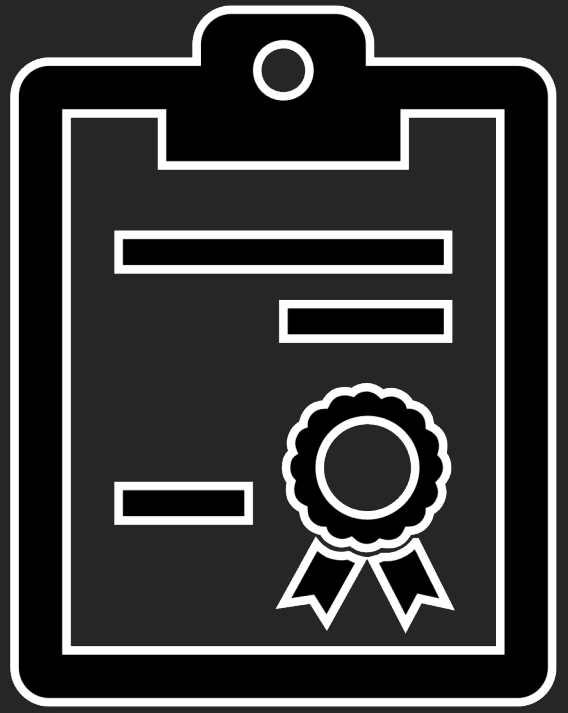
Sumit Asthana, Sabrina Tobar Thommel, Dr. Aaron Halfaker, Dr. Nikola Banovic

WMF research showcase, March 2022

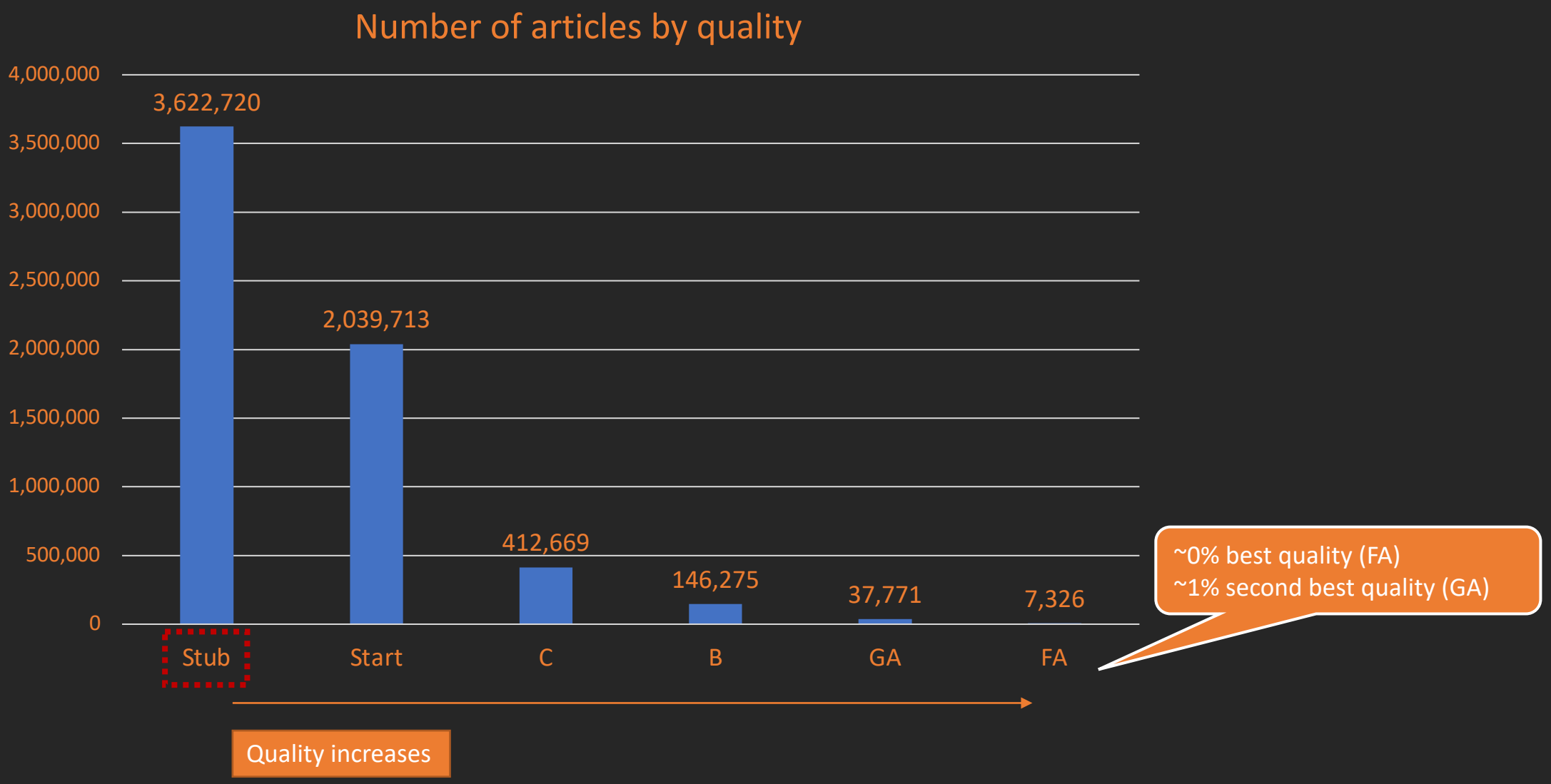
CompHCILab | EECS ELECTRICAL ENGINEERING AND COMPUTER SCIENCE UNIVERSITY OF MICHIGAN

Wikipedia’s Goal as an encyclopedia

Wikipedia aims to be an encyclopedia of high quality with well written, grammatically correct and neutral articles



English Wikipedia article quality statistics



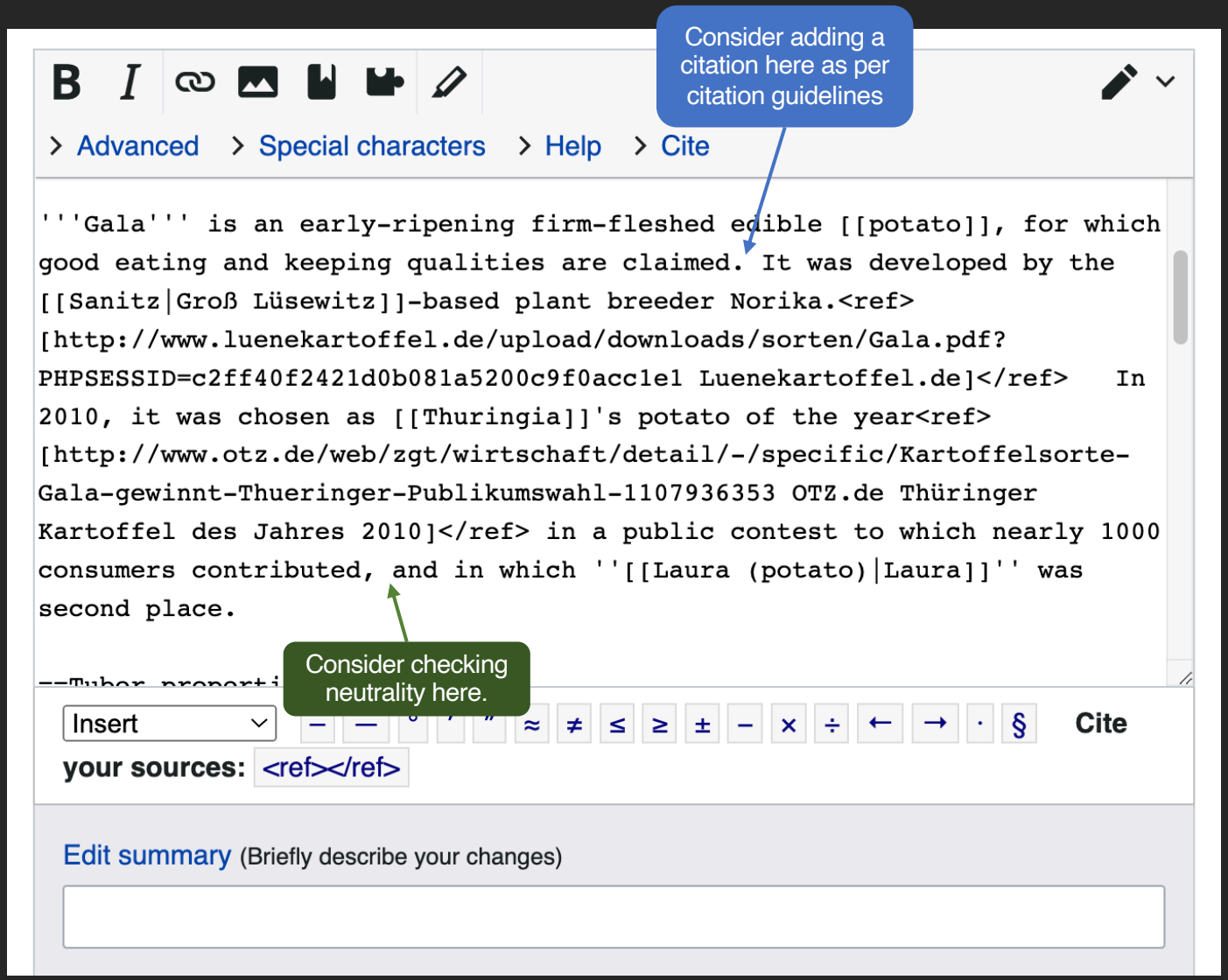
Challenges to Wikipedia's quality

Too many articles and too few editors has made article improvement a challenging task

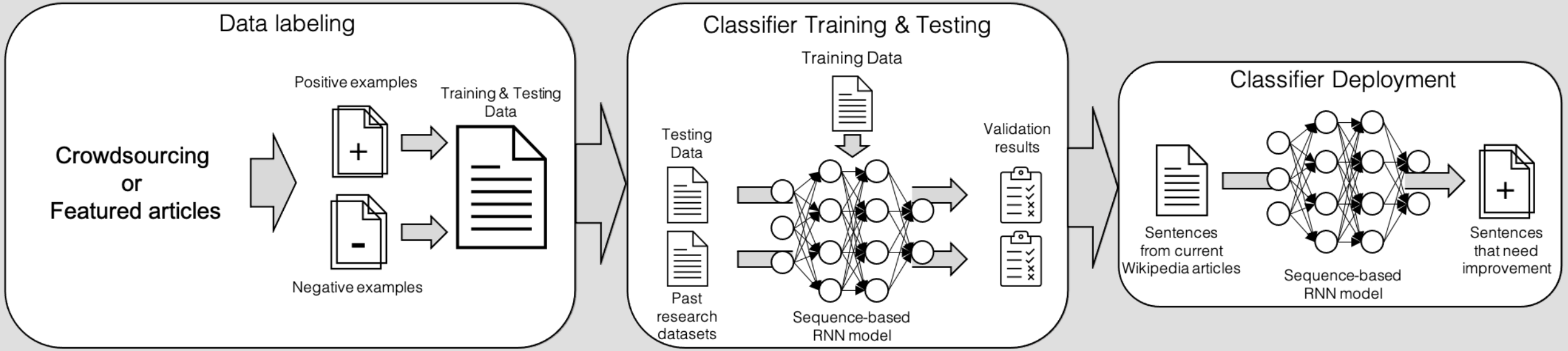
Credits: https://sciencenordic.com

Article improvement – add, assess, improve

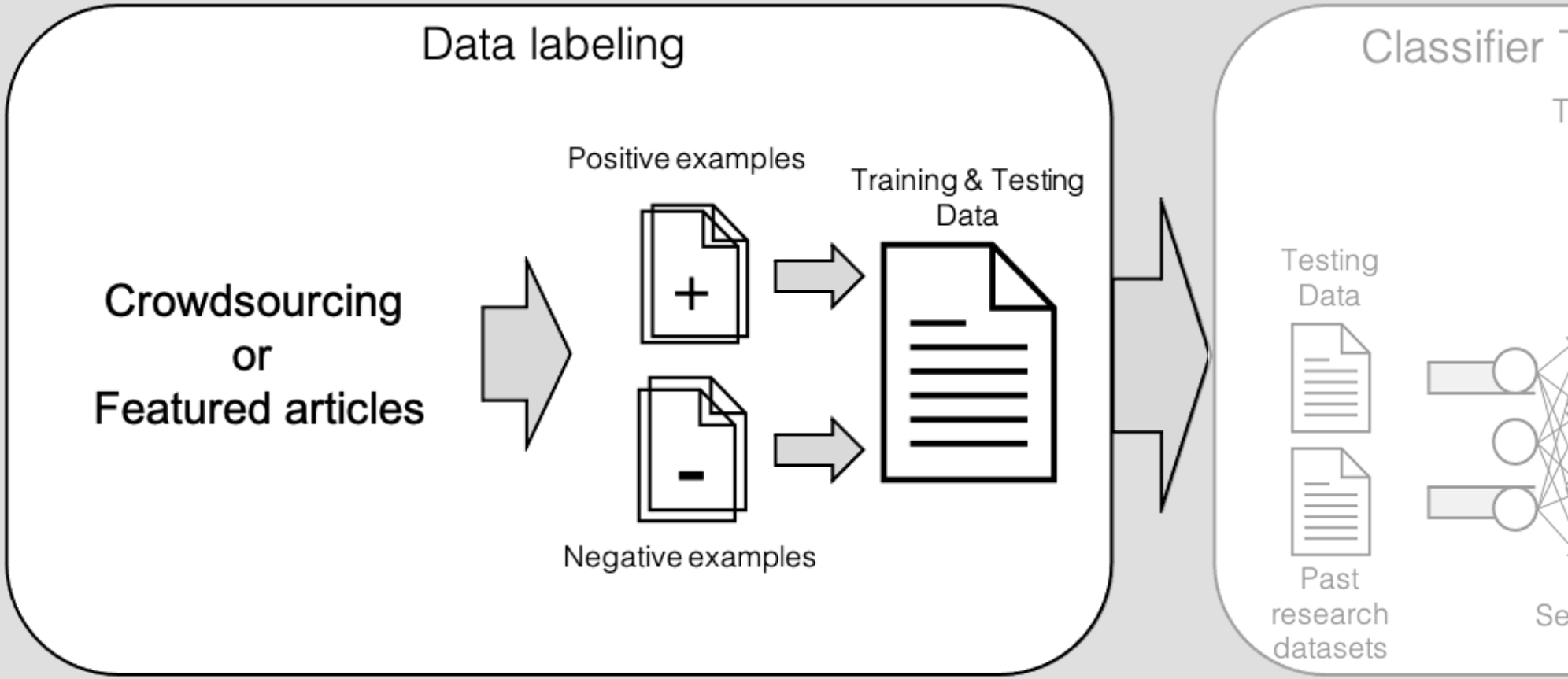
We need automation



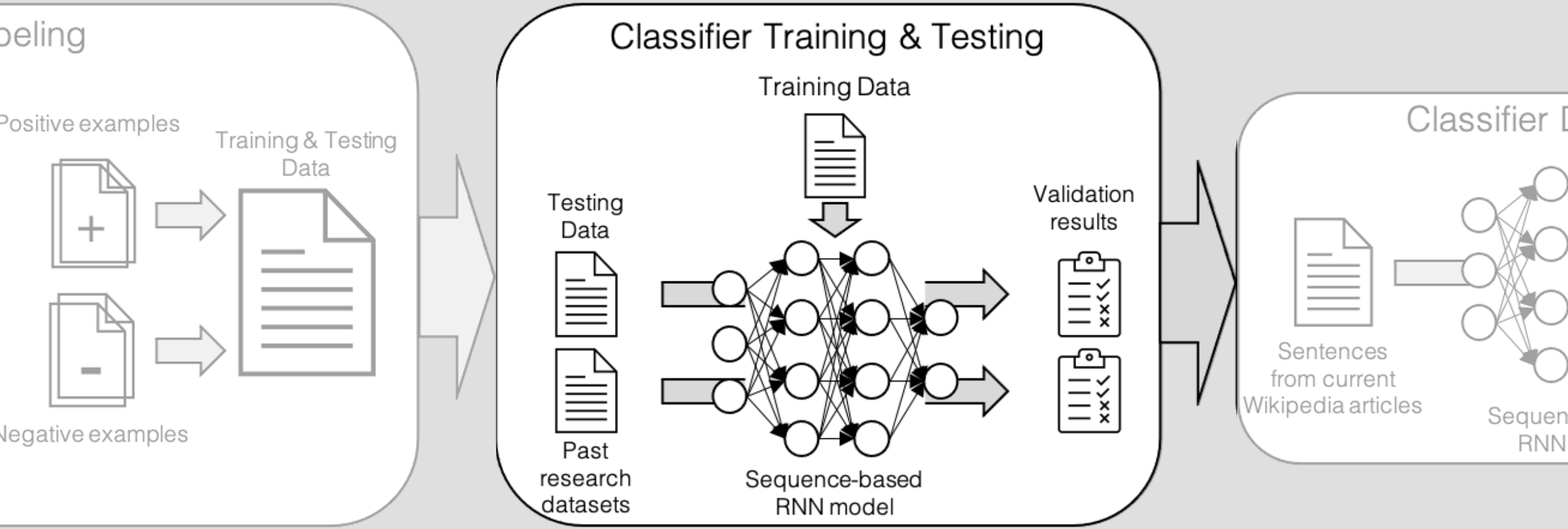
Automating article improvements using ML



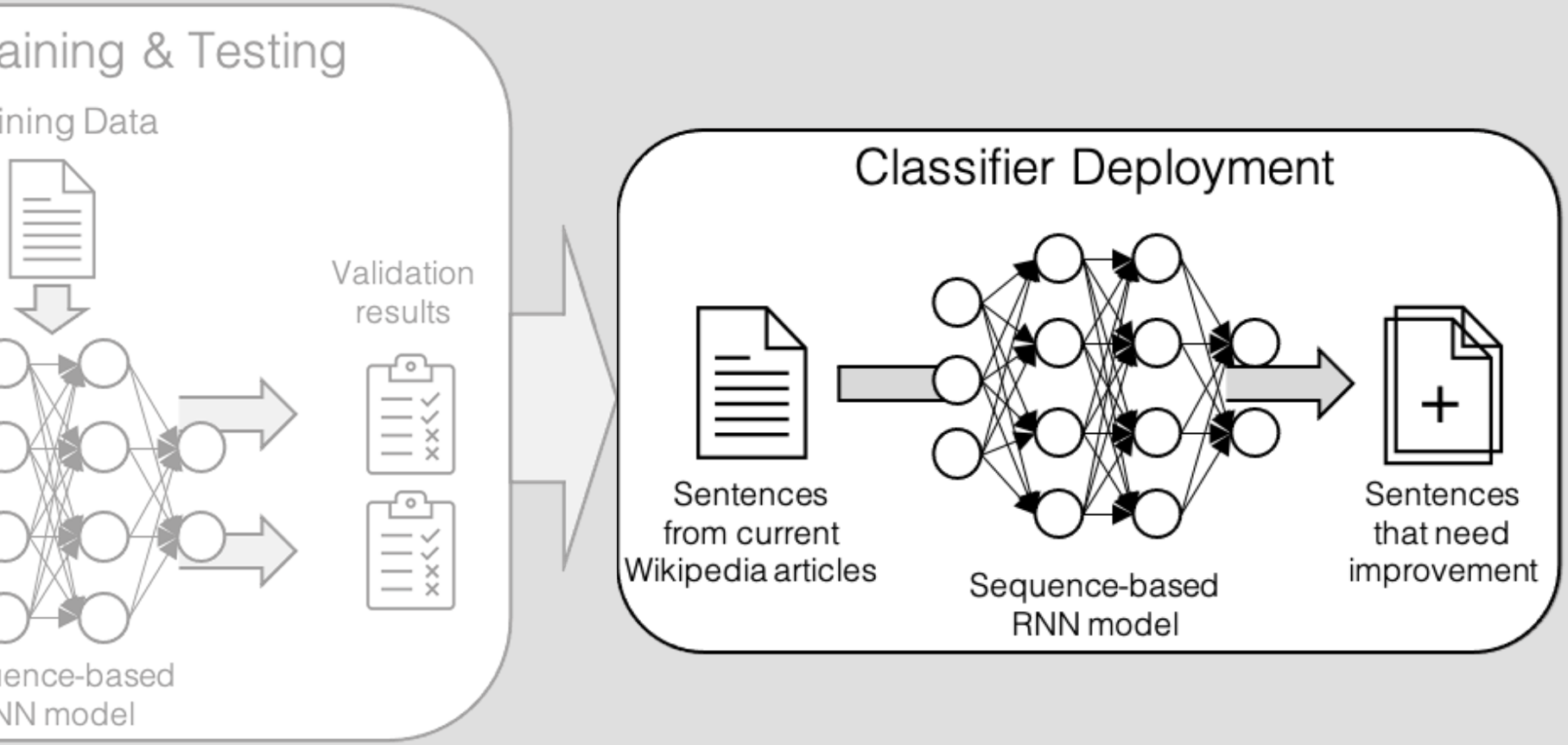
Automating article improvements using ML



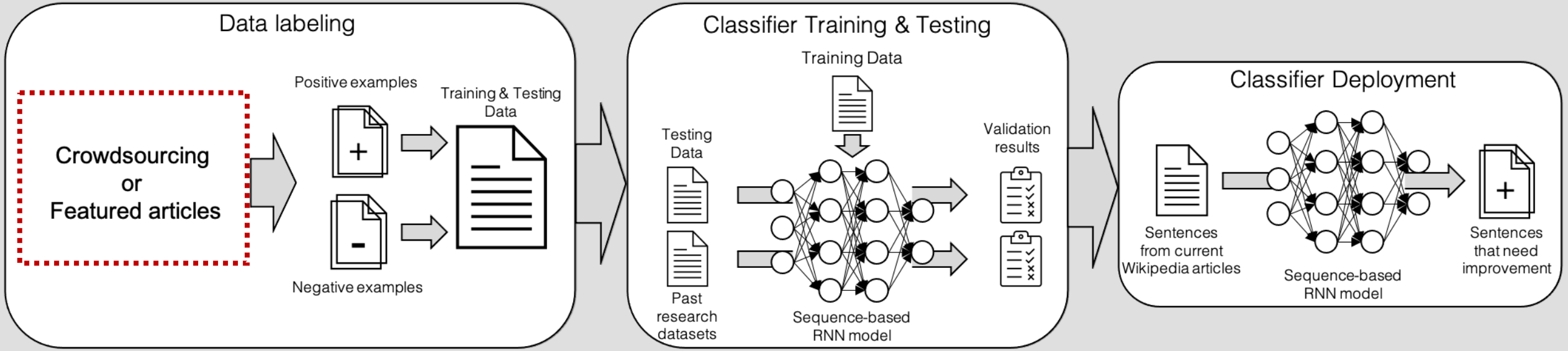
Automating article improvements using ML



Automating article improvements using ML



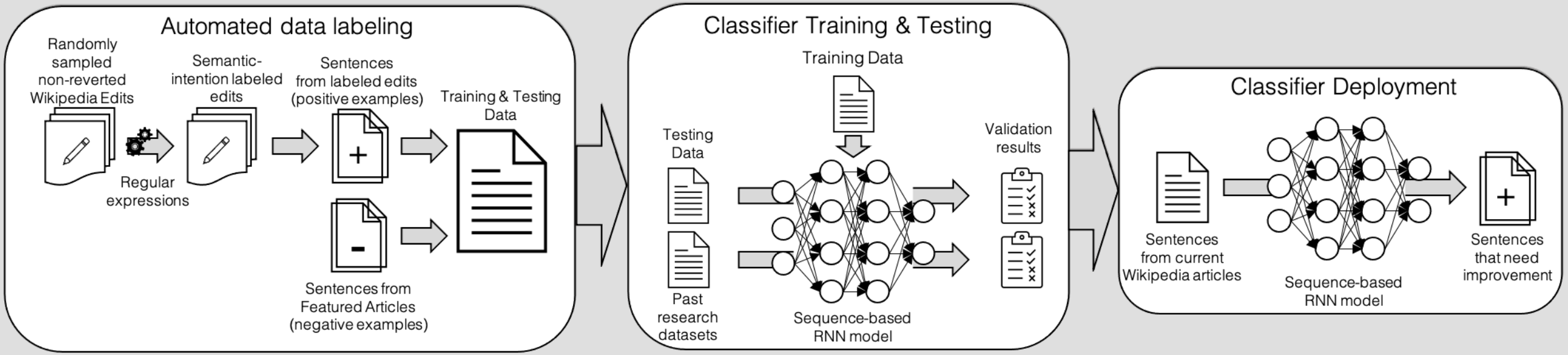
Challenges to automating article improvement using ML



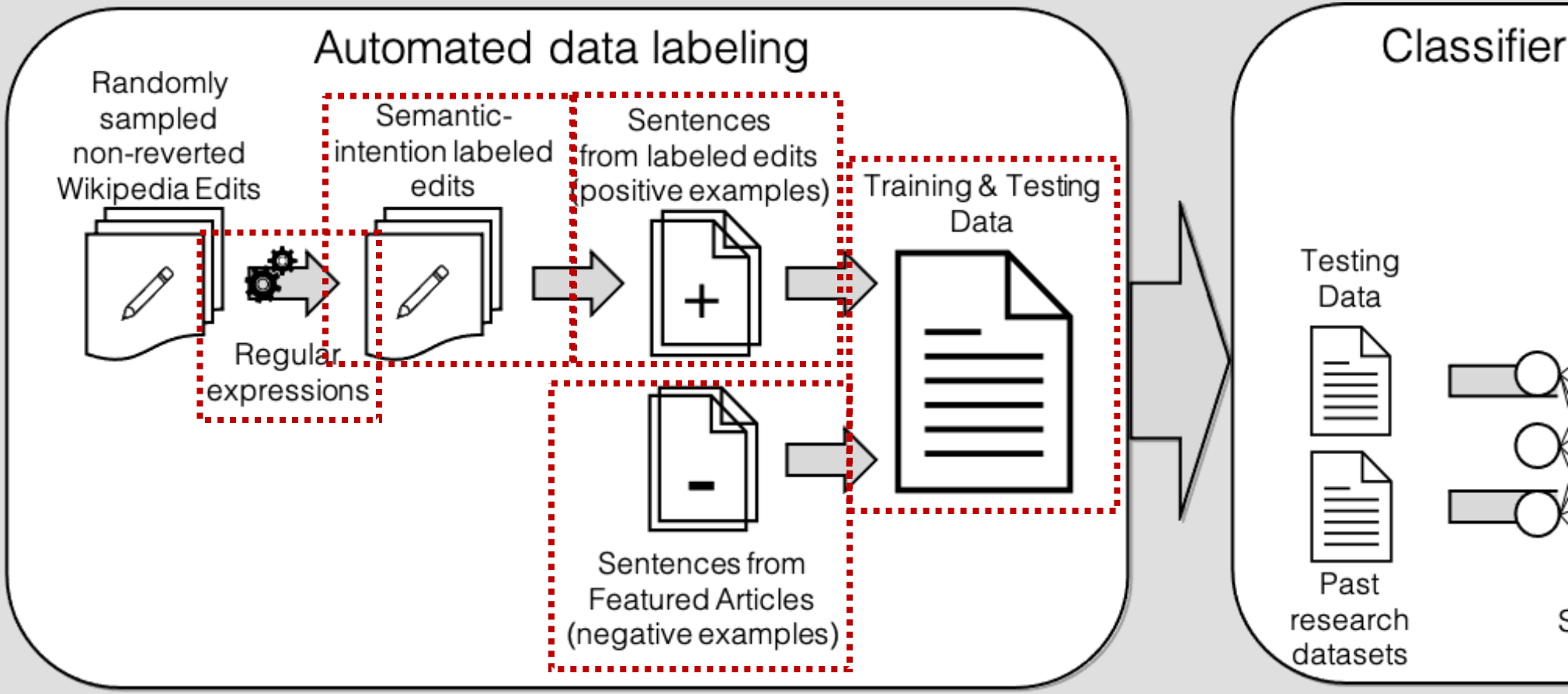
A new data quality labeling approach

A labeling approach that leverages implicit editing behavior to automatically label the quality of Wikipedia sentences

Automated labeling of sentence quality



Automated labeling of sentence quality



Regular expressions

..was born on July 27th, 1948 into the suburbs of… [0-9]+

…the compound is harmful {{citation-needed}}… {[a-z]+}

What needs improvement?

Does it have a point-of-view

Does it need fact updates?

"It was during 1996 season that Ricketson was first picked for City Origin due to his outstanding defence."

Does it need clarifications?

Does it need to be simplified?

Idea

- Many Wikipedia edits are performed with a specific intention
  - e.g. adding citation, removing bias, clarifying content.

- Sentences modified in Wikipedia edits can be used as positive examples of sentence improvement.

Semantic categories demonstrated

- Citations: adding a citation
- Point-of-view: rewrite using encyclopedic, neutral tone; remove bias; apply due weight.
- Clarification: specify or explain an existing fact or meaning by example or discussion without adding new information.

Diyi Yang, Aaron Halfaker, Robert Kraut, Eduard Hovy. 2017. Identifying Semantic Edit Intentions from Revisions in Wikipedia. In *ACL 2017*.

Semantic category rules on Wikipedia edits

Lateral flow test: Difference between revisions

From Wikipedia, the free encyclopedia

Browse history interactively

Revision as of 02:55, 9 June 2013 (edit)
Hipal (talk | contribs)
m *(Reverted edits by 131.111.184.97 (talk) to last version by 128.243.253.112)*
← Previous edit

Revision as of 14:47, 13 August 2013 (edit) (undo)
Yetisen (talk | contribs)
m *(A reference was added.)*
Next edit →

Line 1:

```
{{Context|date=October 2009}}
```

Line 1:

```
{{Context|date=October 2009}}
```

−

```
'''Lateral flow tests'''<ref>
[http://www.devicelink.com/ivdt/archive/99/11/009.html Concurrent
Engineering for Lateral-Flow Diagnostics (IVDT archive, Nov 99)<!-
- Bot generated title -->]</ref> also known as Lateral Flow
[[Immunochromatographic]] Assays are simple devices intended to
detect the presence (or absence) of a target analyte in sample
(matrix) without the need for specialized and costly equipment,
though many lab based applications exist that are supported by
reading equipment. Typically, these tests are used for medical
diagnostics either for home testing, [[point of care]] testing, or
laboratory use. A widely spread and well known application is the
home pregnancy test.
```

+

```
'''Lateral flow tests'''<ref>
[http://www.devicelink.com/ivdt/archive/99/11/009.html Concurrent
Engineering for Lateral-Flow Diagnostics (IVDT archive, Nov 99)<!-
- Bot generated title -->]</ref> also known as Lateral Flow
[[Immunochromatographic]] Assays are simple devices intended to
detect the presence (or absence) of a target analyte in sample
(matrix) without the need for specialized and costly equipment,
though many lab based applications exist that are supported by
reading equipment.<ref name="Yetisen2013">{{cite journal | author
= Yetisen A. K. | year = 2013 | title = Paper-based microfluidic
point-of-care diagnostic devices | url =
http://pubs.rsc.org/en/content/articlelanding/2013/lc/c3lc50169h |
journal = Lab on a Chip |volume= 13 |issue= 12 |pages= 2210–2251
|doi=10.1039/C3LC50169H}}</ref> Typically, these tests are used
for medical diagnostics either for home testing, [[point of care]]
testing, or laboratory use. A widely spread and well known
application is the home pregnancy test.
```

Semantic category rules on Wikipedia edits

Lateral flow test: Difference between revisions

From Wikipedia, the free encyclopedia

Browse history interactively

Revision as of 02:55, 9 June 2013 (edit)
Hipal (talk | contribs)
m *(Reverted edits by 131.111.184.97 (talk) to last version by 128.243.253.112)*
← Previous edit

Revision as of 14:47, 13 August 2013 (edit) (undo)
Yetisen (talk | contribs)
m *(A reference was added.)*
Next edit →

Intent: Citation

Inserted contains "<ref>"

Line 1:

```
{{Context|date=October 2009}}
```

Line 1:

```
{{Context|date=October 2009}}
```

−

```
'''Lateral flow tests'''<ref>
[http://www.devicelink.com/ivdt/archive/99/11/009.html Concurrent
Engineering for Lateral-Flow Diagnostics (IVDT archive, Nov 99)<!-
- Bot generated title -->]</ref> also known as Lateral Flow
[[Immunochromatographic]] Assays are simple devices intended to
detect the presence (or absence) of a target analyte in sample
(matrix) without the need for specialized and costly equipment,
though many lab based applications exist that are supported by
reading equipment. Typically, these tests are used for medical
diagnostics either for home testing, [[point of care]] testing, or
laboratory use. A widely spread and well known application is the
home pregnancy test.
```

+

```
'''Lateral flow tests'''<ref>
[http://www.devicelink.com/ivdt/archive/99/11/009.html Concurrent
Engineering for Lateral-Flow Diagnostics (IVDT archive, Nov 99)<!-
- Bot generated title -->]</ref> also known as Lateral Flow
[[Immunochromatographic]] Assays are simple devices intended to
detect the presence (or absence) of a target analyte in sample
(matrix) without the need for specialized and costly equipment,
though many lab based applications exist that are supported by
reading equipment.<ref name="Yetisen2013">{{cite journal | author
= Yetisen A. K. | year = 2013 | title = Paper-based microfluidic
point-of-care diagnostic devices | url =
http://pubs.rsc.org/en/content/articlelanding/2013/lc/c3lc50169h |
journal = Lab on a Chip |volume= 13 |issue= 12 |pages= 2210–2251
|doi=10.1039/C3LC50169H}}</ref> Typically, these tests are used
for medical diagnostics either for home testing, [[point of care]]
testing, or laboratory use. A widely spread and well known
application is the home pregnancy test.
```

Semantic category rules – citations

Rule	Description
Inserted contains “<ref>”	Presence of citation tag in added content
NOT is_citation_inserted_or_deleted	No <ref> tag inserted or removed
NOT is_template_inserted_or_deleted	No templates inserted or removed
NOT is_wikilink_inserted_or_deleted	No wikilinks inserted or removed
NOT is_infobox_inserted_or_deleted	No infoboxes inserted or removed
NOT is_multiline_inserted_or_deleted	Multiple lines not changed

Semantic intent to sentence quality label

Lateral flow test: Difference between revisions

From Wikipedia, the free encyclopedia

Browse history interactively

Revision as of 02:55, 9 June 2013 (edit)
Hipal (talk | contribs)
m *(Reverted edits by 131.111.184.97 (talk) to last version by 128.243.253.112)*
← Previous edit

Revision as of 14:47, 13 August 2013 (edit) (undo)
Yetisen (talk | contribs)
m *(A reference was added.)*
Next edit →

Intent: Citation

Line 1:

```
{{Context|date=October 2009}}
```

−

```
'''Lateral flow tests'''<ref>
[http://www.devicelink.com/ivdt/archive/99/11/009.html Concurrent
Engineering for Lateral-Flow Diagnostics (IVDT archive, Nov 99)<!-
- Bot generated title -->]</ref> also known as Lateral Flow
[[Immunochromatographic]] Assays are simple devices intended to
detect the presence (or absence) of a target analyte in sample
(matrix) without the need for specialized and costly equipment,
though many lab based applications exist that are supported by
reading equipment. Typically, these tests are used for medical
diagnostics either for home testing, [[point of care]] testing, or
laboratory use. A widely spread and well known application is the
home pregnancy test.
```

"needs citations": 1

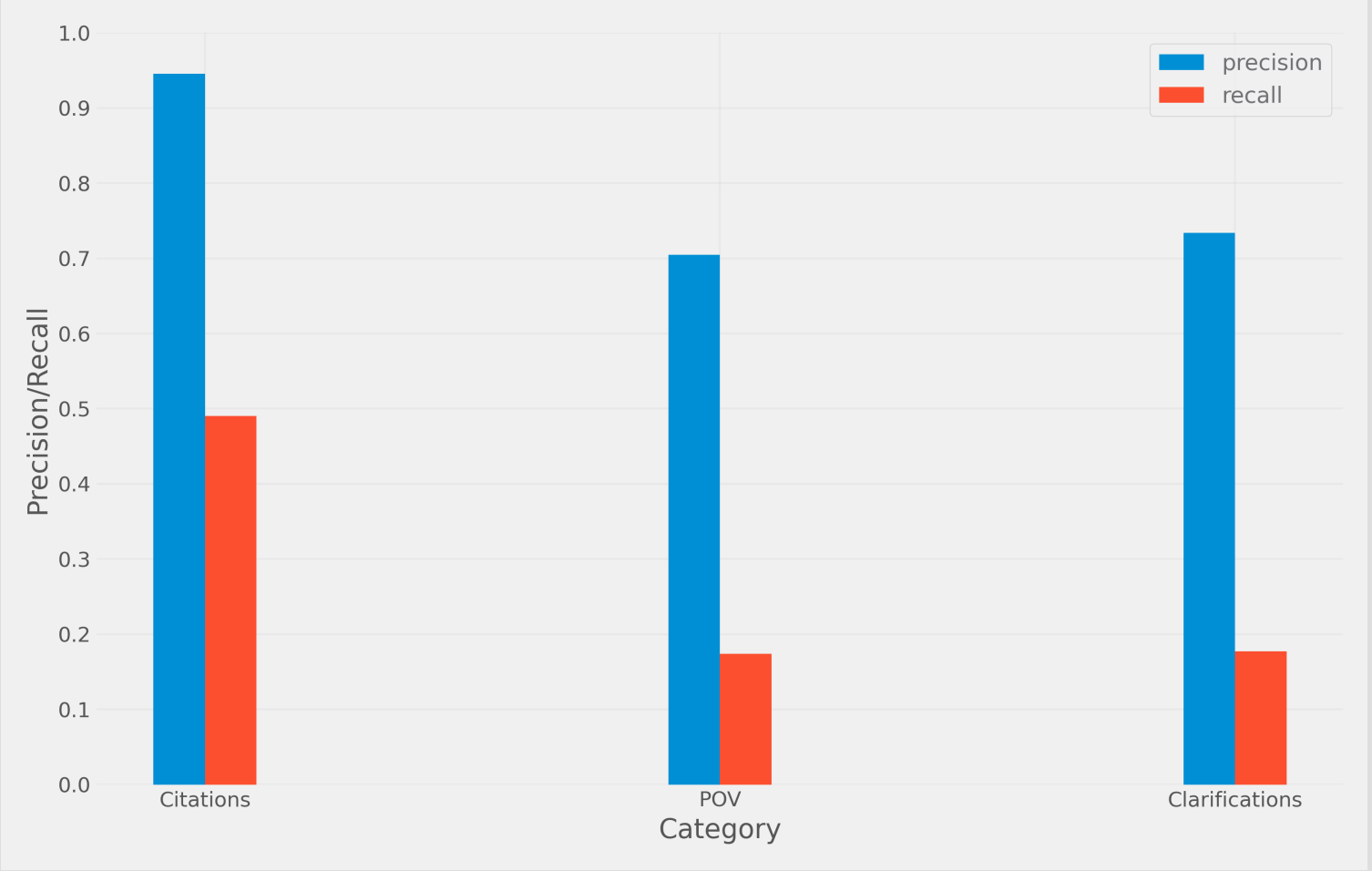
Line 1:

```
{{Context|date=October 2009}}
```

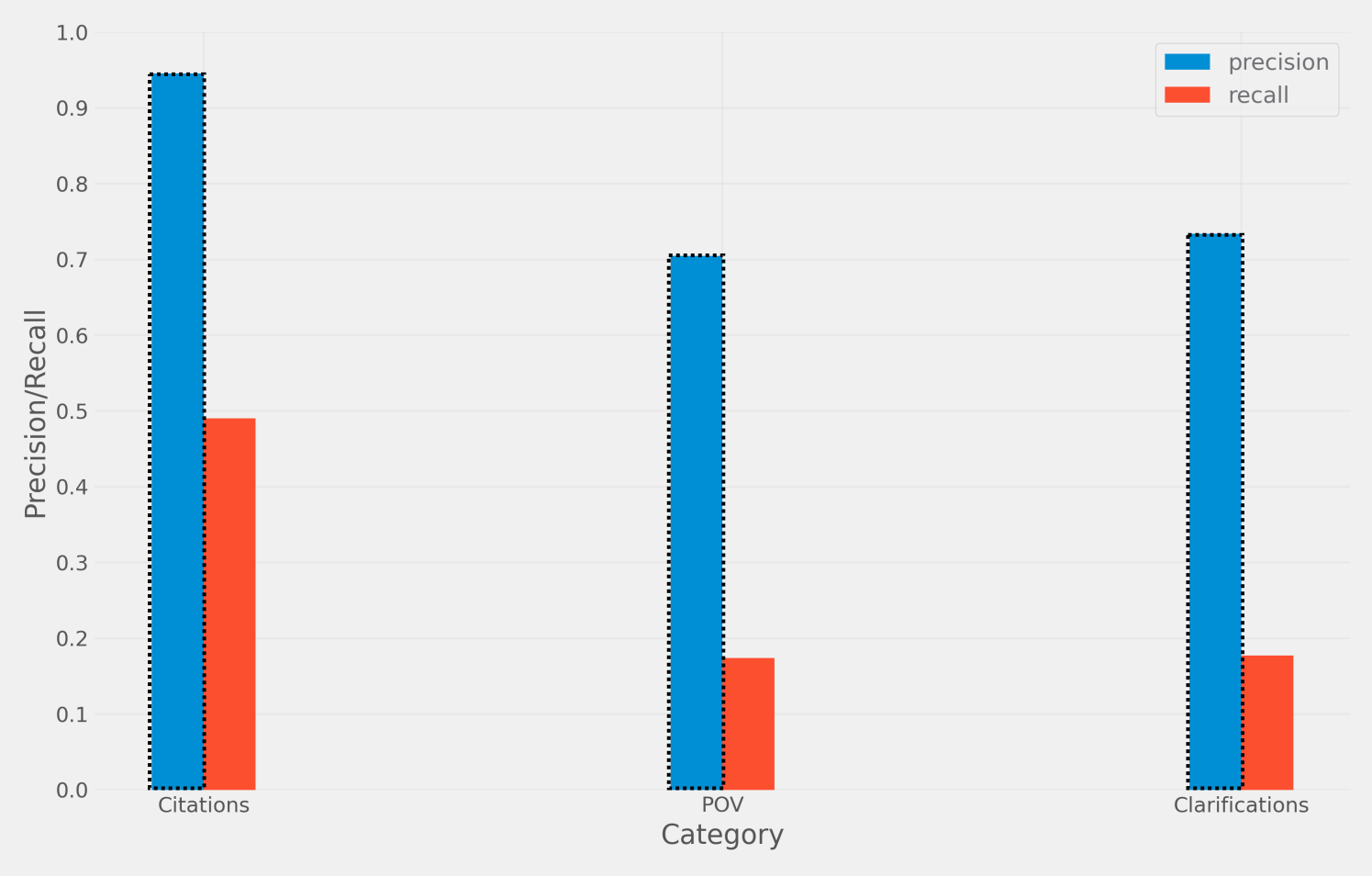
+

```
'''Lateral flow tests'''<ref>
[http://www.devicelink.com/ivdt/archive/99/11/009.html Concurrent
Engineering for Lateral-Flow Diagnostics (IVDT archive, Nov 99)<!-
- Bot generated title -->]</ref> also known as Lateral Flow
[[Immunochromatographic]] Assays are simple devices intended to
detect the presence (or absence) of a target analyte in sample
(matrix) without the need for specialized and costly equipment,
though many lab based applications exist that are supported by
reading equipment.<ref name="Yetisen2013">{{cite journal | author
= Yetisen A. K. | year = 2013 | title = Paper-based microfluidic
point-of-care diagnostic devices | url =
http://pubs.rsc.org/en/content/articlelanding/2013/lc/c3lc50169h |
journal = Lab on a Chip |volume= 13 |issue= 12 |pages= 2210–2251
|doi=10.1039/C3LC50169H}}</ref> Typically, these tests are used
for medical diagnostics either for home testing, [[point of care]]
testing, or laboratory use. A widely spread and well known
application is the home pregnancy test.
```

Evaluation of semantic edit intention labels

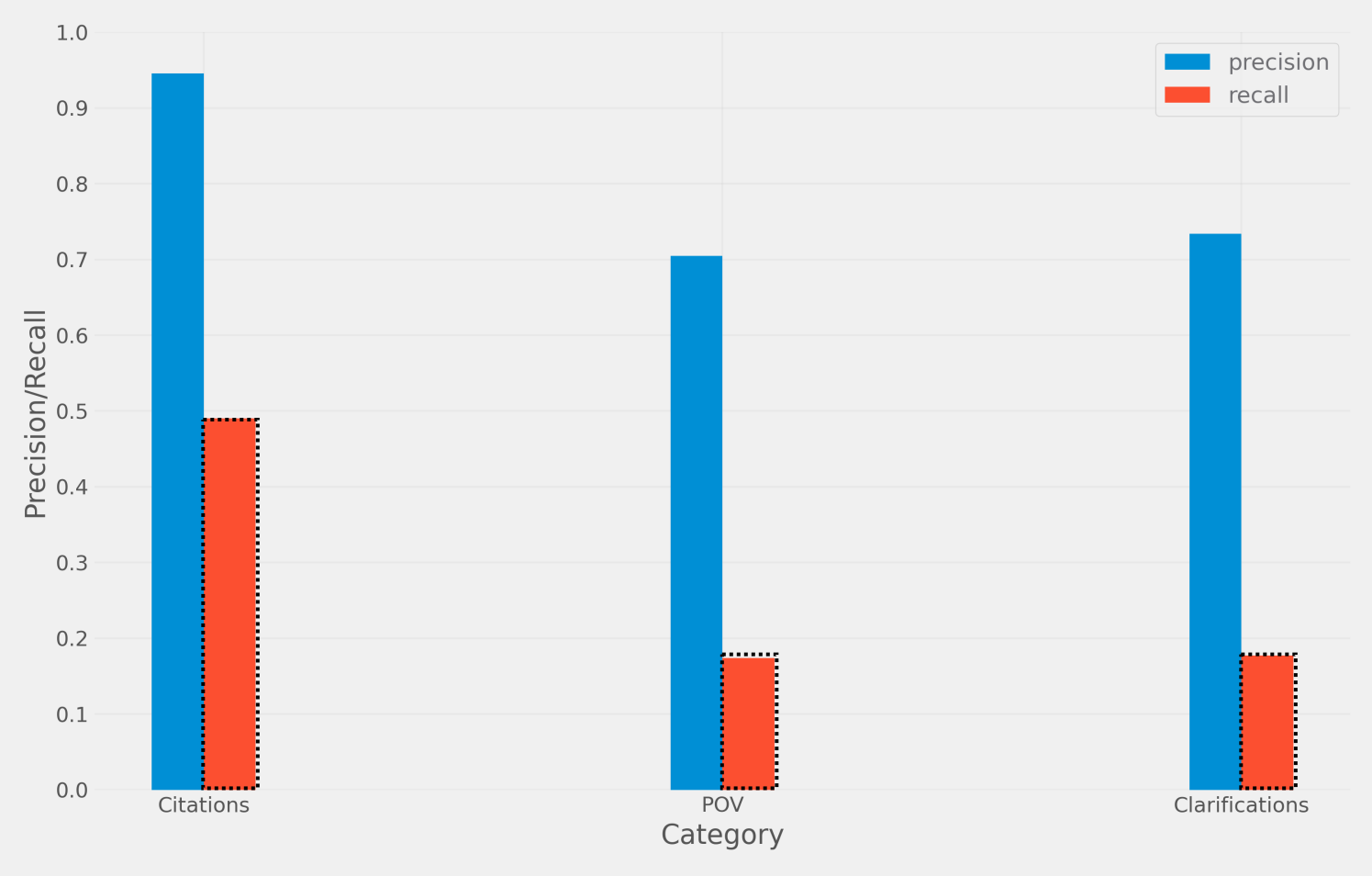


Evaluation of semantic edit intention labels



High precision

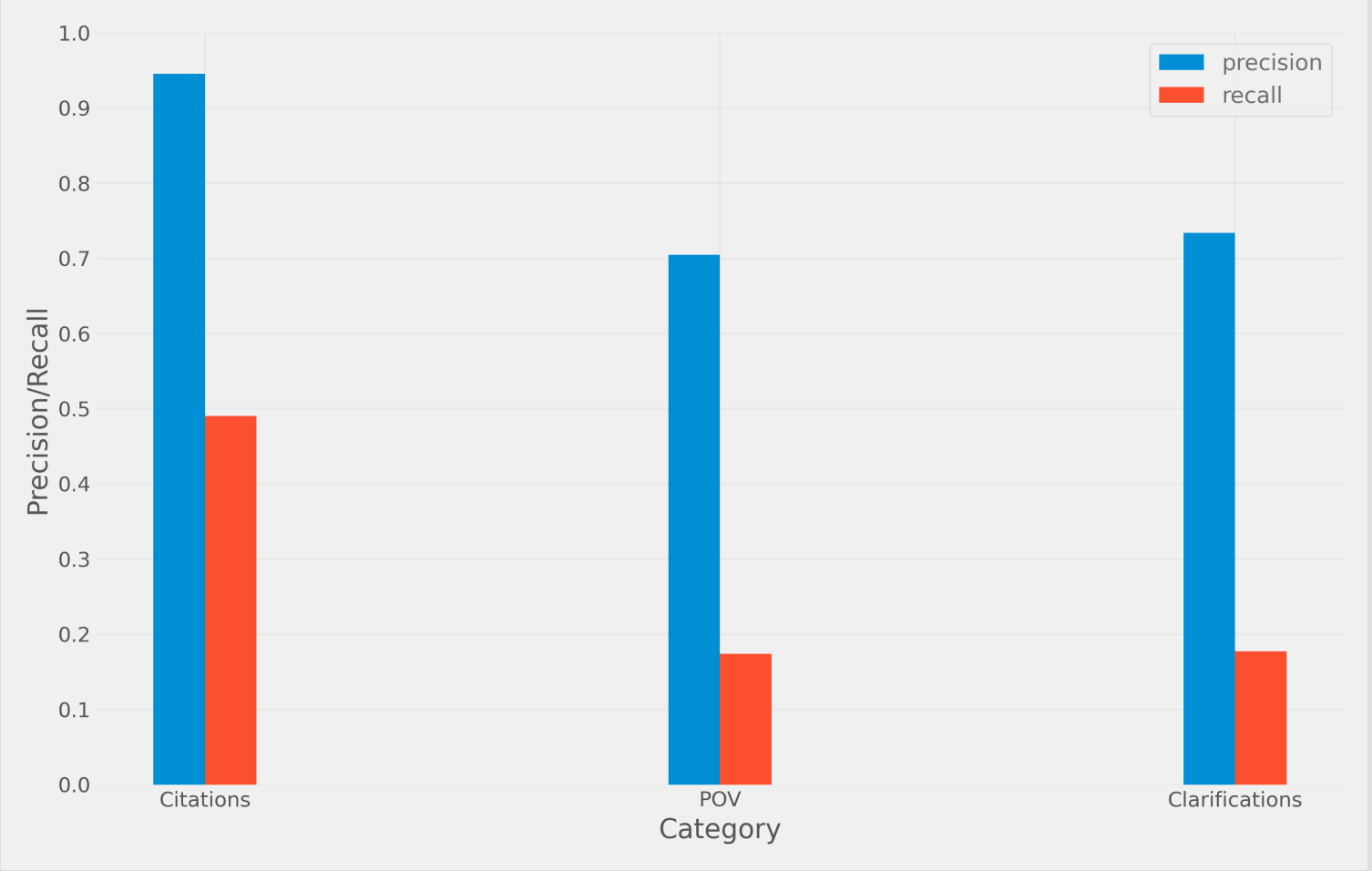
Evaluation of semantic edit intention labels



High precision

Low recall

Evaluation of semantic edit intention labels



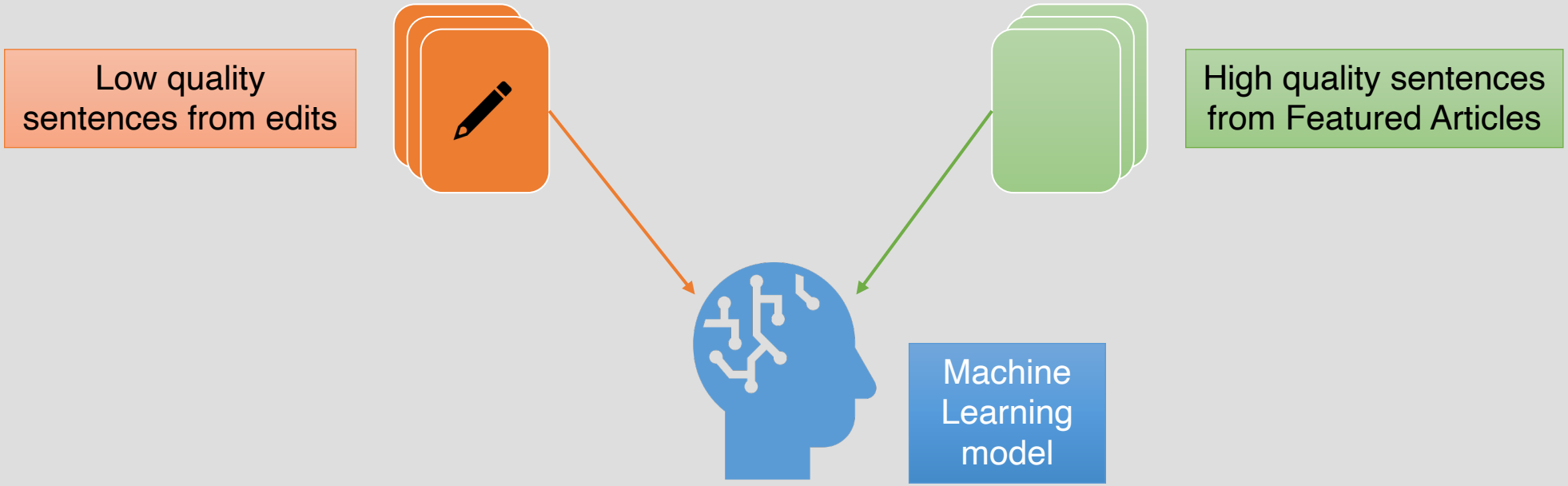
High precision

Low recall

Considerable disagreements

Inter-annotator agreement: 0.62 0.02 0.17

Evaluation of sentence quality labels



Needs Citations

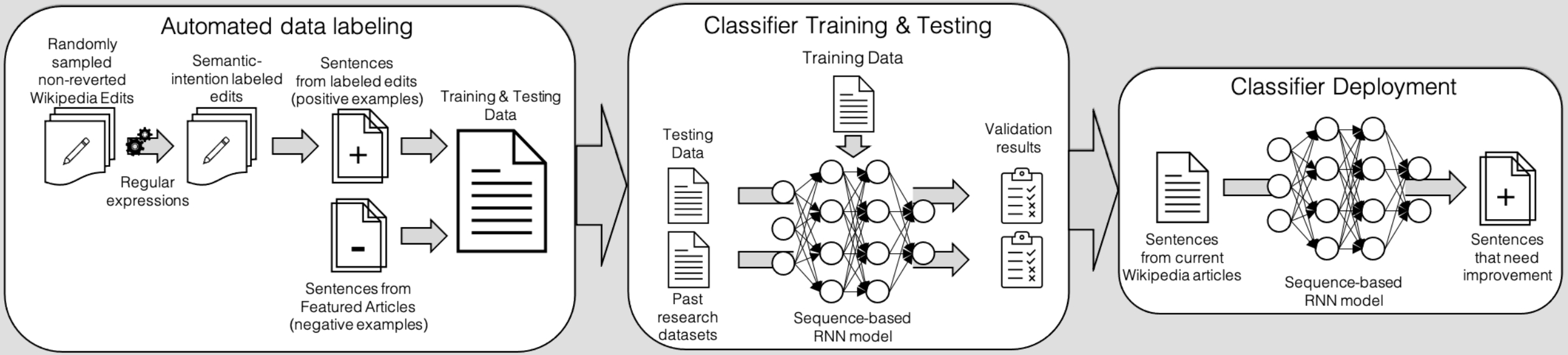
1946 : ford sues the allies for damages done to his factories in dresden during the infamous bombing and wins compensation

Has point-of-view

sharky and george was a very popular humorous show about two fish private detective.

Implications of learning from Wikipedia editing behaviors

Implications of automated labeling of sentence quality



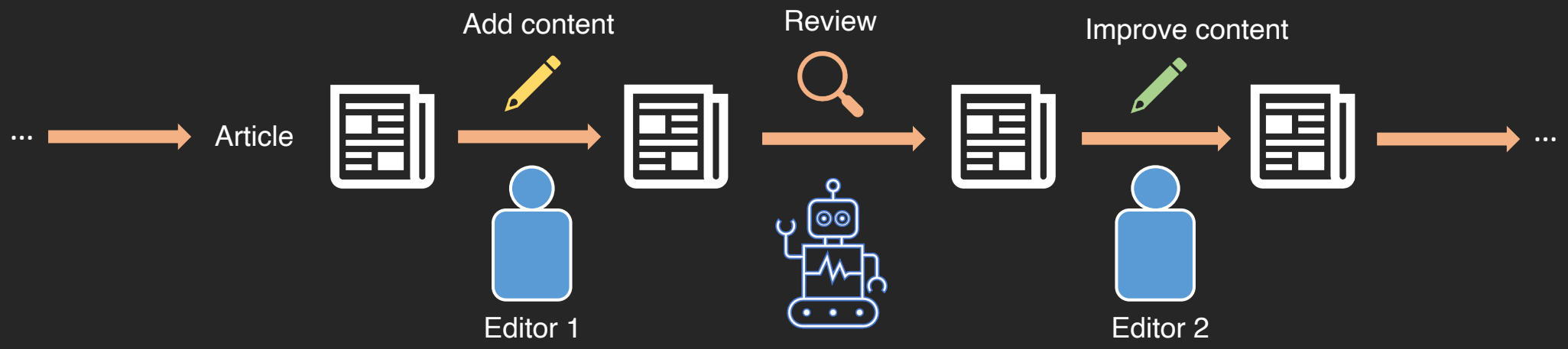
Fast

Flexible

Language independent

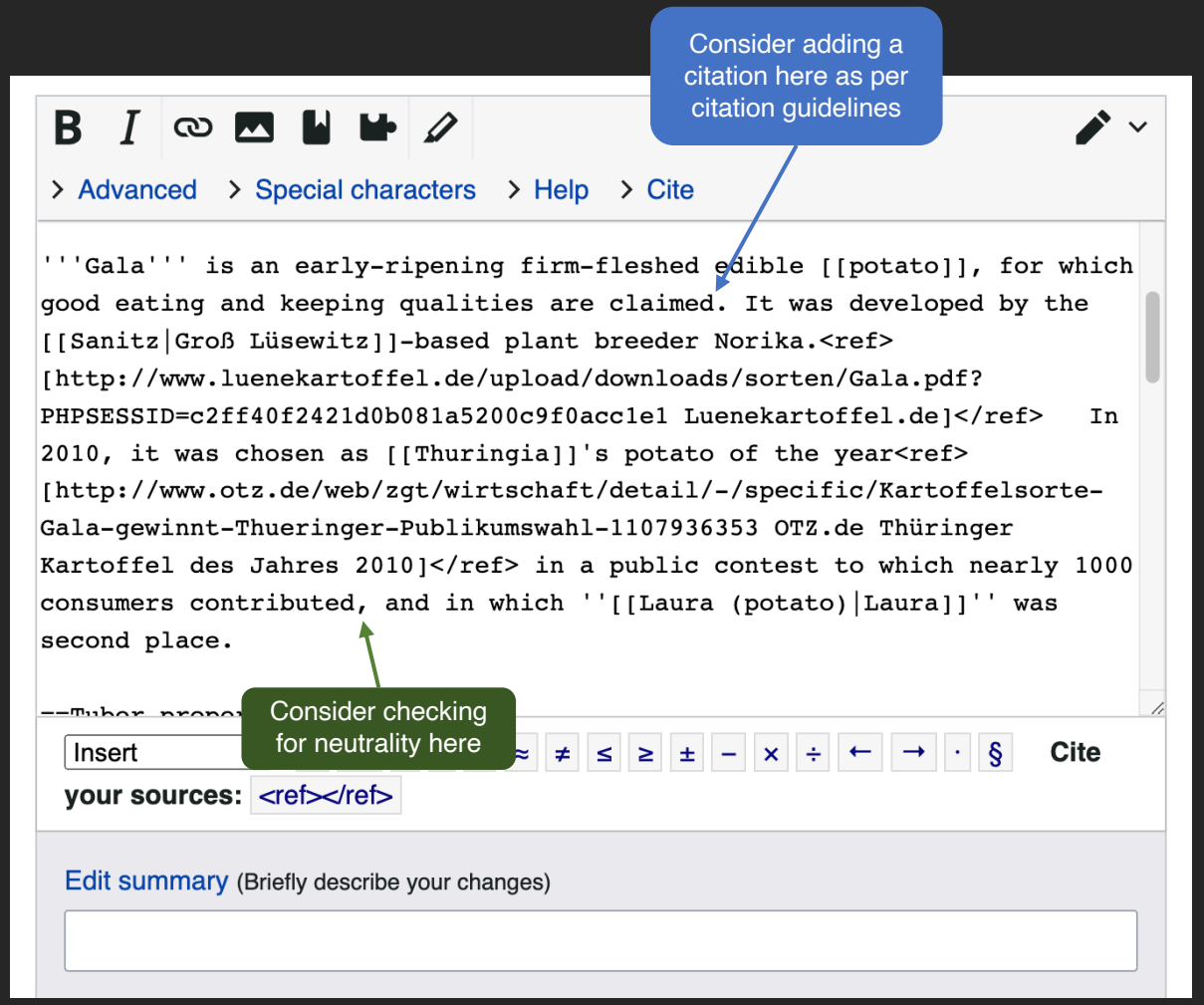
Article improvement – add, assess, improve

Article improvement – add, assess, improve

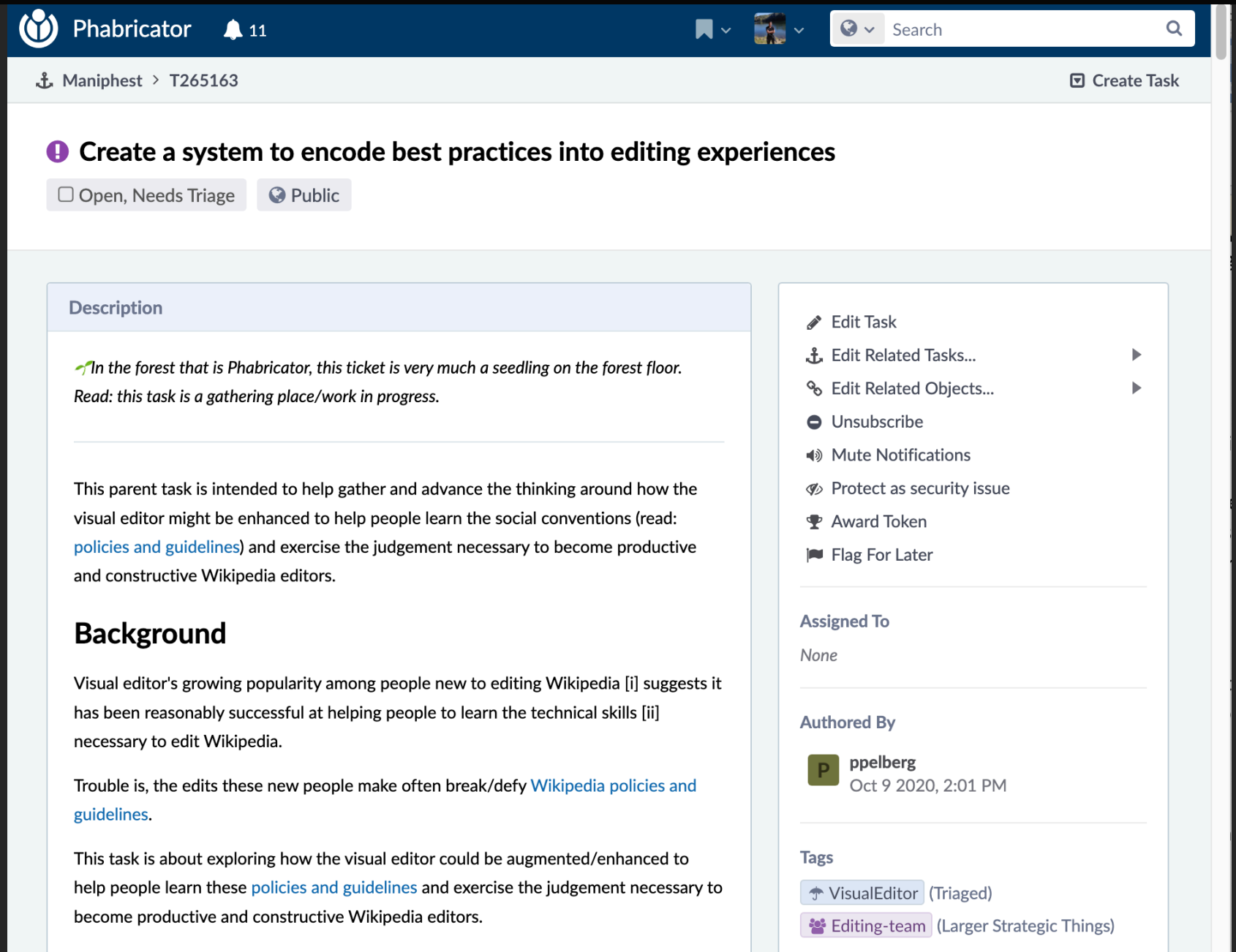


Assisting new editors on common practices

- Recommending best practice changes to new editors.
- Can prevent follow-up edits and save other editor's valuable time!



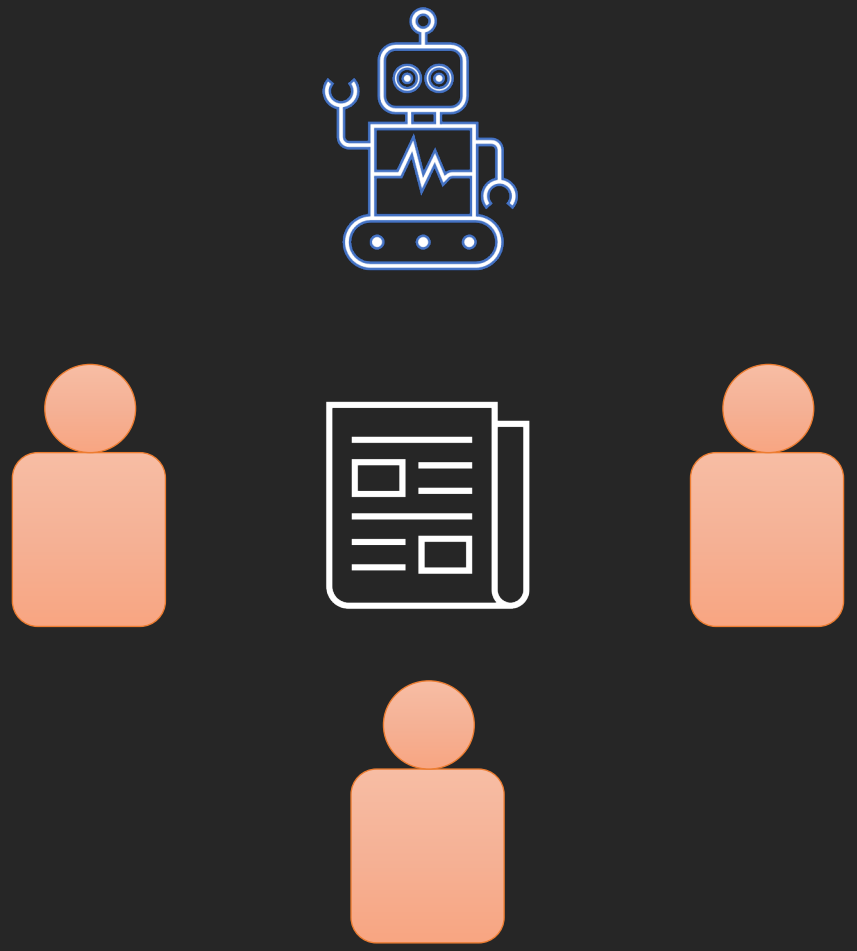
Assisting new editors on common practices



Inherent Ambiguity

- Some semantic categories can have multiple viewpoints.
- E.g., Neutral point-of-view is a highly debatable policy
- Models need to incorporate such ambiguity.

How can AI and humans work together in a task? Whose voice matters?



Gordon, Mitchell L., et al. "Jury Learning: Integrating Dissenting Voices into Machine Learning Models." arXiv preprint arXiv:2202.02950 (2022).

Technical probe of Wikipedia norms

- Models trained on implicit editing behaviors capture descriptive behaviors (how editors interpret and operationalize Wikipedia norms and practices).

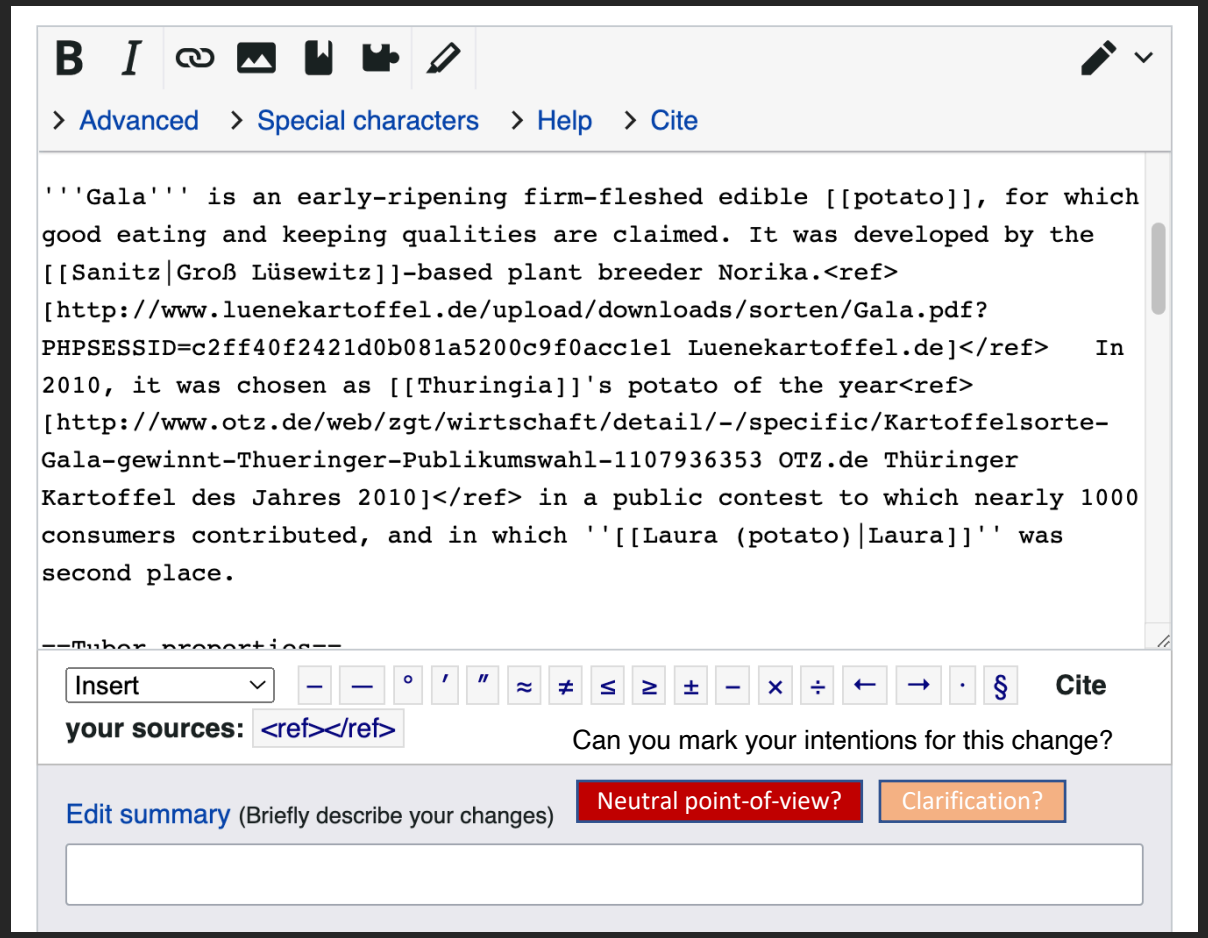
- Thus, these models can be used to understand the application of a policy in context (e.g., looking at the prediction of new sentences to understand the policy better)

Technical probe of Wikipedia norms

“He had a very jovial personality”

Prediction: ”biased” (91%)

What do we need for better AI?



More expressive UI

Better data

Better AI

Thank You!

Sumit Asthana
asumit@umich.edu
http://sumitasthana.xyz