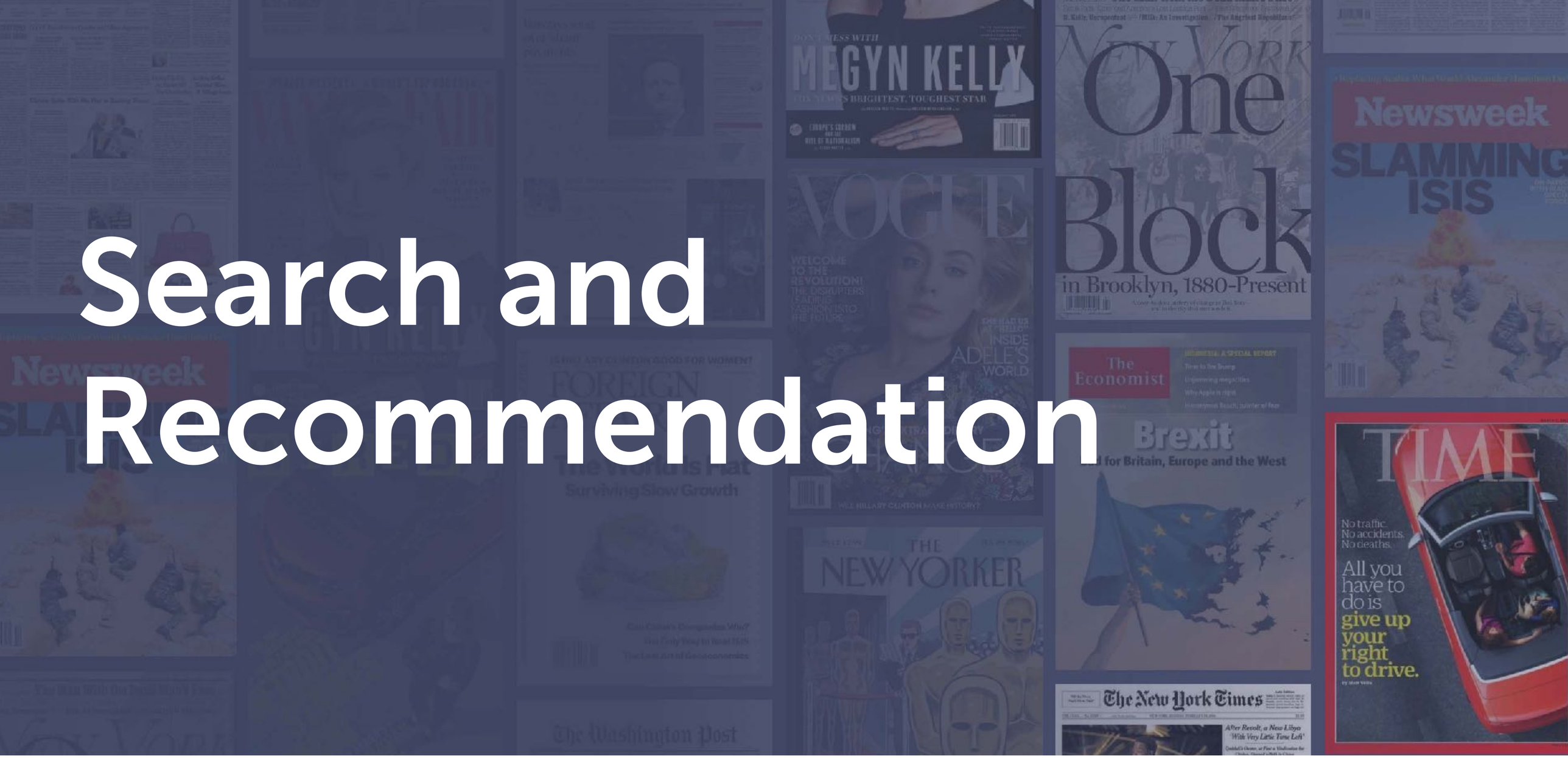
Search and Recommendation



Daan Odijk & Anne Schuth
data scientists

Blendle

Plan today

- **Search Engines that Learn from their Users**

Plan today

- **Search Engines that Learn from their Users**
- **Struggling and Success in Web Search**

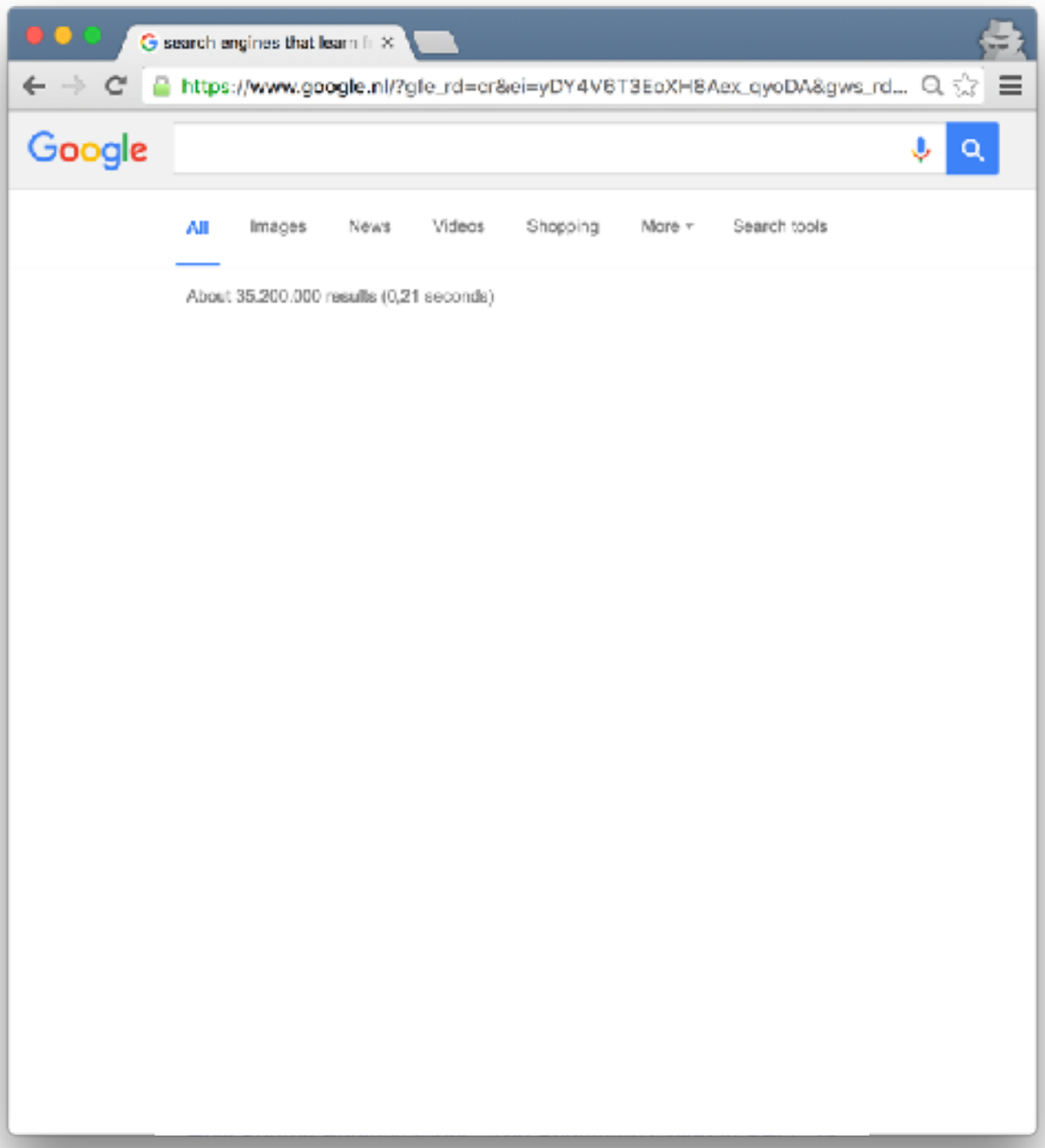
Plan today

- **Search Engines that Learn from their Users**
- **Struggling and Success in Web Search**
- **Real-time recommendations at Blendle**
  - Article- and User profile enrichment
  - Streaming Architecture *(Docker, Kubernetes and Kafka)*
  - Demo

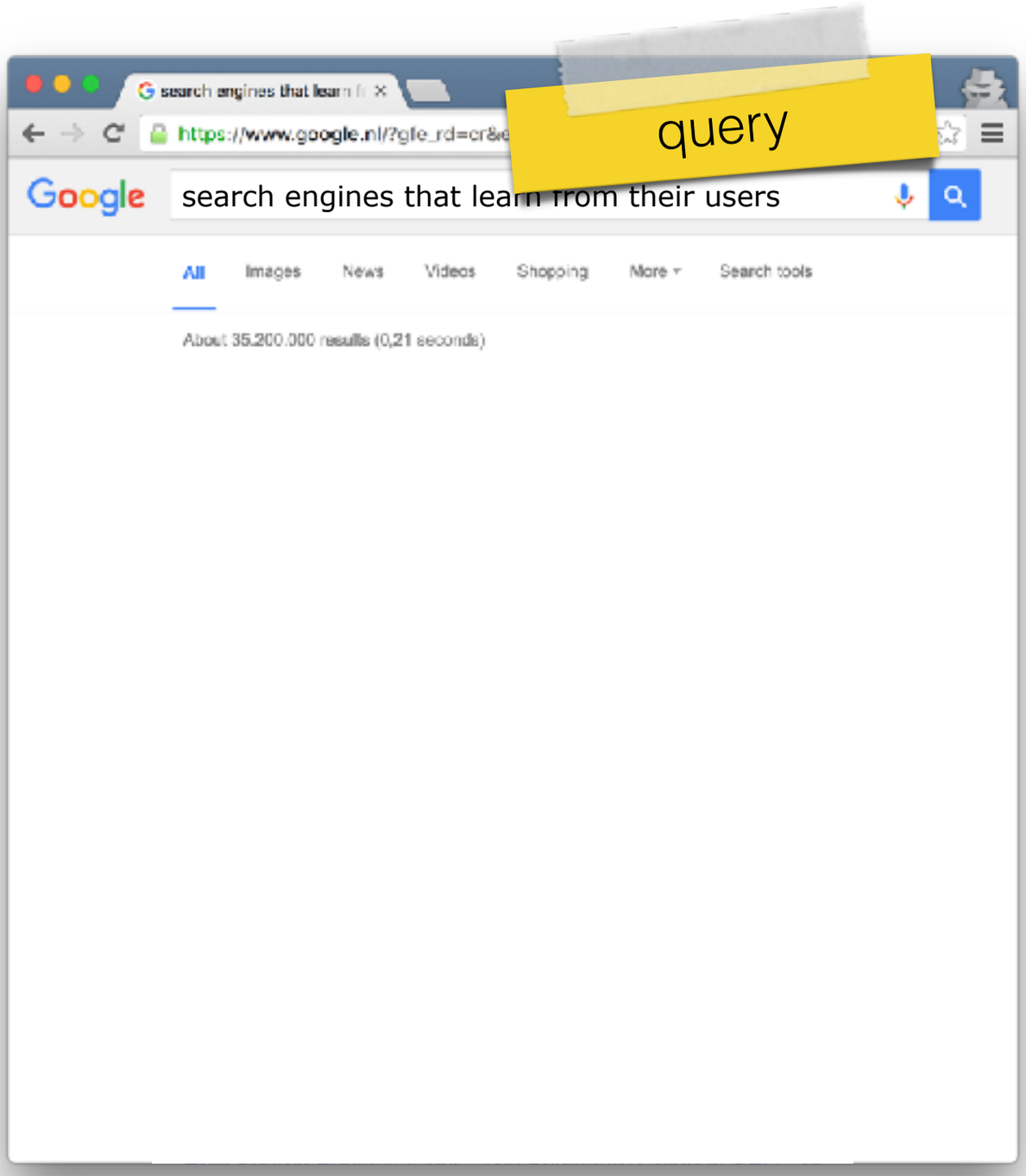
SEARCH ENGINES THAT LEARN FROM THEIR USERS

ANNE SCHUTH

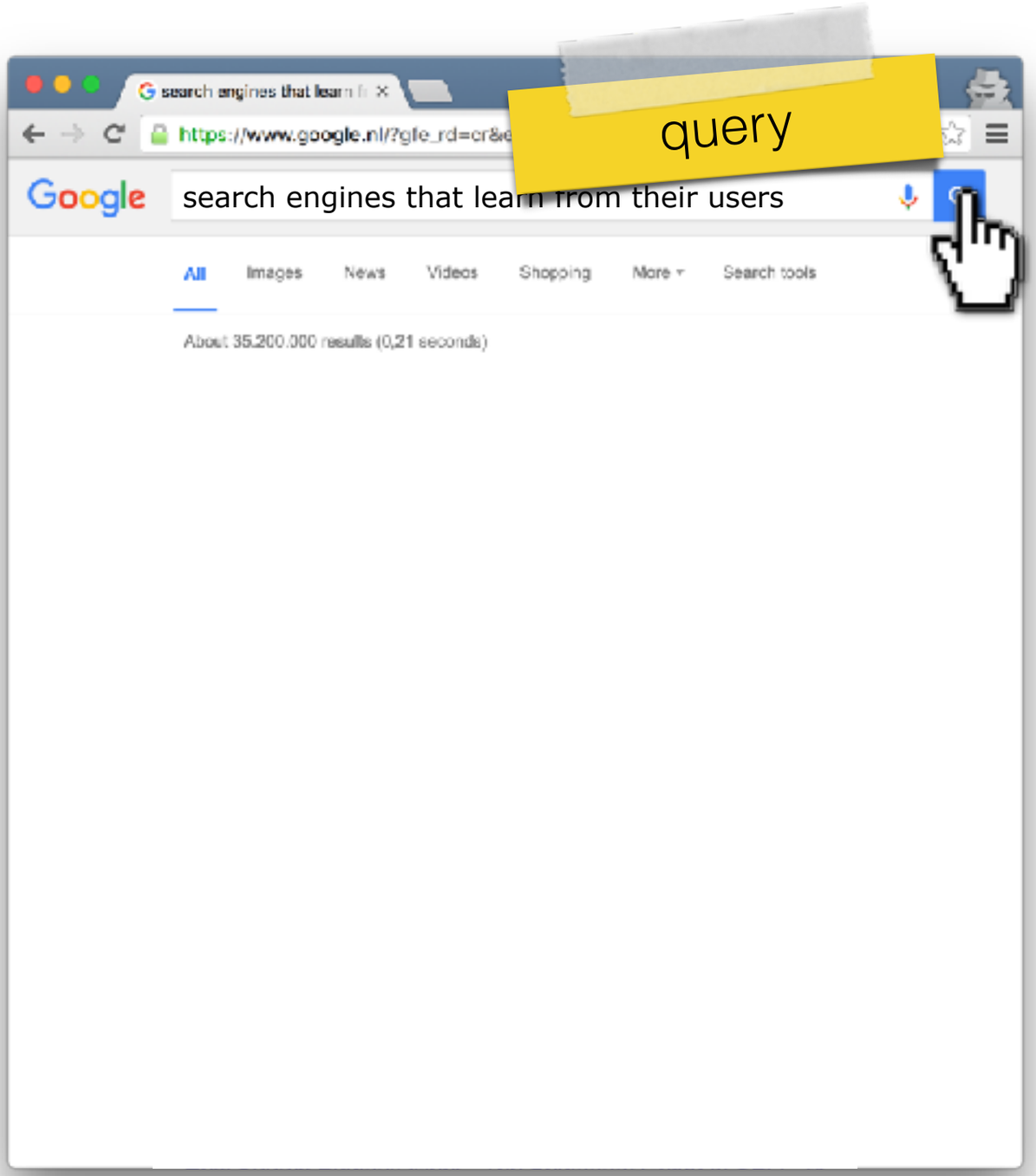
Search Engines



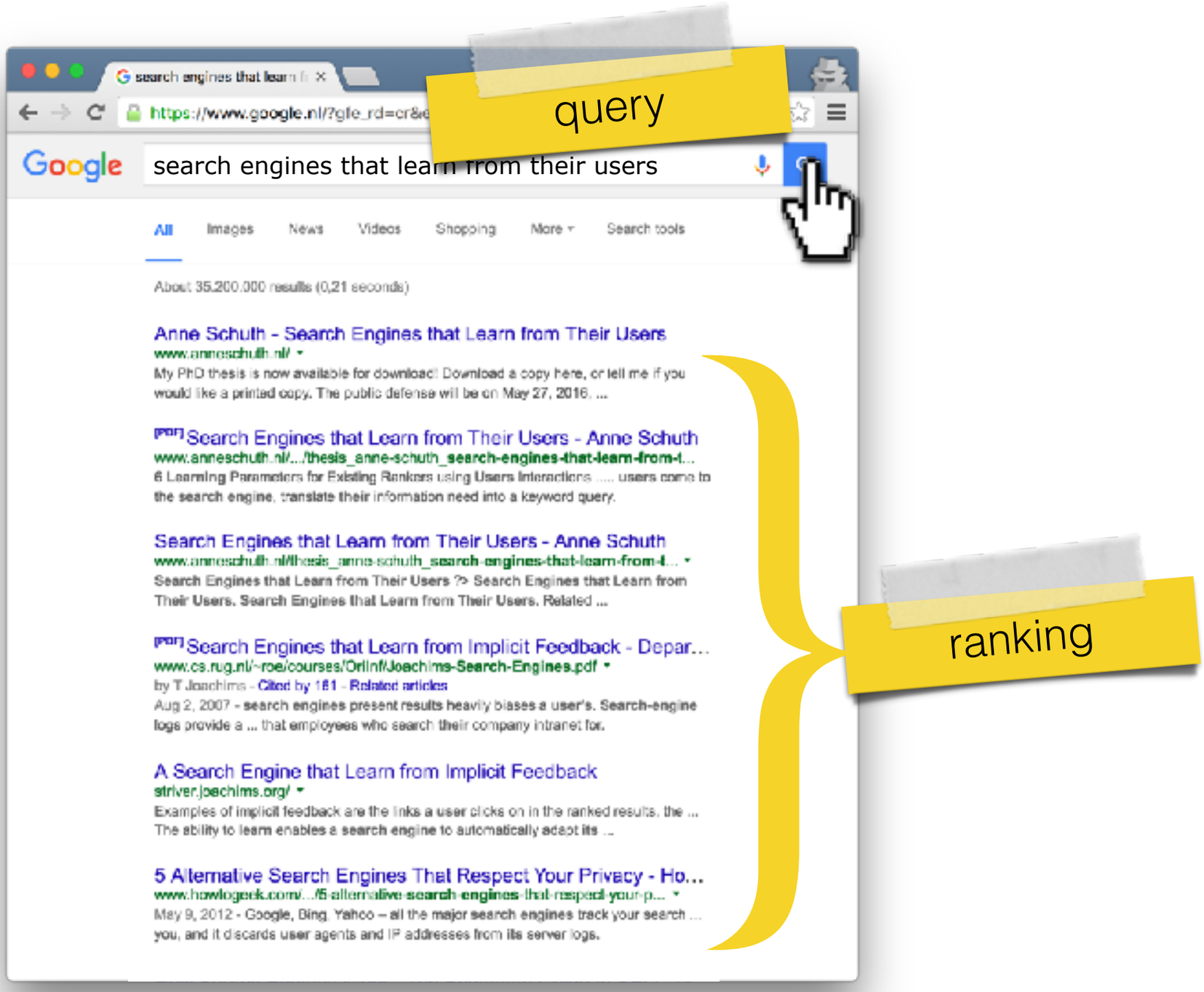
SEARCH ENGINES



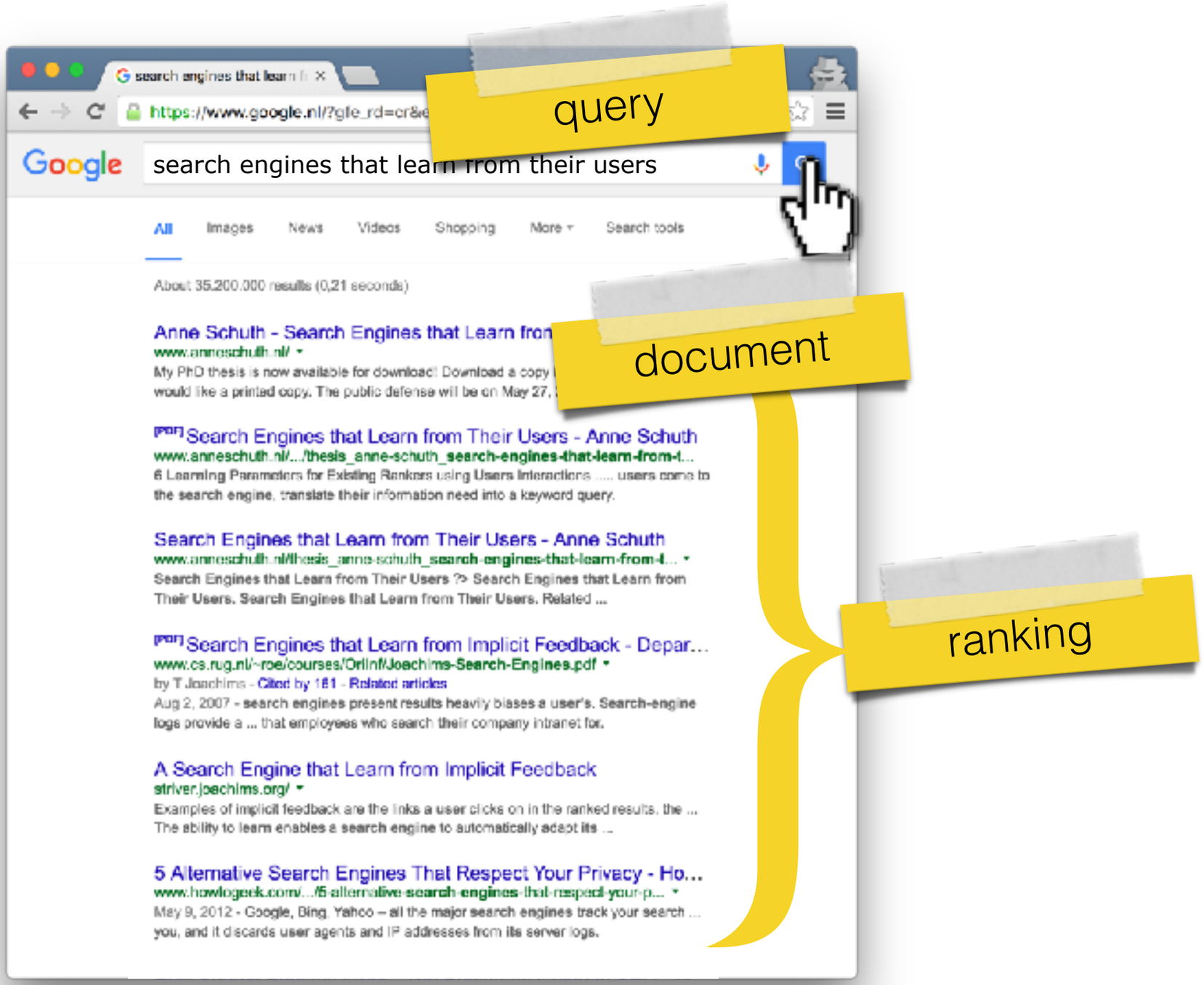
SEARCH ENGINES



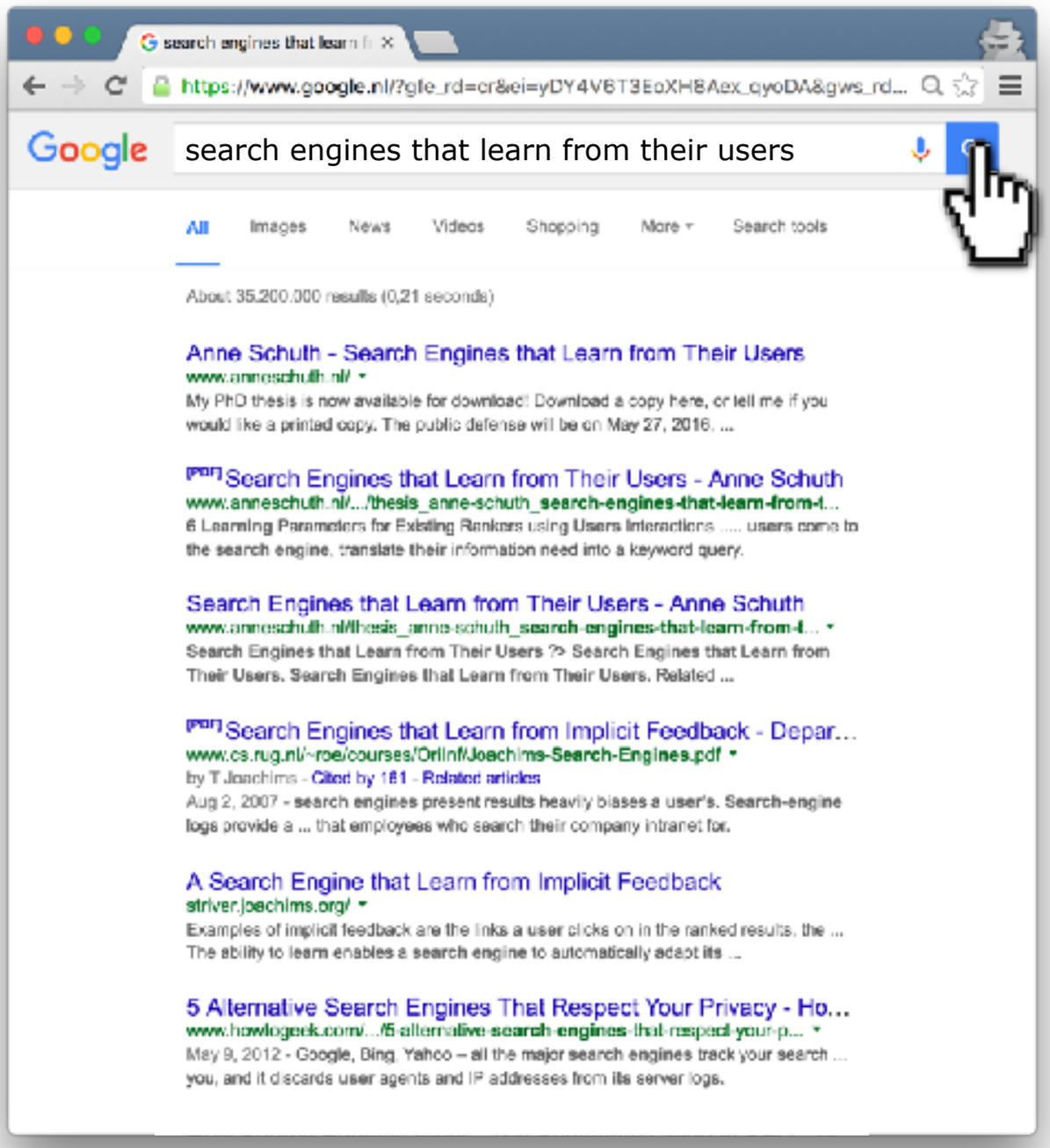
SEARCH ENGINES



Search Engines



RANKING


search engines that learn fr
https://www.google.nl/?gfe_rd=cr&ei=yDY4V6T3EoXH8Aex_oyoDA&gws_rd...
Google
search engines that learn from their users
All
Images
News
Videos
Shopping
More
Search tools
About 35.200.000 results (0,21 seconds)
Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...
[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.
Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...
[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.
A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...
5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

RANKING

search engines that learn from their users

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results, the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

RANKING

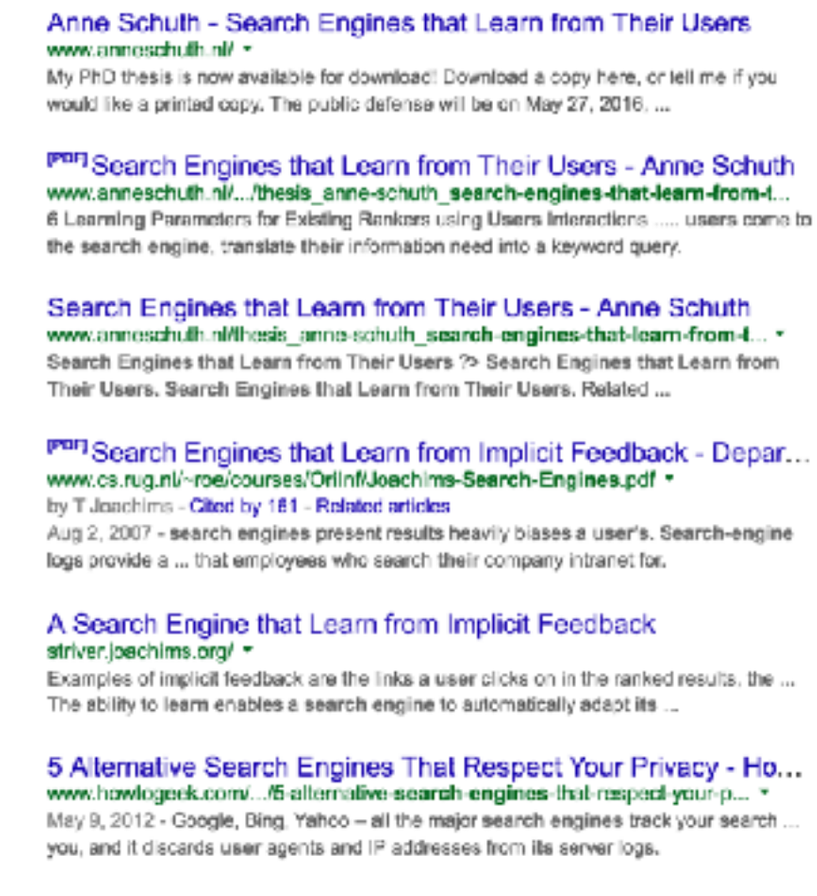
search engines that learn from their users

feature

	#query terms
Anne Schuth - Search Engines that Learn from Their Users www.anneschuth.nl/ My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...	7
[PDF] Search Engines that Learn from Their Users - Anne Schuth www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t... 6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.	7
Search Engines that Learn from Their Users - Anne Schuth www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...	7
[PDF] Search Engines that Learn from Implicit Feedback - Depar... www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf by T Joachims - Cited by 181 - Related articles Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.	7
A Search Engine that Learn from Implicit Feedback striver.joachims.org/ Examples of implicit feedback are the links a user clicks on in the ranked results, the ... The ability to learn enables a search engine to automatically adapt its ...	7
5 Alternative Search Engines That Respect Your Privacy - Ho... www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.	7

RANKING

search engines that learn from their users

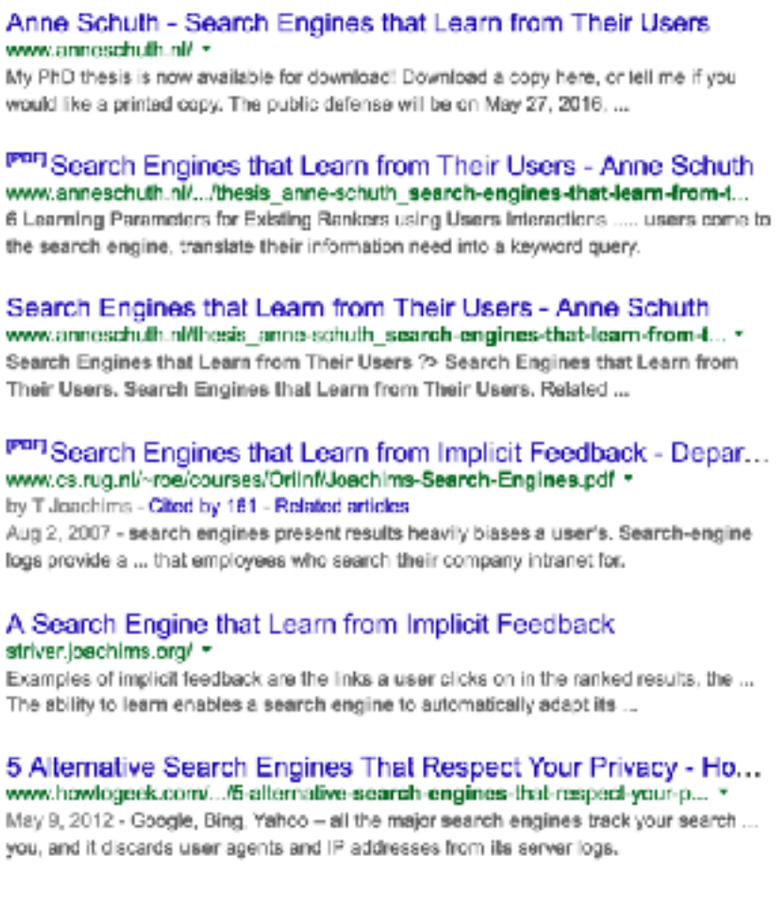


feature

#query terms	#document terms
7	500
7	80000
7	80000
7	5000
7	2000
7	1000

RANKING

search engines that learn from their users

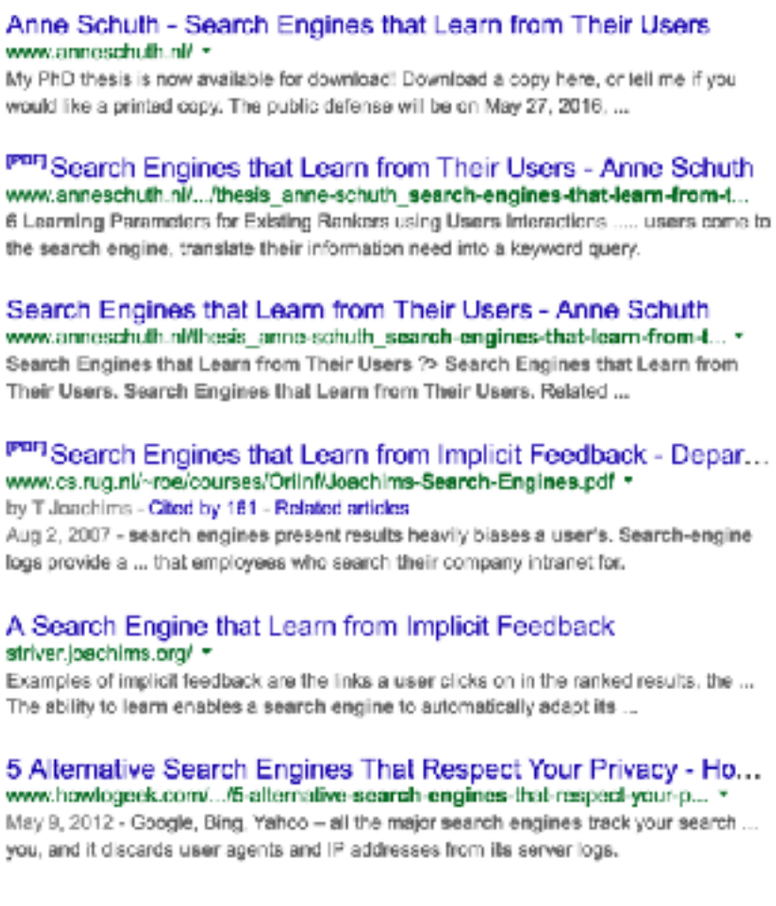


feature

#query terms	#document terms	pagerank
7	500	0.2
7	80000	0.2
7	80000	0.2
7	5000	0.5
7	2000	0.4
7	1000	0.6

RANKING

search engines that learn from their users

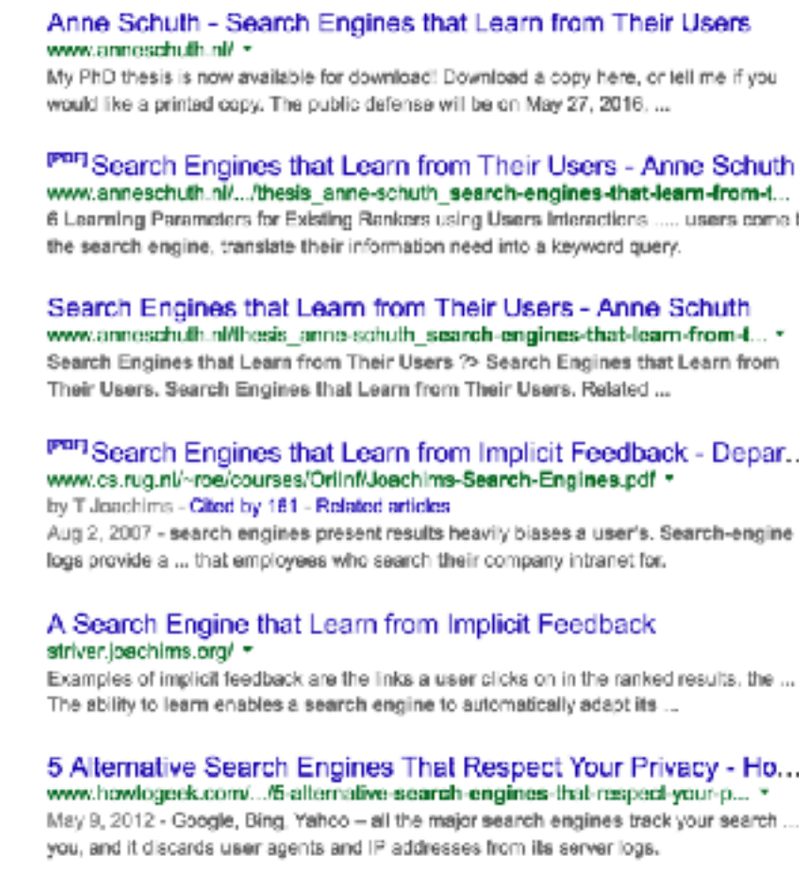


feature

#query terms	#document terms	pagerank	#query terms in document
7	500	0.2	7
7	80000	0.2	7
7	80000	0.2	7
7	5000	0.5	7
7	2000	0.4	7
7	1000	0.6	6

RANKING

search engines that learn from their users

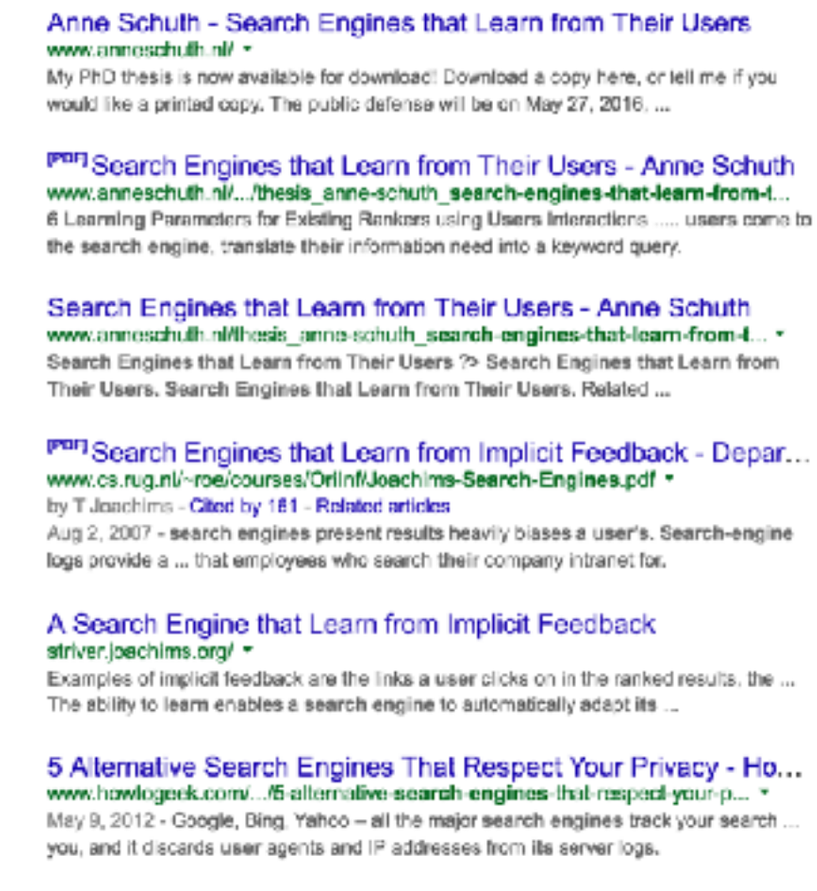


feature

#query terms	#document terms	pagerank	#query terms in document	#query terms in title
7	500	0.2	7	7
7	80000	0.2	7	7
7	80000	0.2	7	7
7	5000	0.5	7	5
7	2000	0.4	7	5
7	1000	0.6	6	3

RANKING

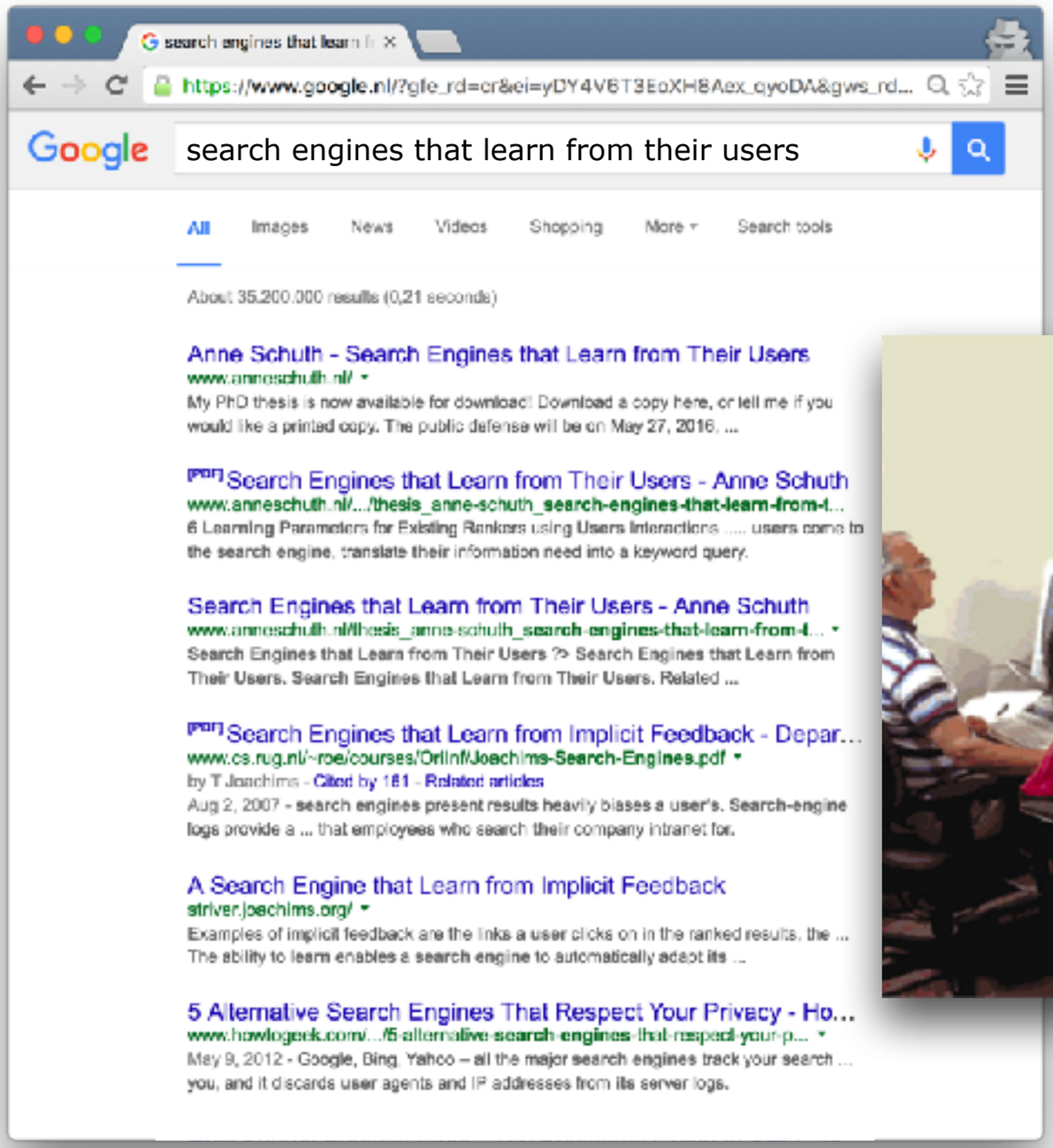
search engines that learn from their users



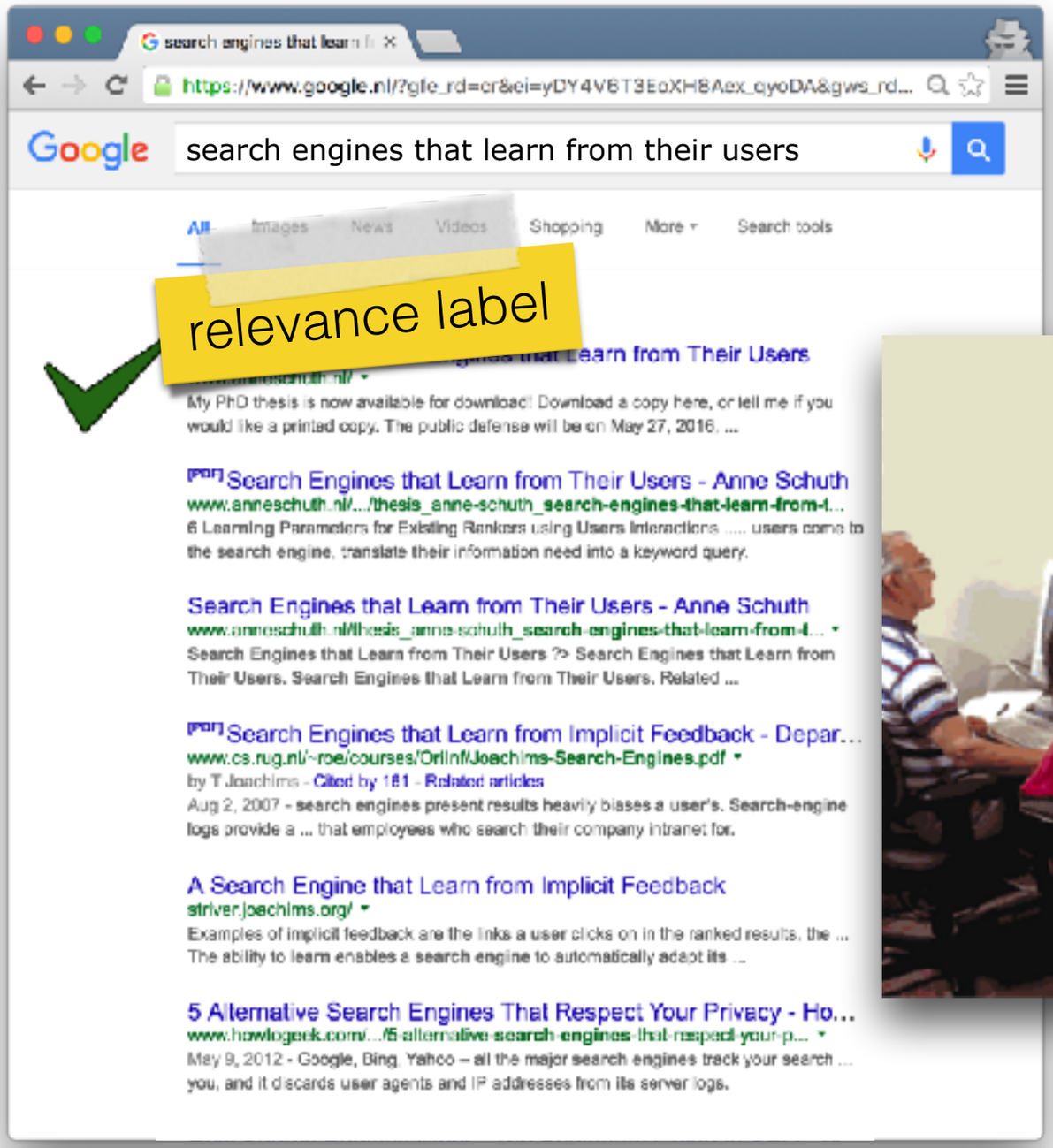
feature

#query terms	#document terms	pagerank	#query terms in document	#query terms in title	...
7	500	0.2	7	7	
7	80000	0.2	7	7	
7	80000	0.2	7	7	
7	5000	0.5	7	5	
7	2000	0.4	7	5	
7	1000	0.6	6	3	

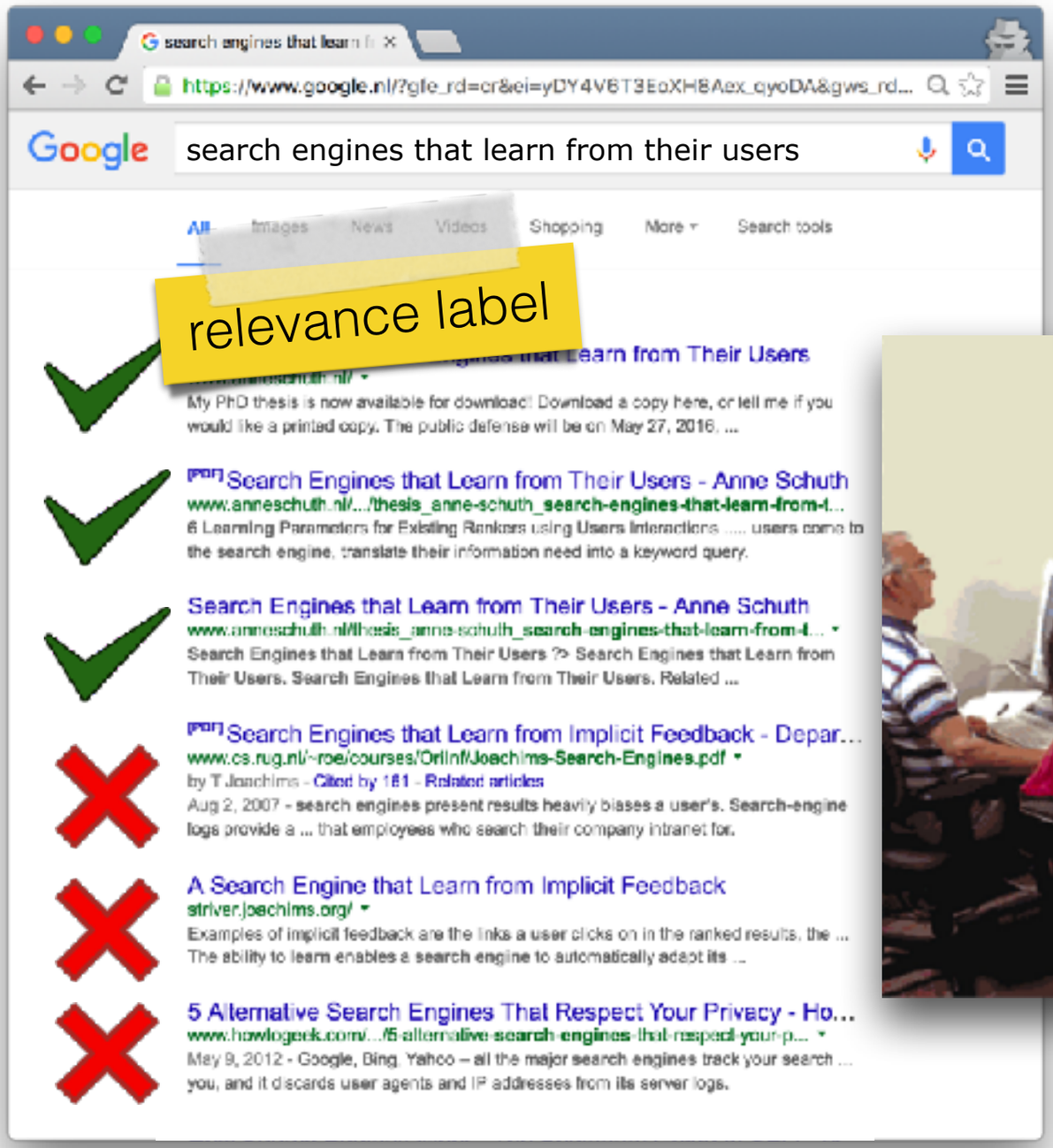
EVALUATION OF SEARCH



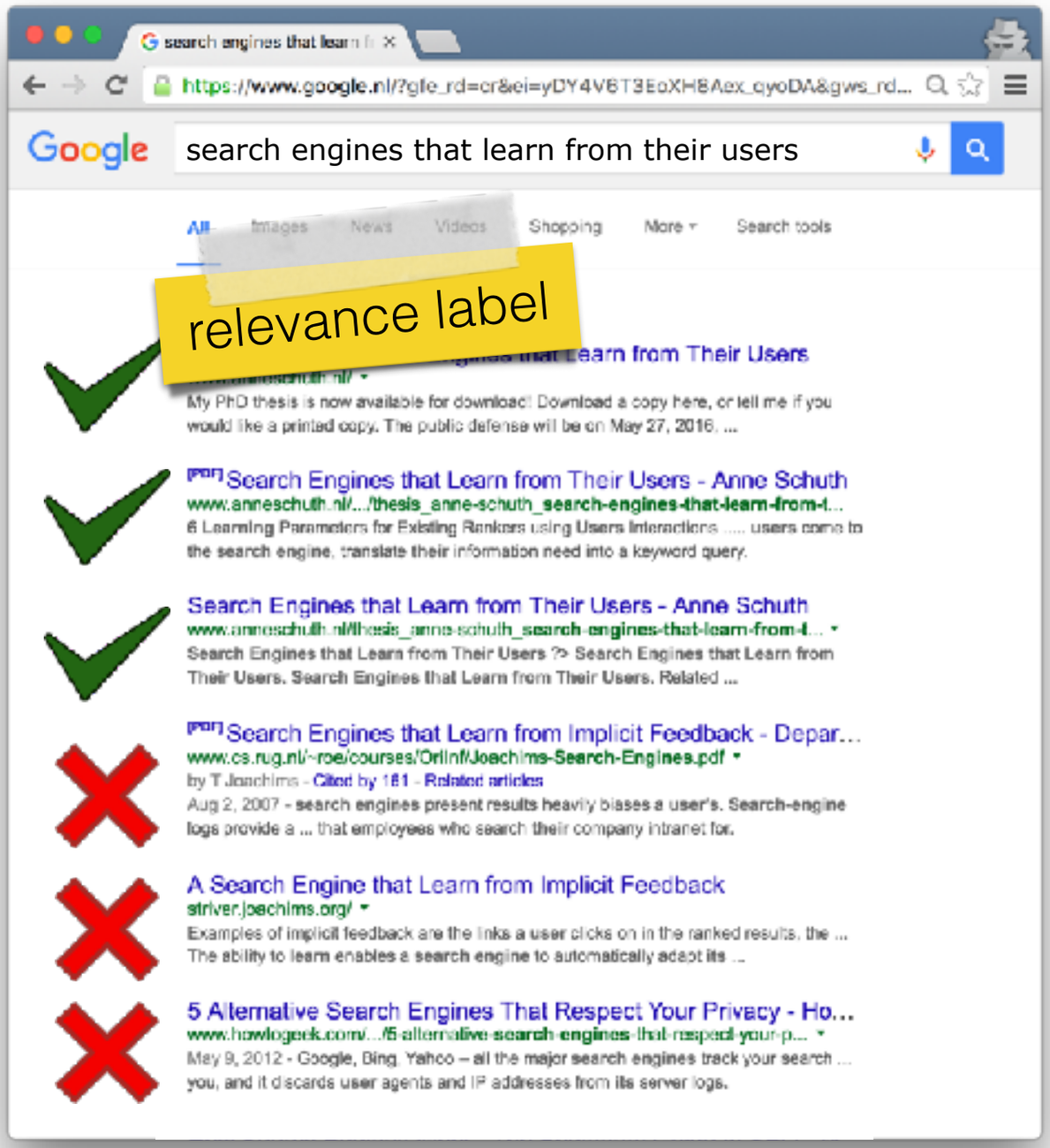
EVALUATION OF SEARCH



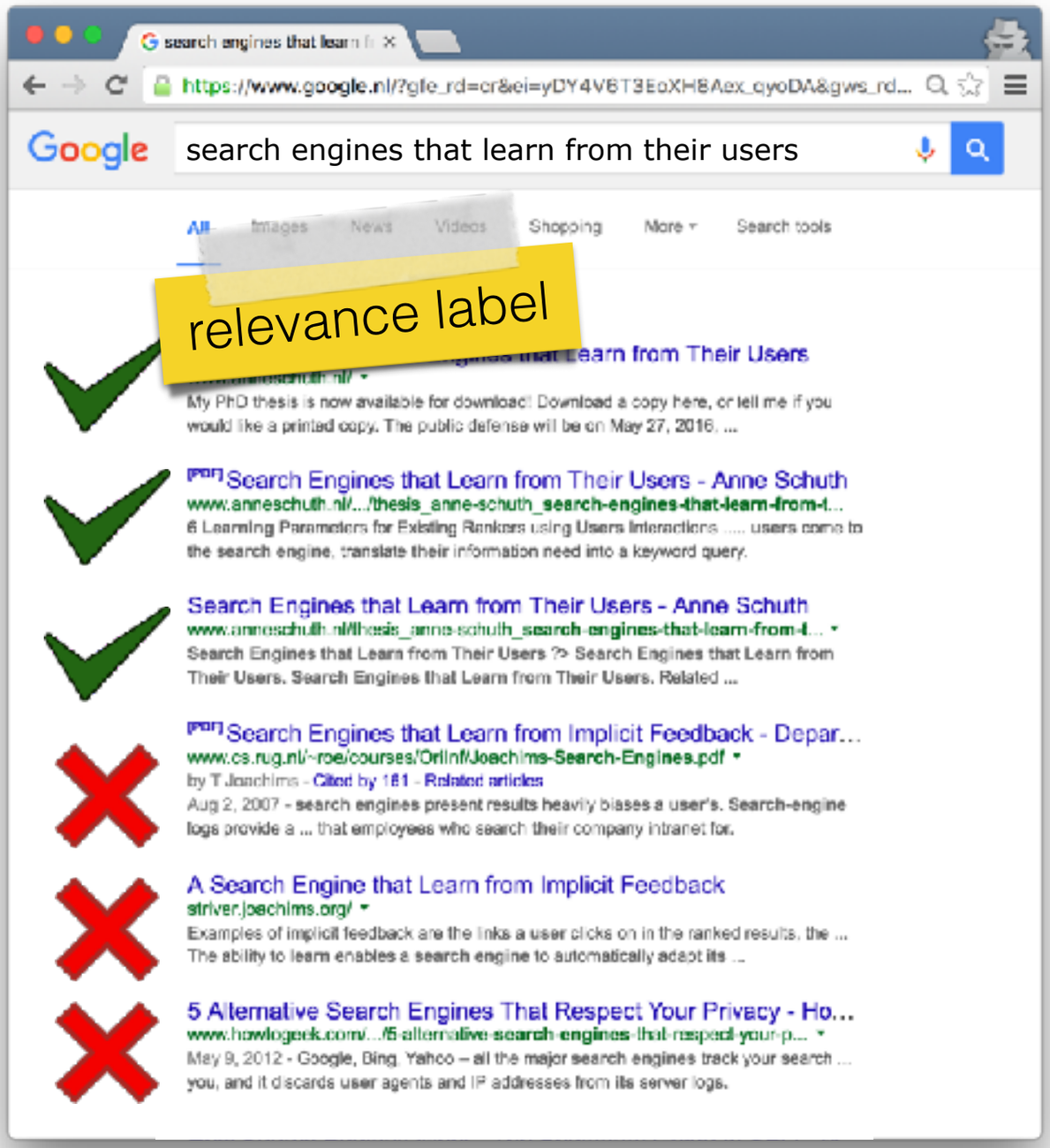
EVALUATION OF SEARCH

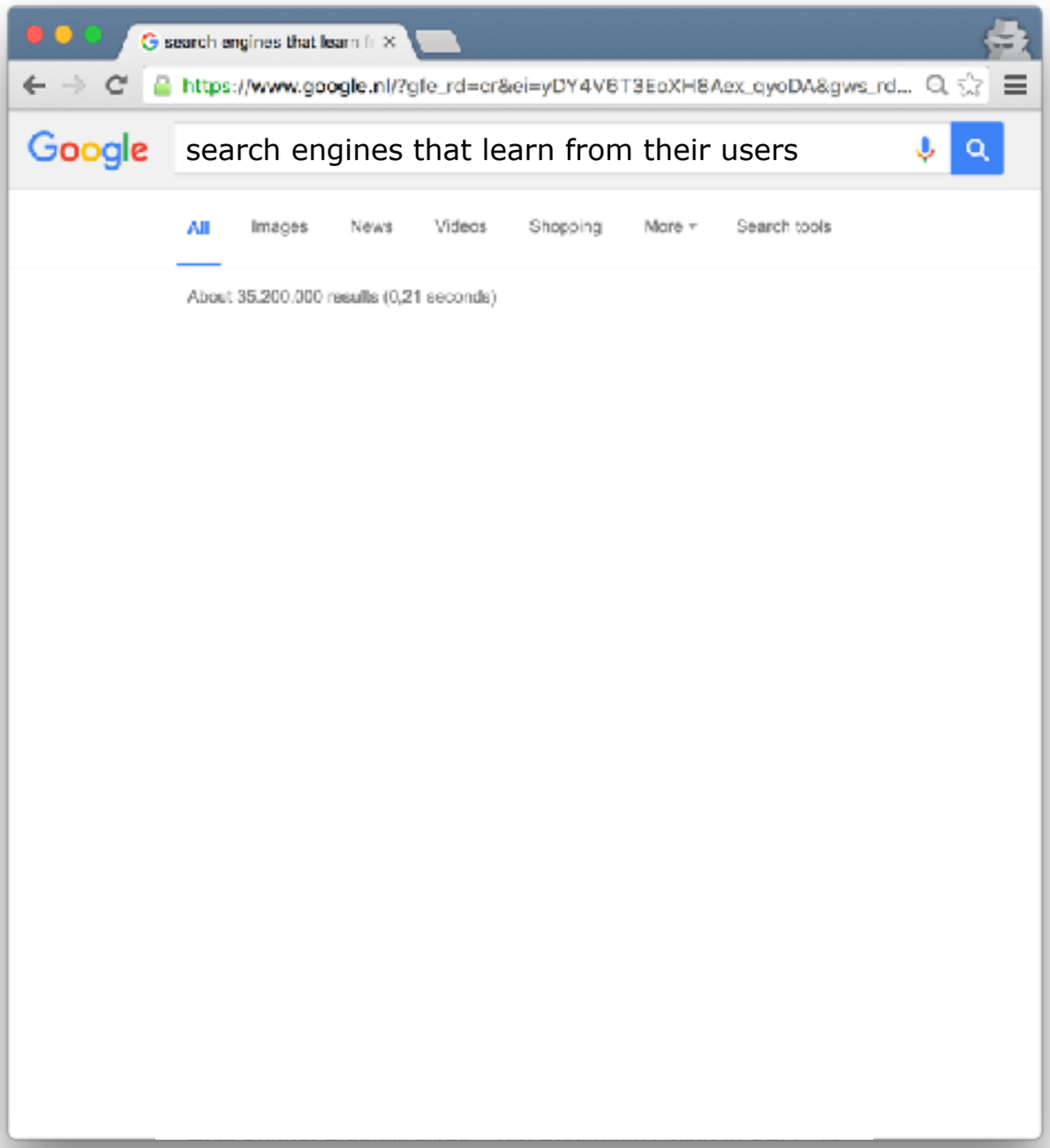


EVALUATION OF SEARCH

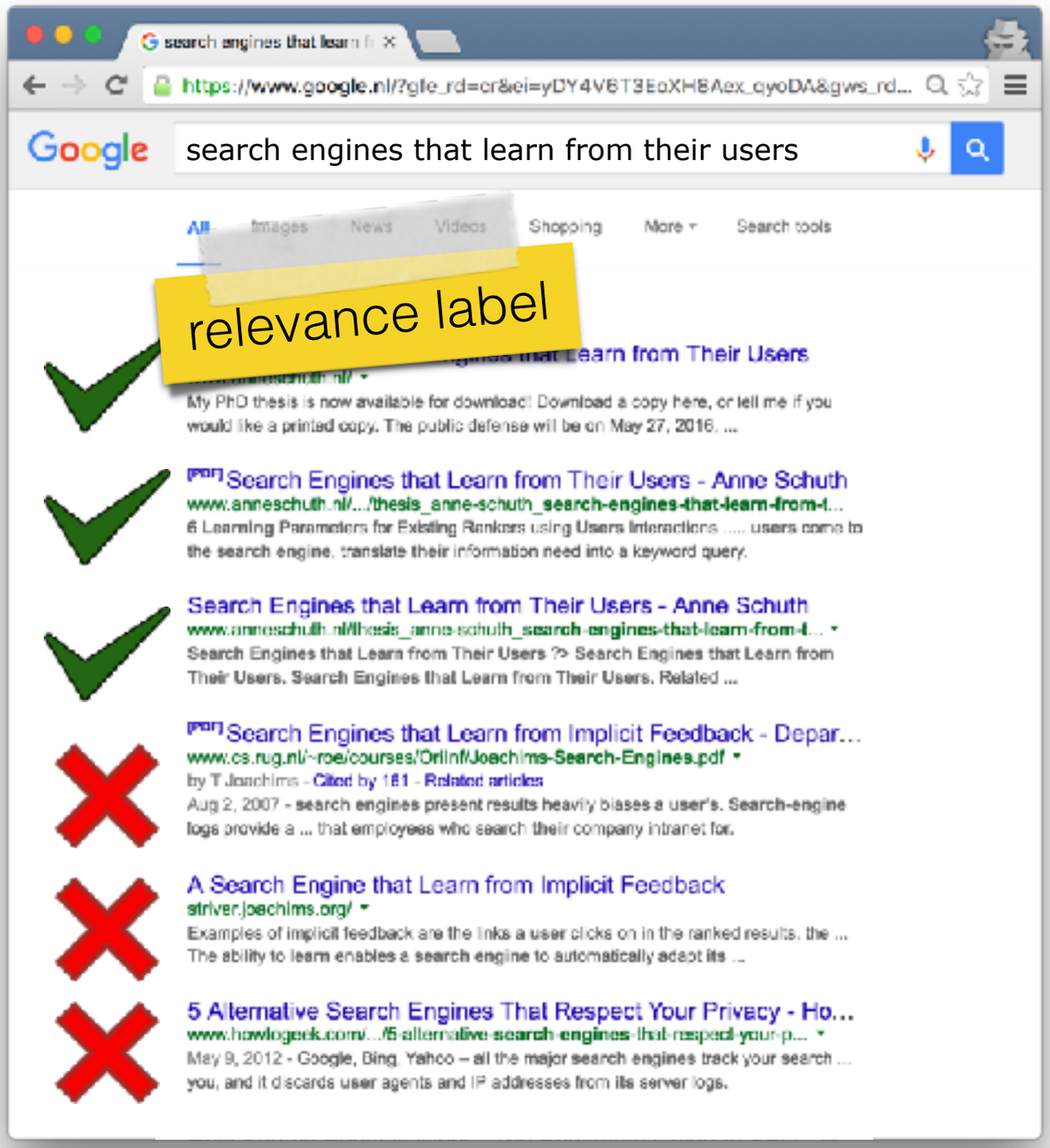


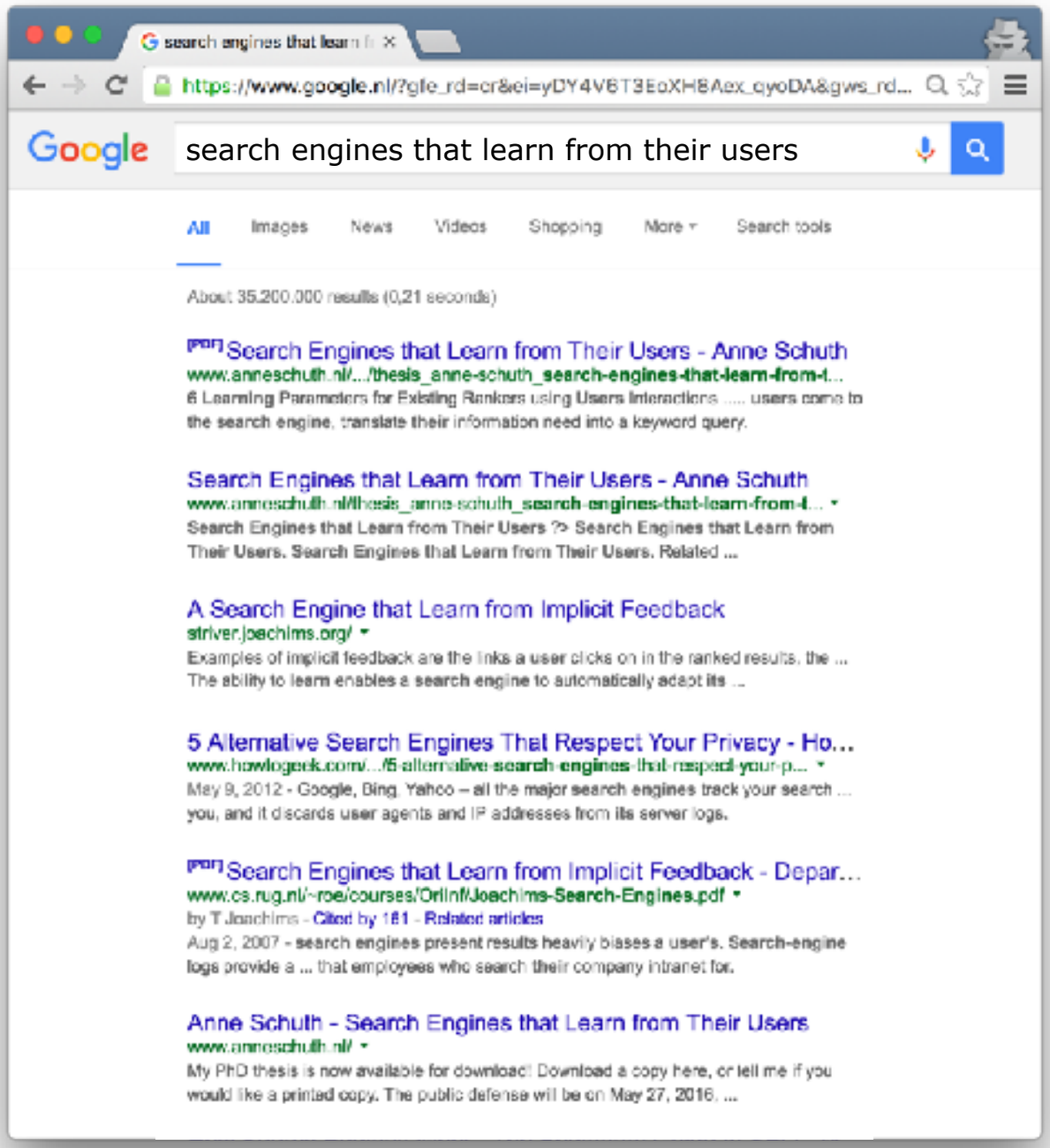
EVALUATION OF SEARCH



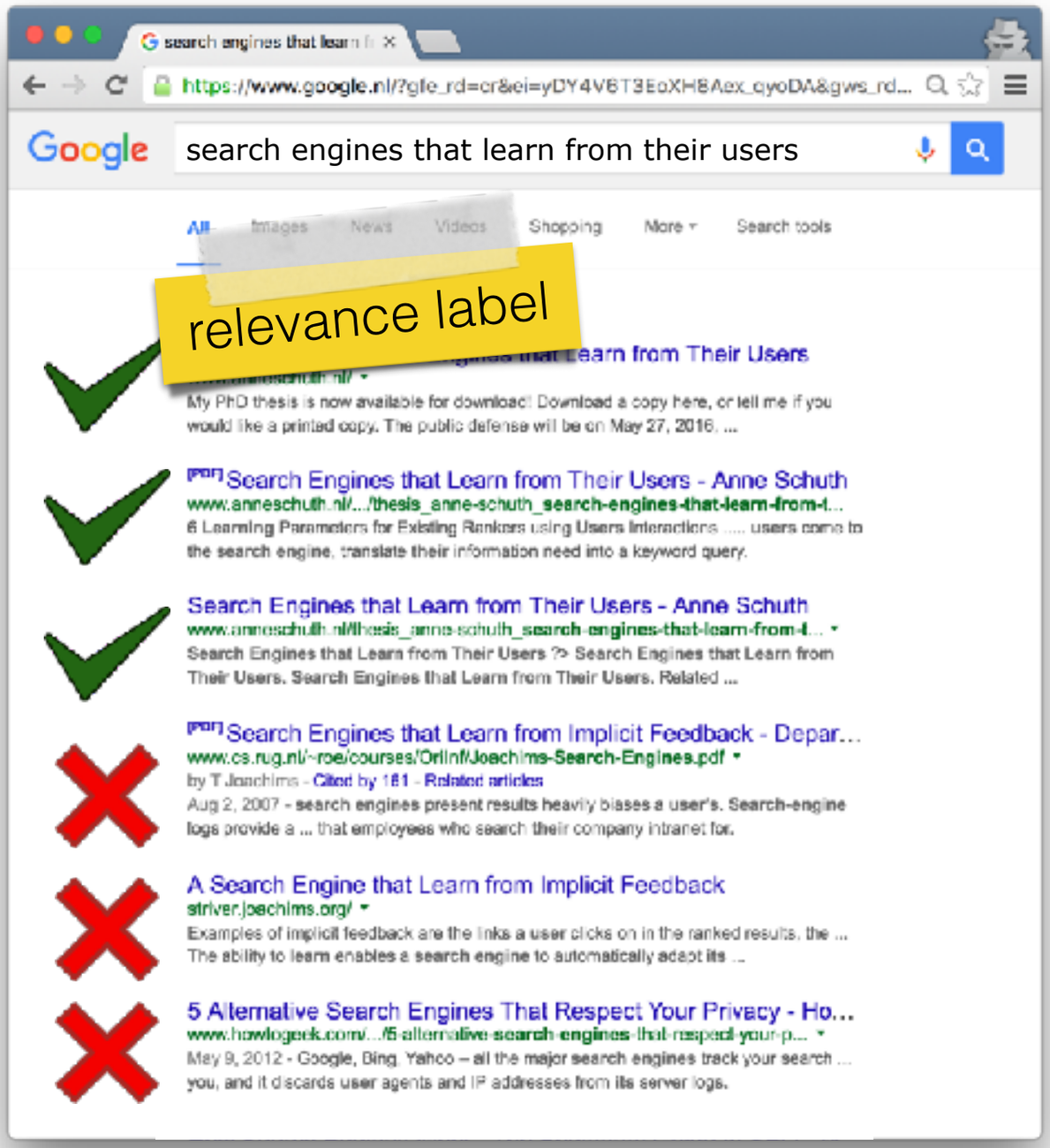


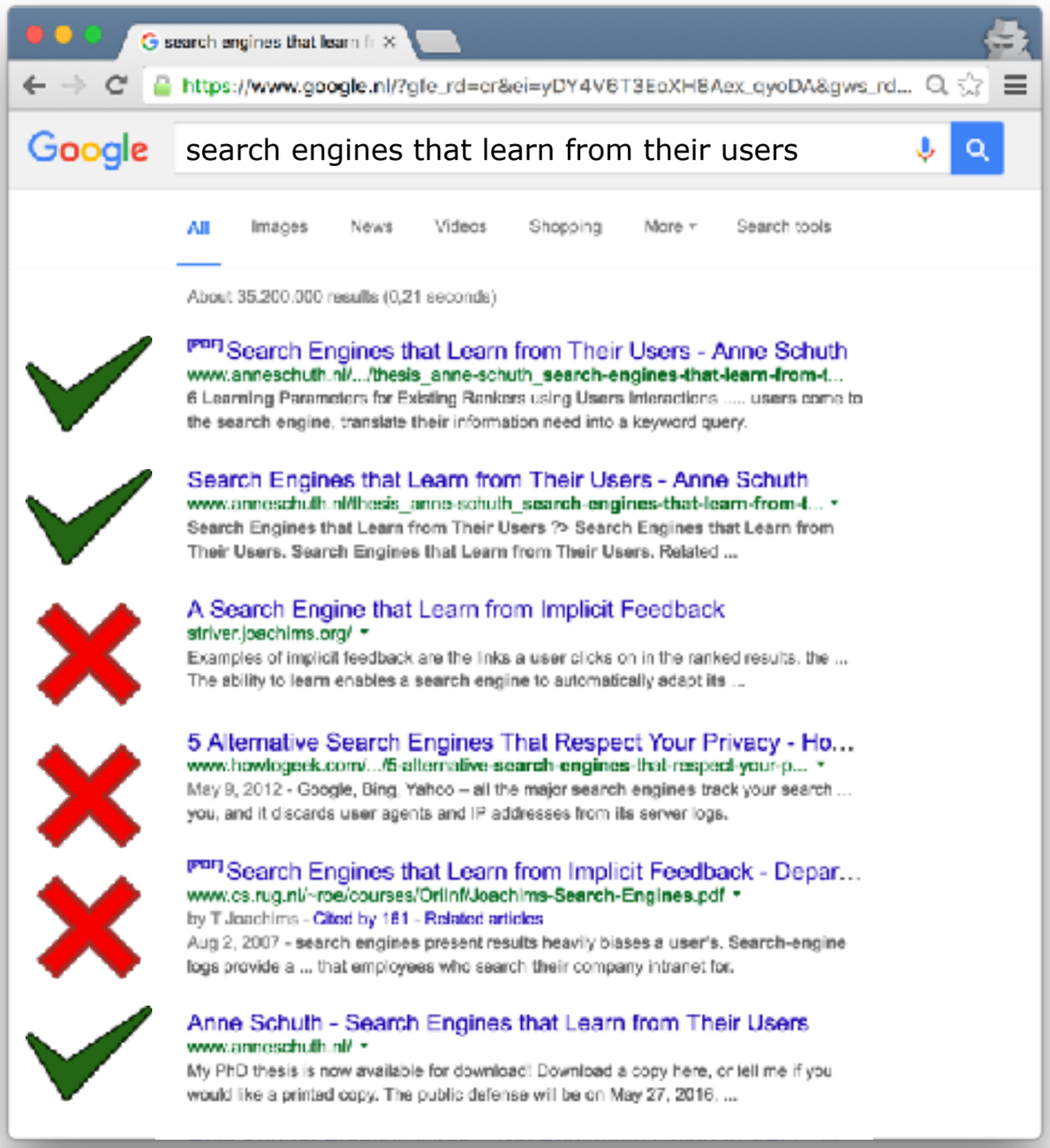
EVALUATION OF SEARCH



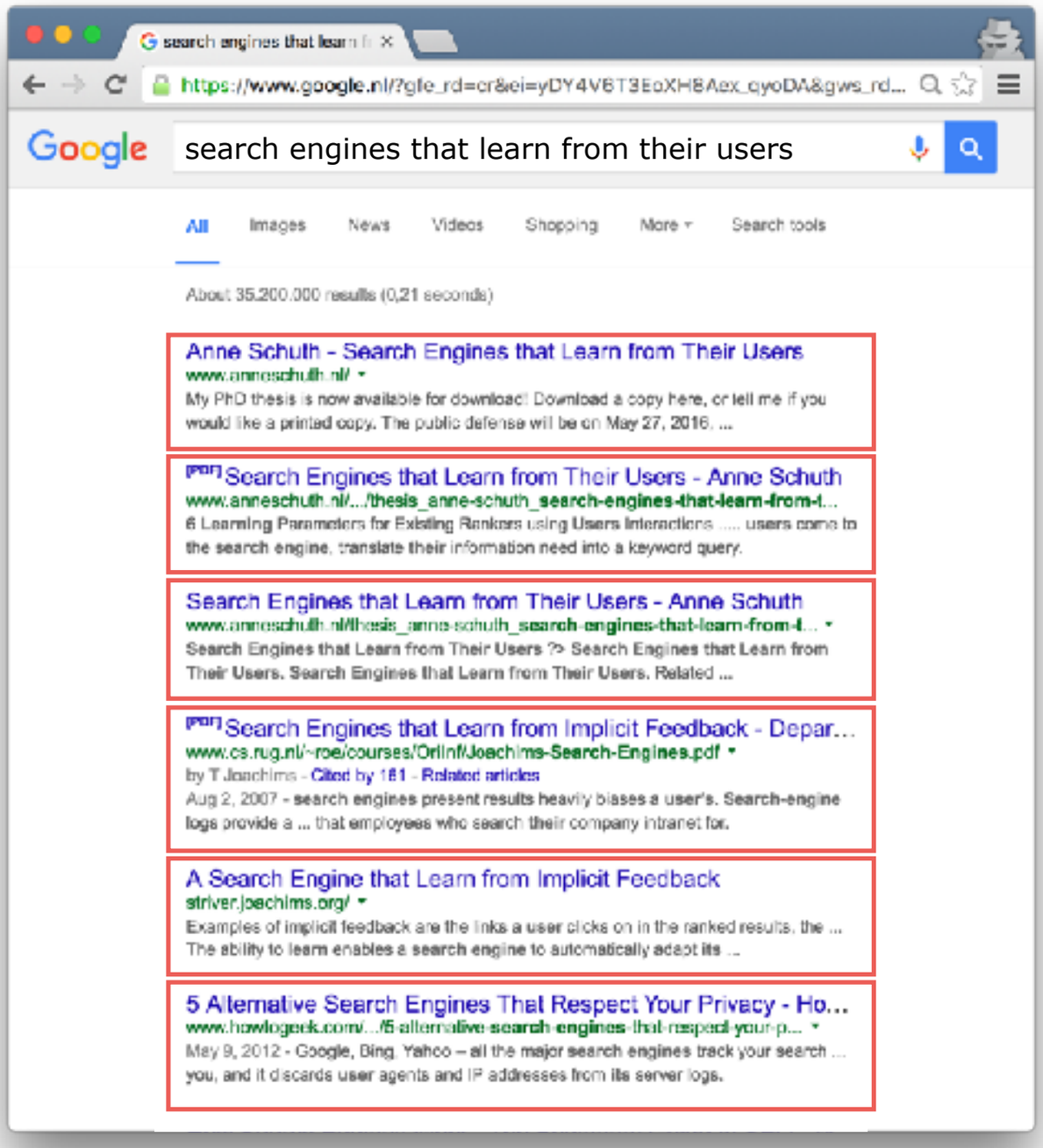


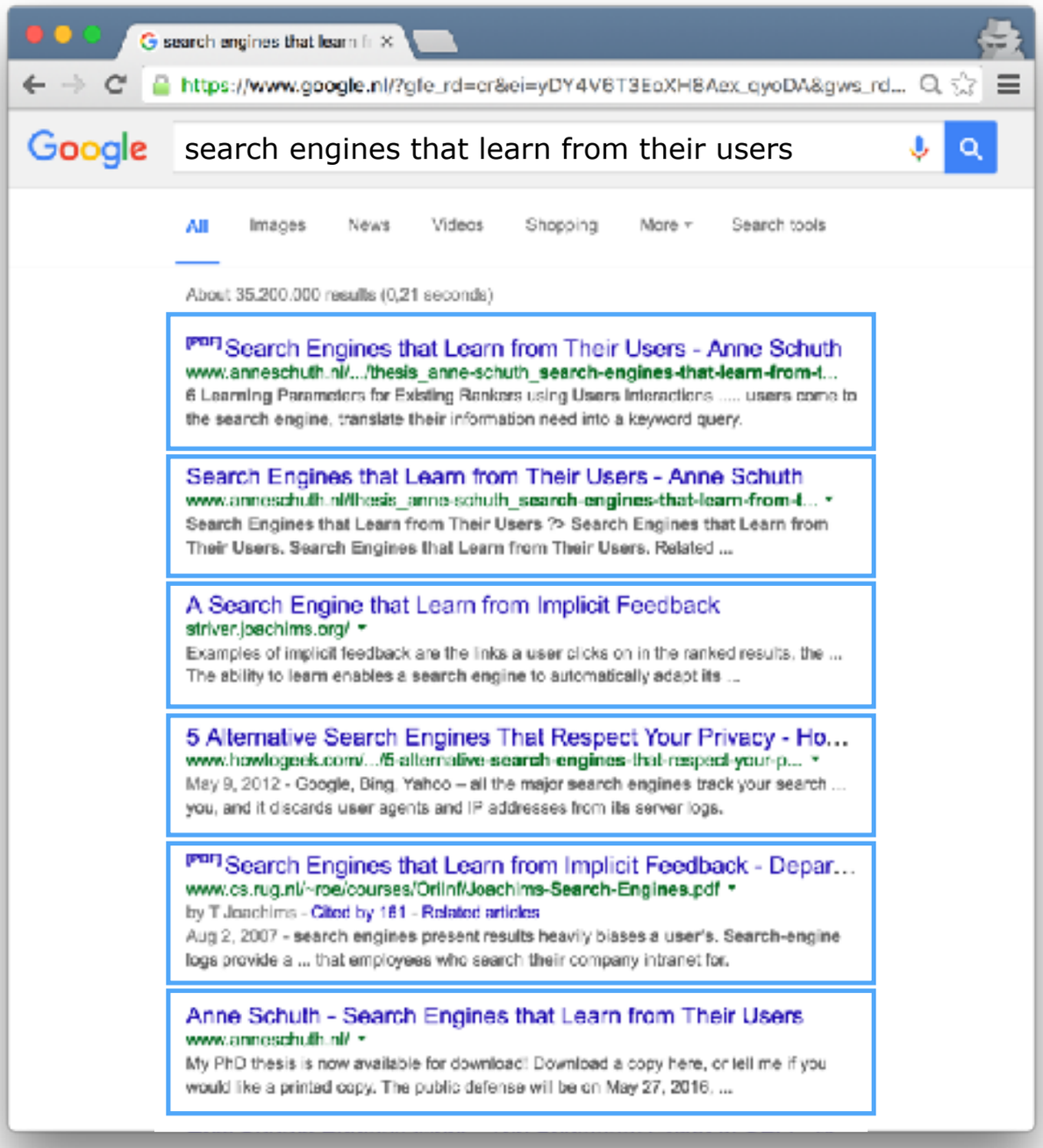
EVALUATION OF SEARCH



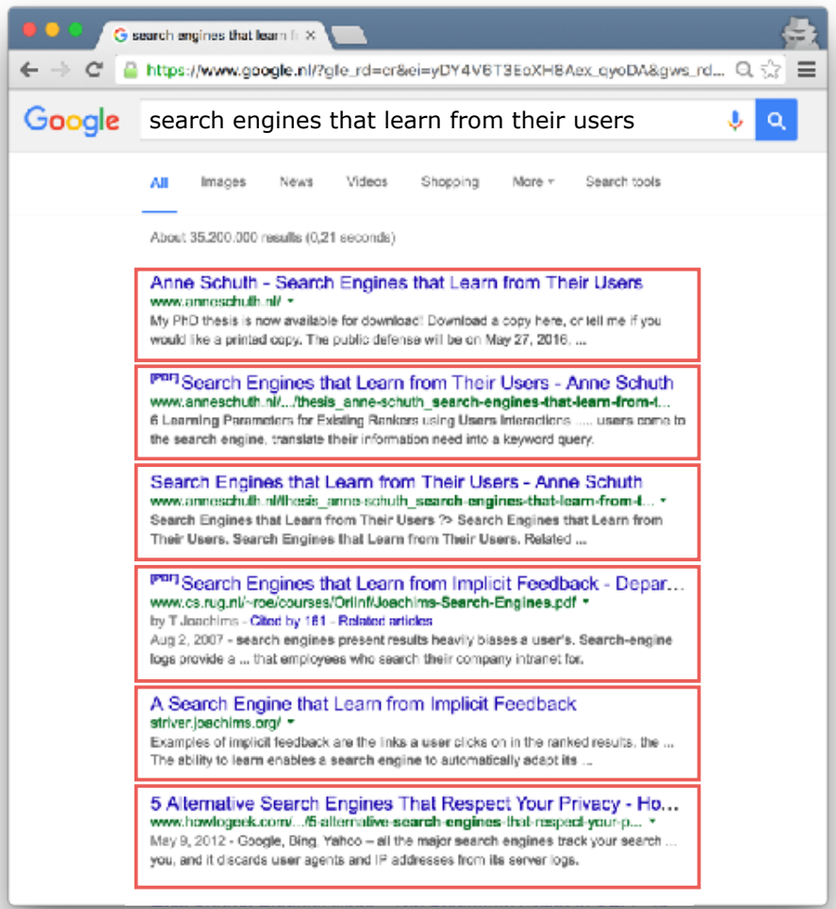


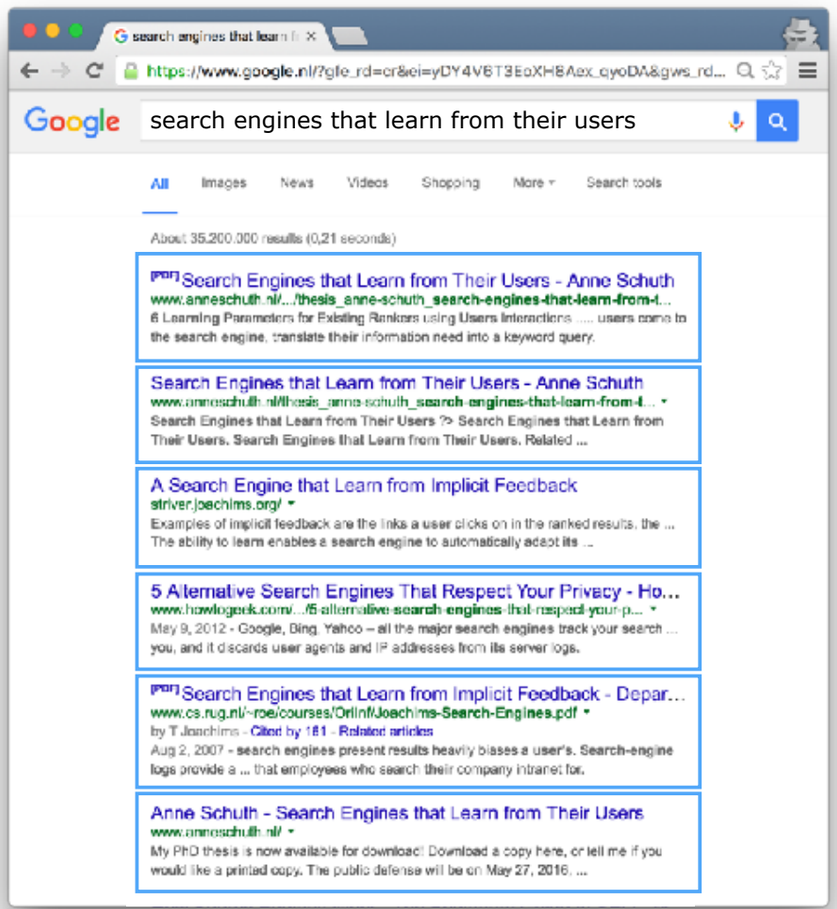
INTERLEAVING



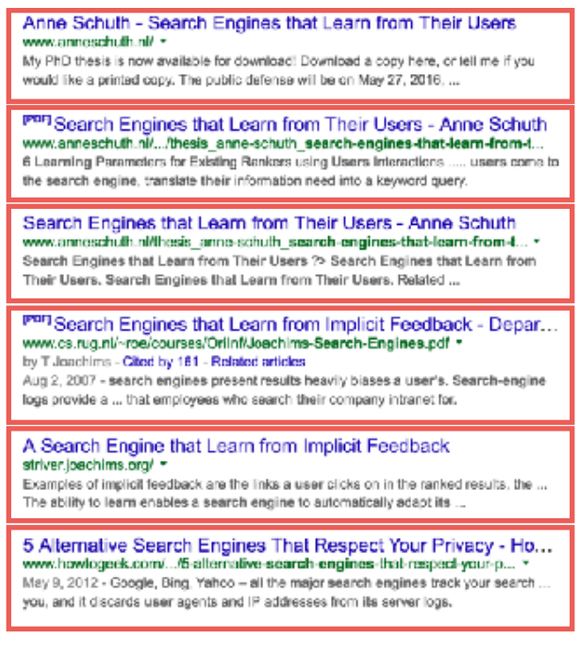


INTERLEAVING





INTERLEAVING



INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

versus

Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

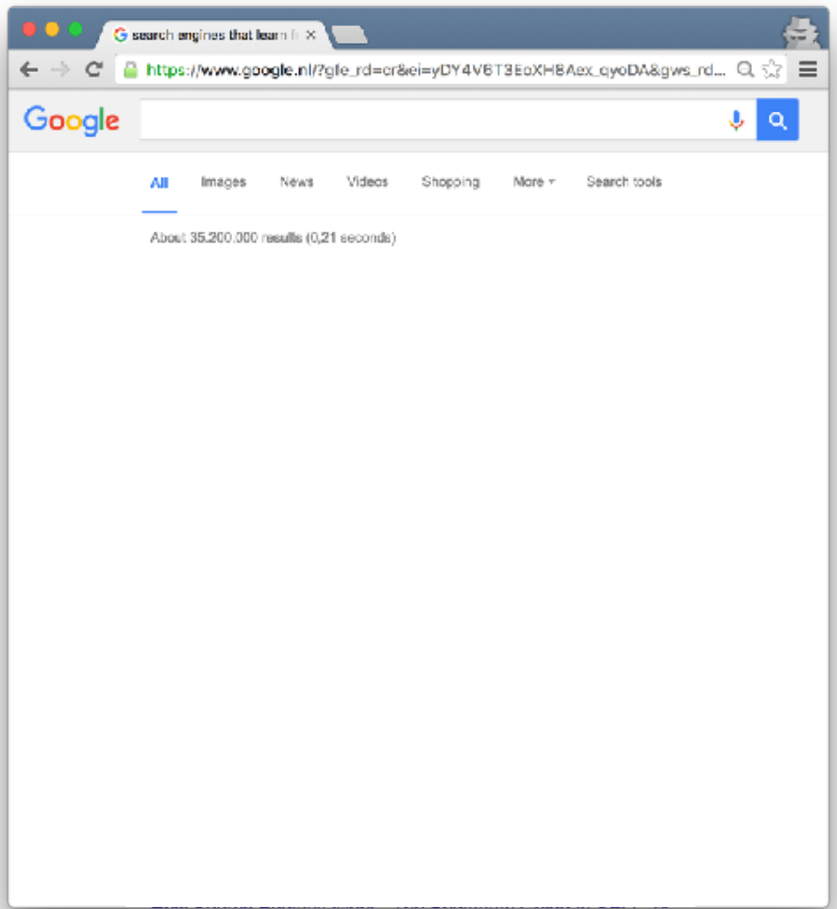
[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

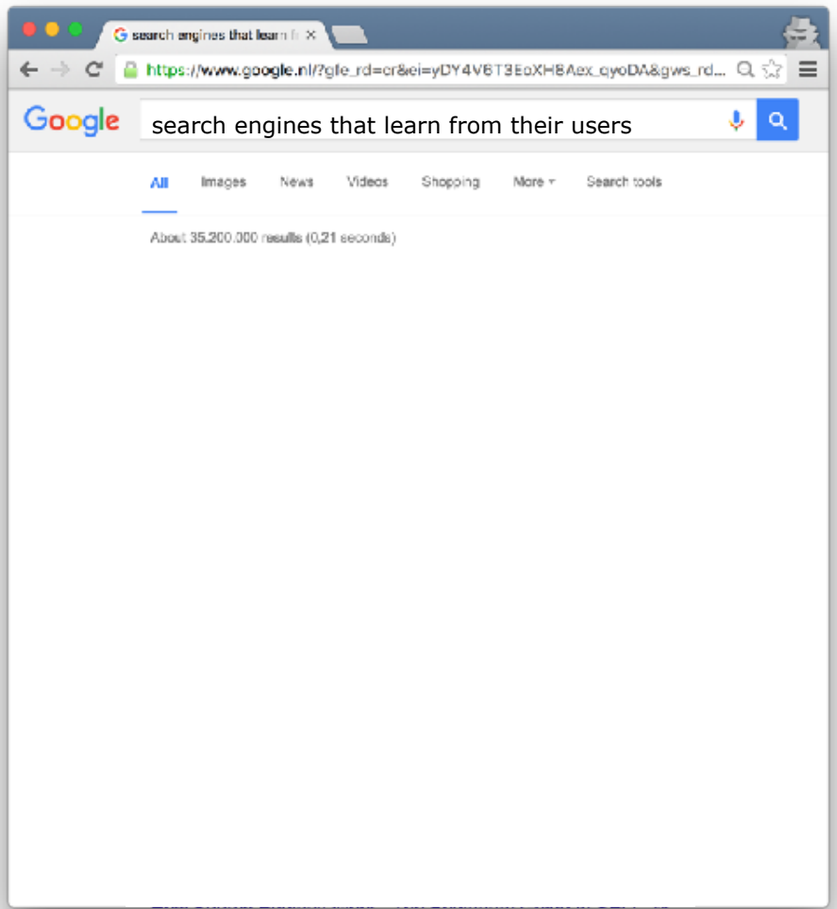
[PDF] **Search Engines that Learn from Their Users - Anne Schuth**
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] **Search Engines that Learn from Implicit Feedback - Depar...**
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google B

[PDF] **Search Engines that Learn from Their Users - Anne Schuth**
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] **Search Engines that Learn from Implicit Feedback - Depar...**
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

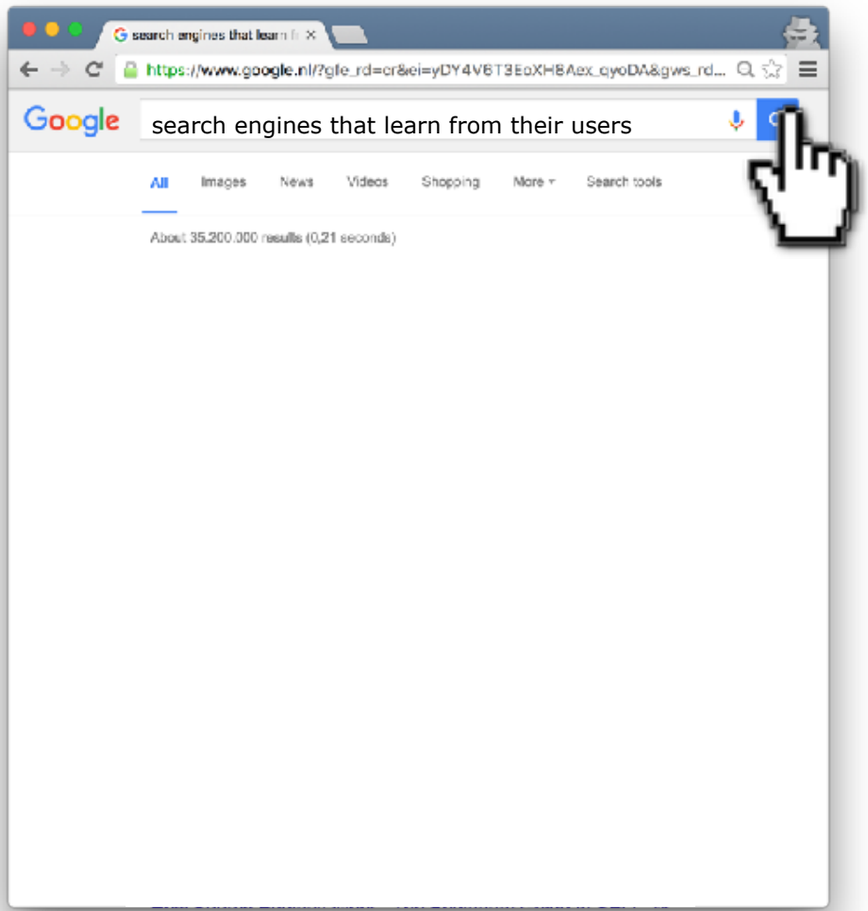
[PDF] **Search Engines that Learn from Their Users - Anne Schuth**
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] **Search Engines that Learn from Implicit Feedback - Depar...**
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google B

[PDF] **Search Engines that Learn from Their Users - Anne Schuth**
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] **Search Engines that Learn from Implicit Feedback - Depar...**
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

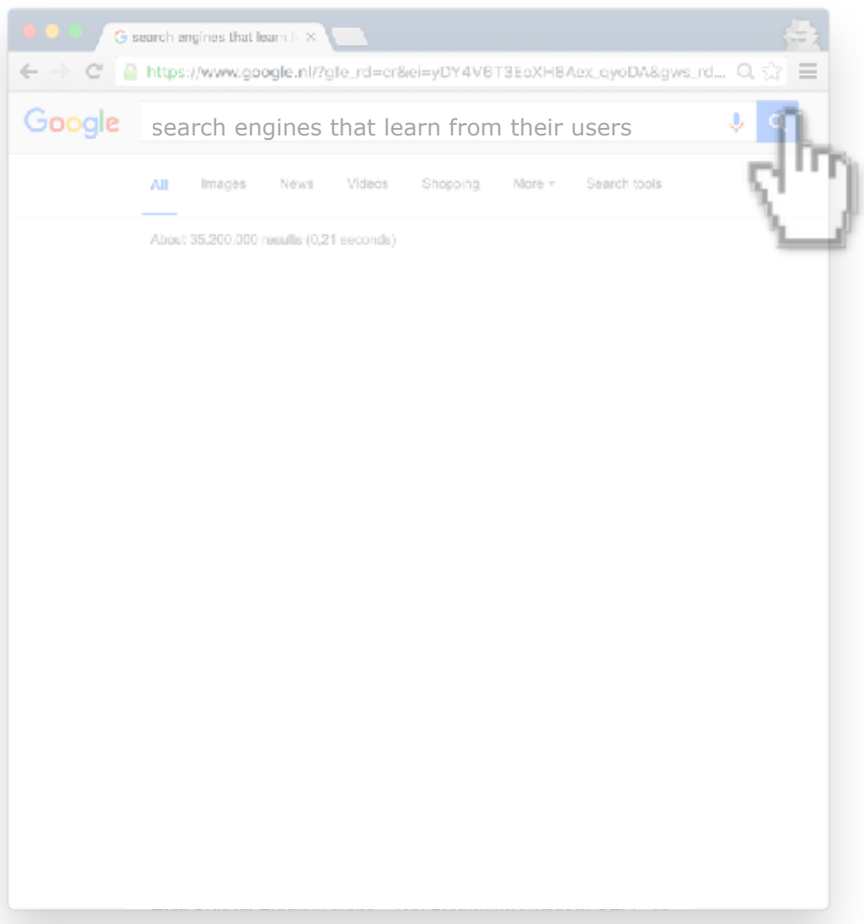
[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

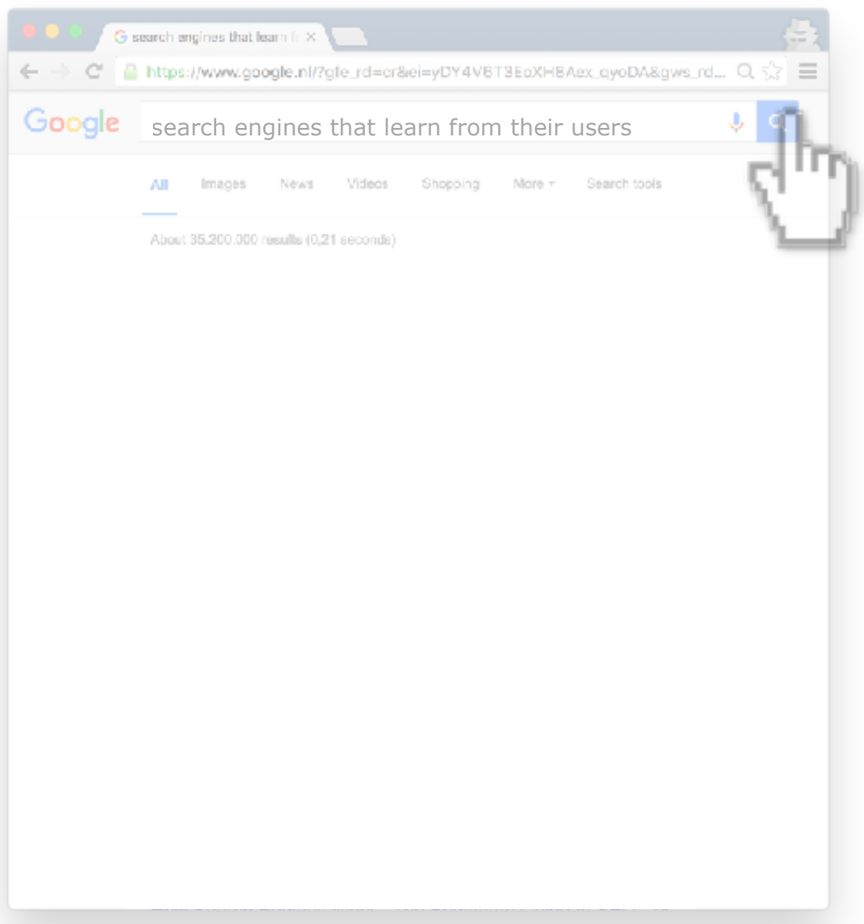
[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

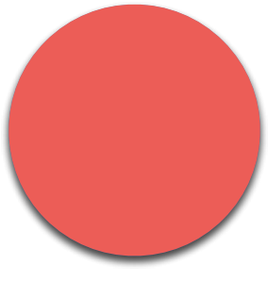
Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...



INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

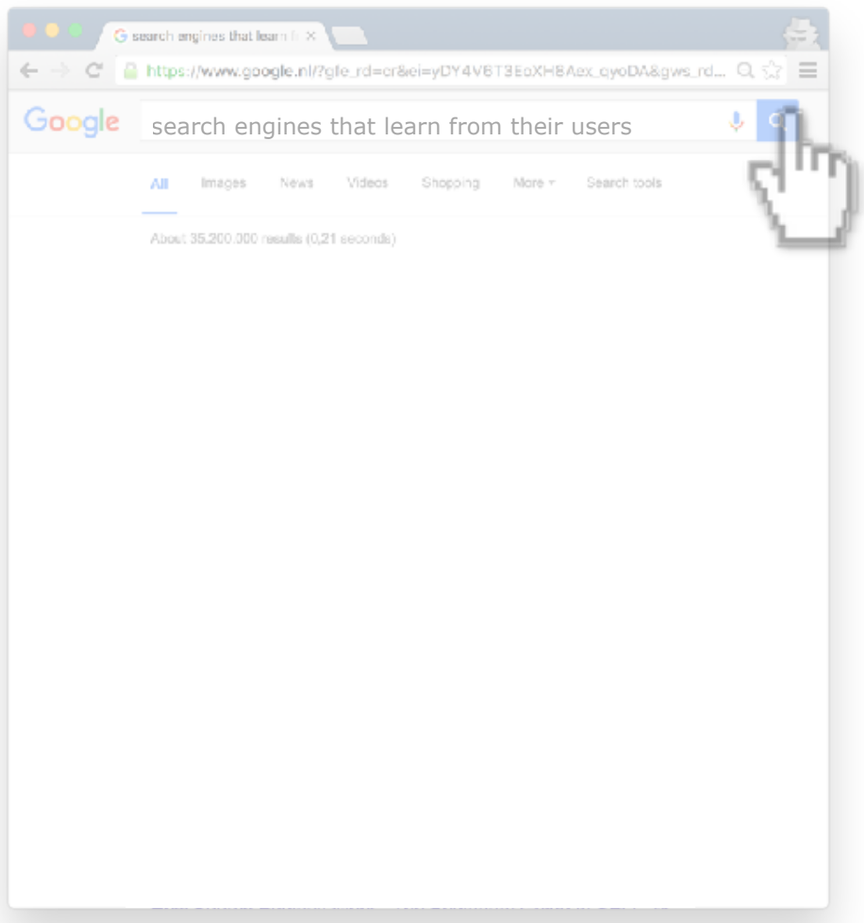
[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

Google **A**

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

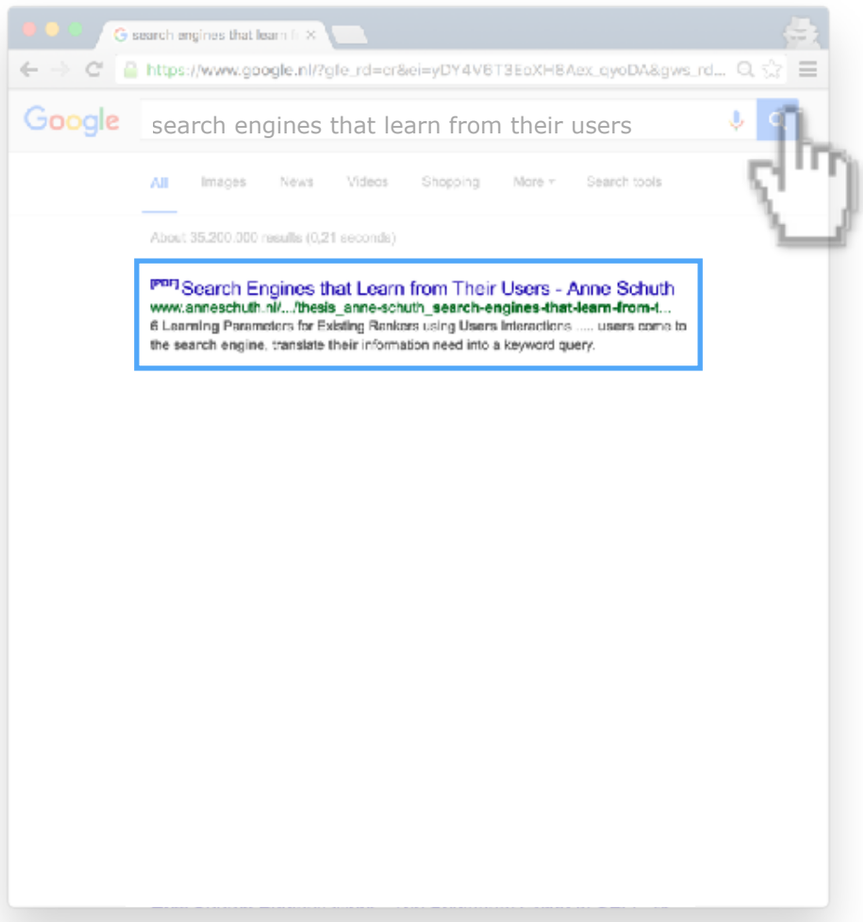
[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t... ▾
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google **B**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/ ▾
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p... ▾
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf ▾
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/ ▾
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

Google **A**

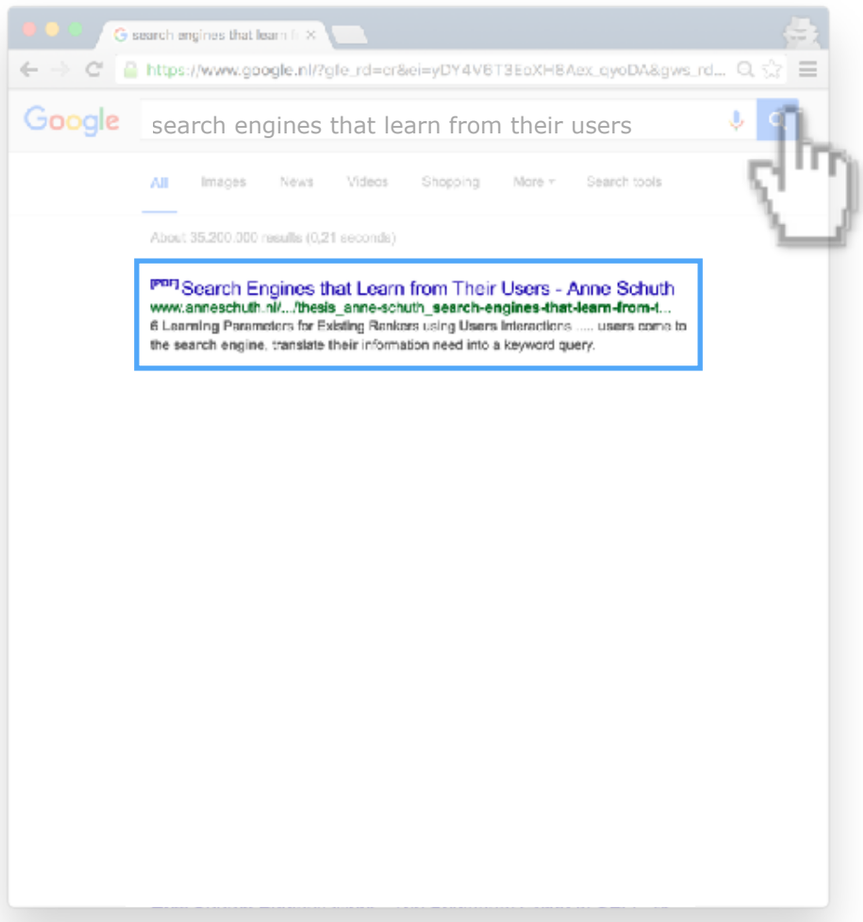
Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google **B**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

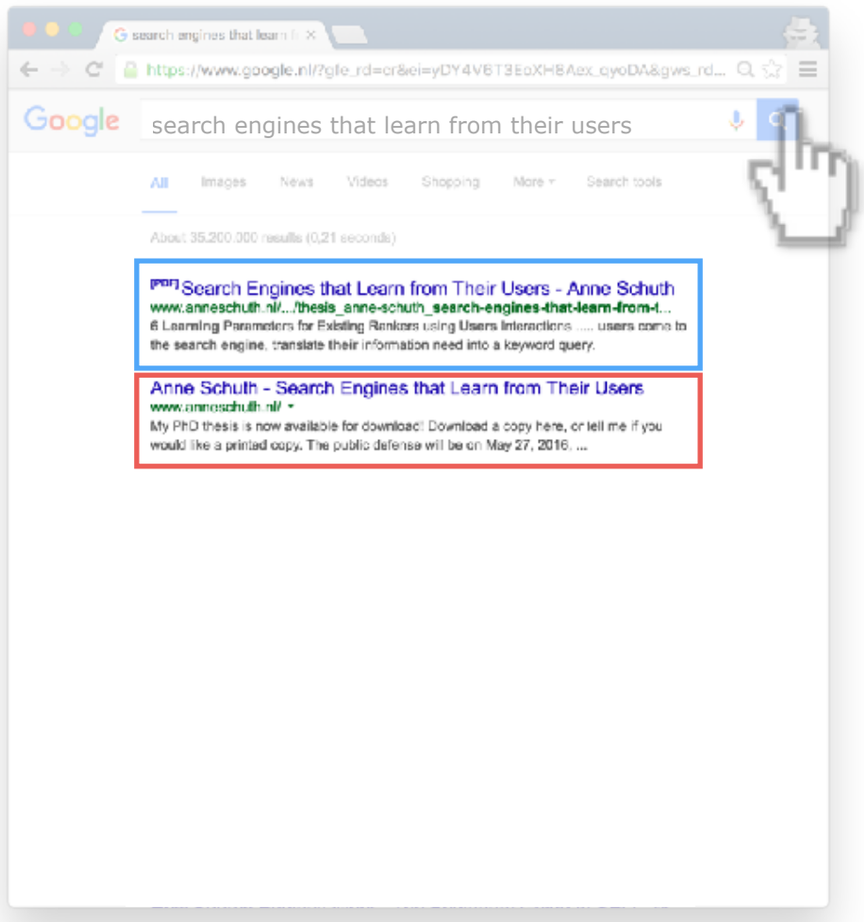
Google **A**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google **B**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

INTERLEAVING

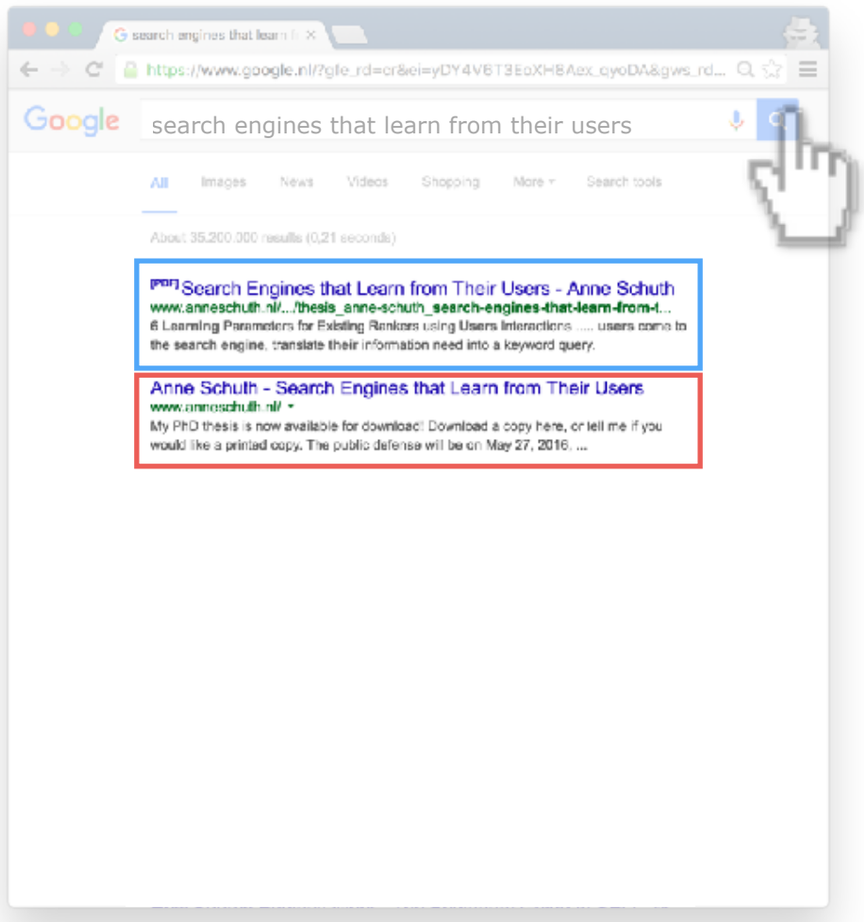
Google **A**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google **B**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

INTERLEAVING

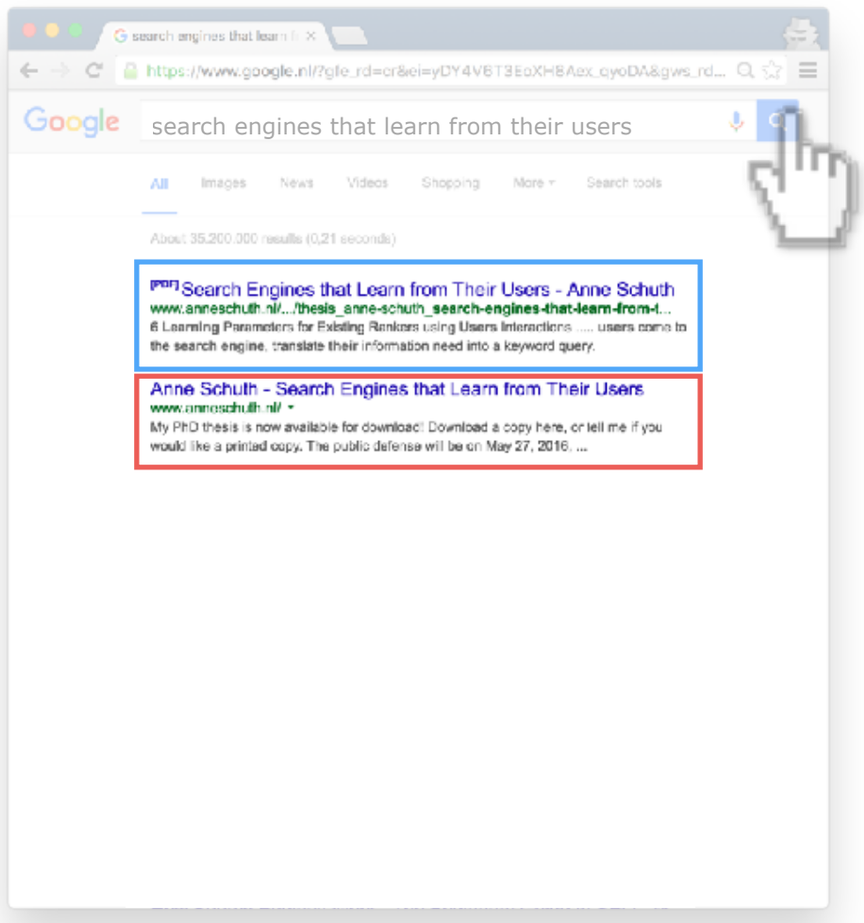
Google **A**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



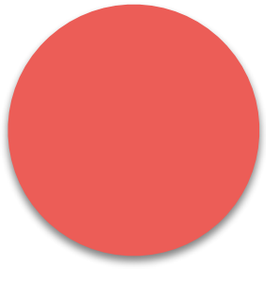
Google **B**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

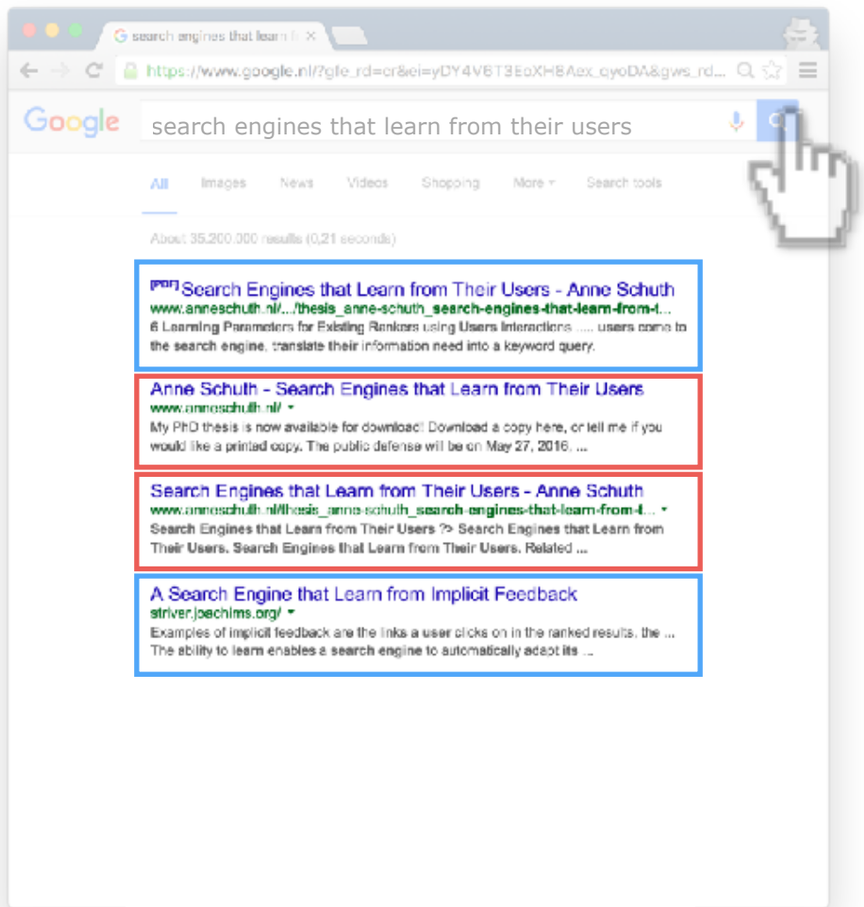


INTERLEAVING

Google **A**

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

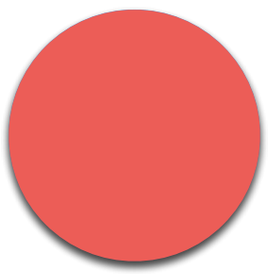
5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google **B**

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

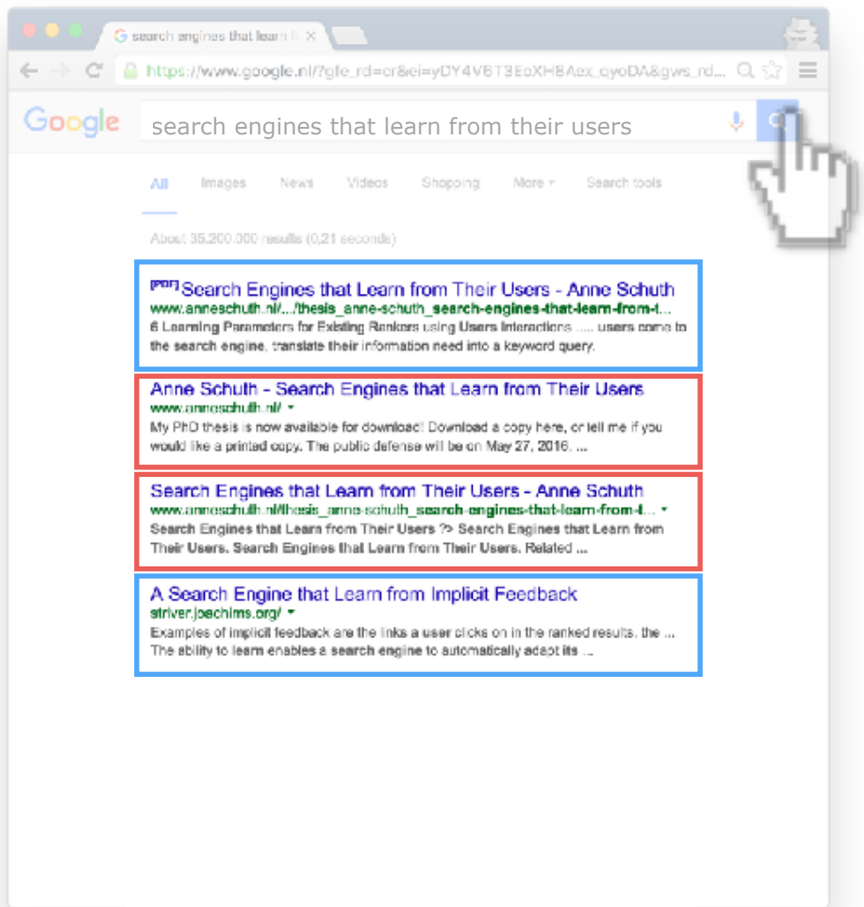


INTERLEAVING

Google **A**

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

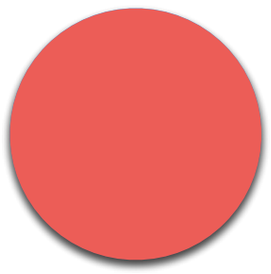
5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.



Google **B**

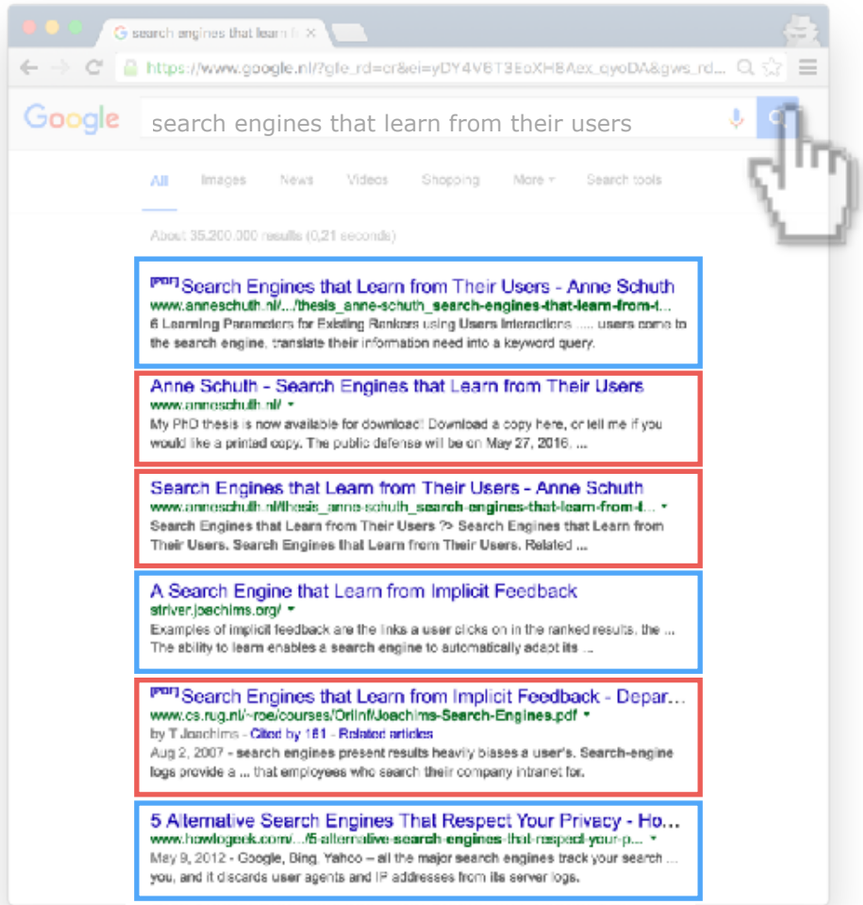
5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.



INTERLEAVING

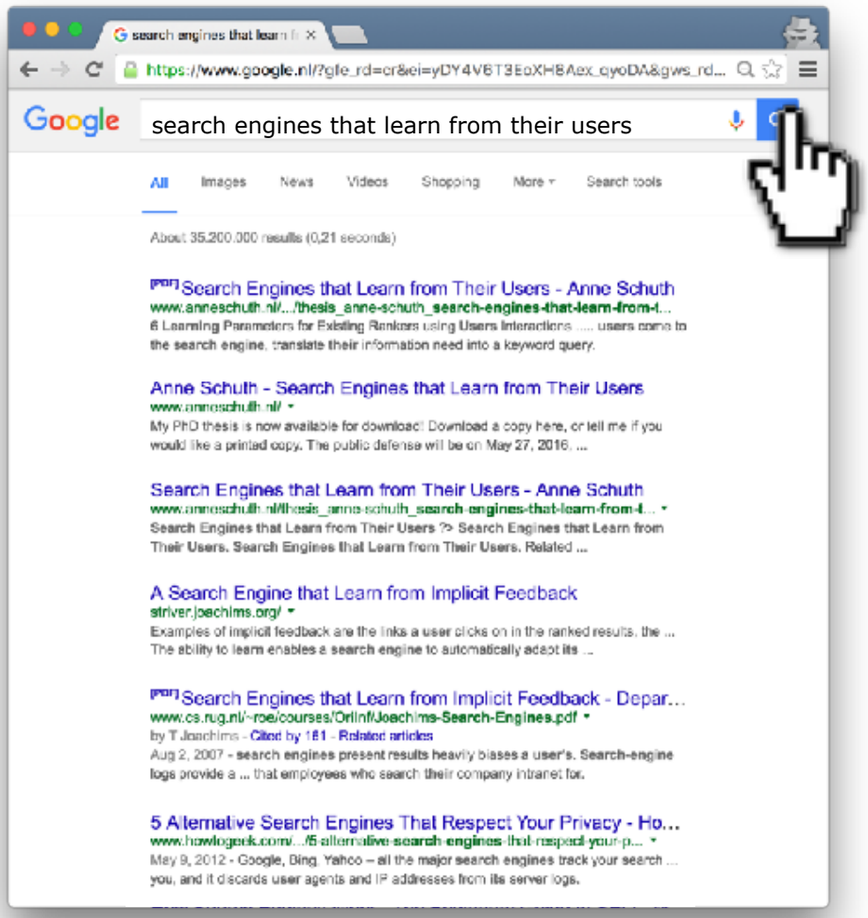
Google **A**



Google **B**

INTERLEAVING

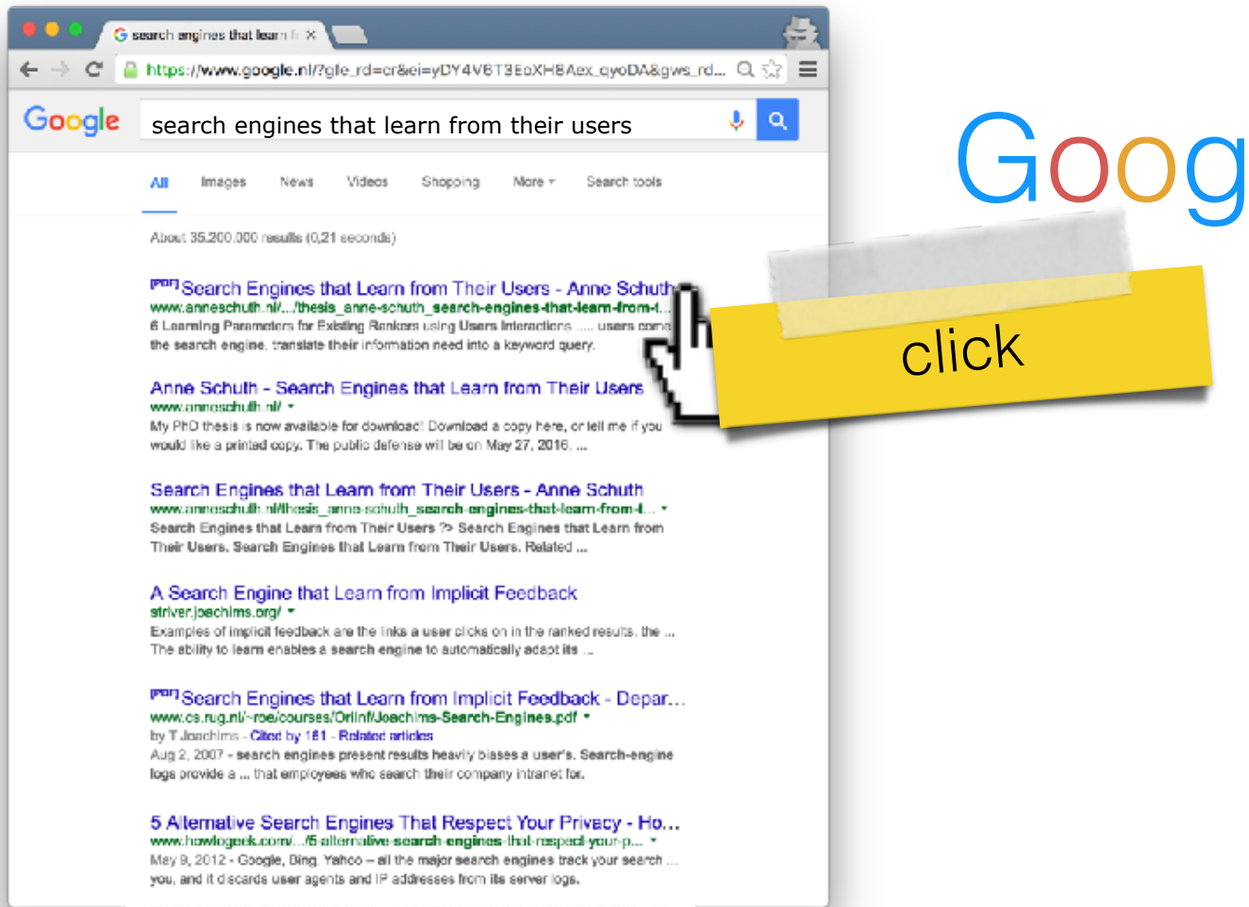
Google **A**



Google **B**

INTERLEAVING

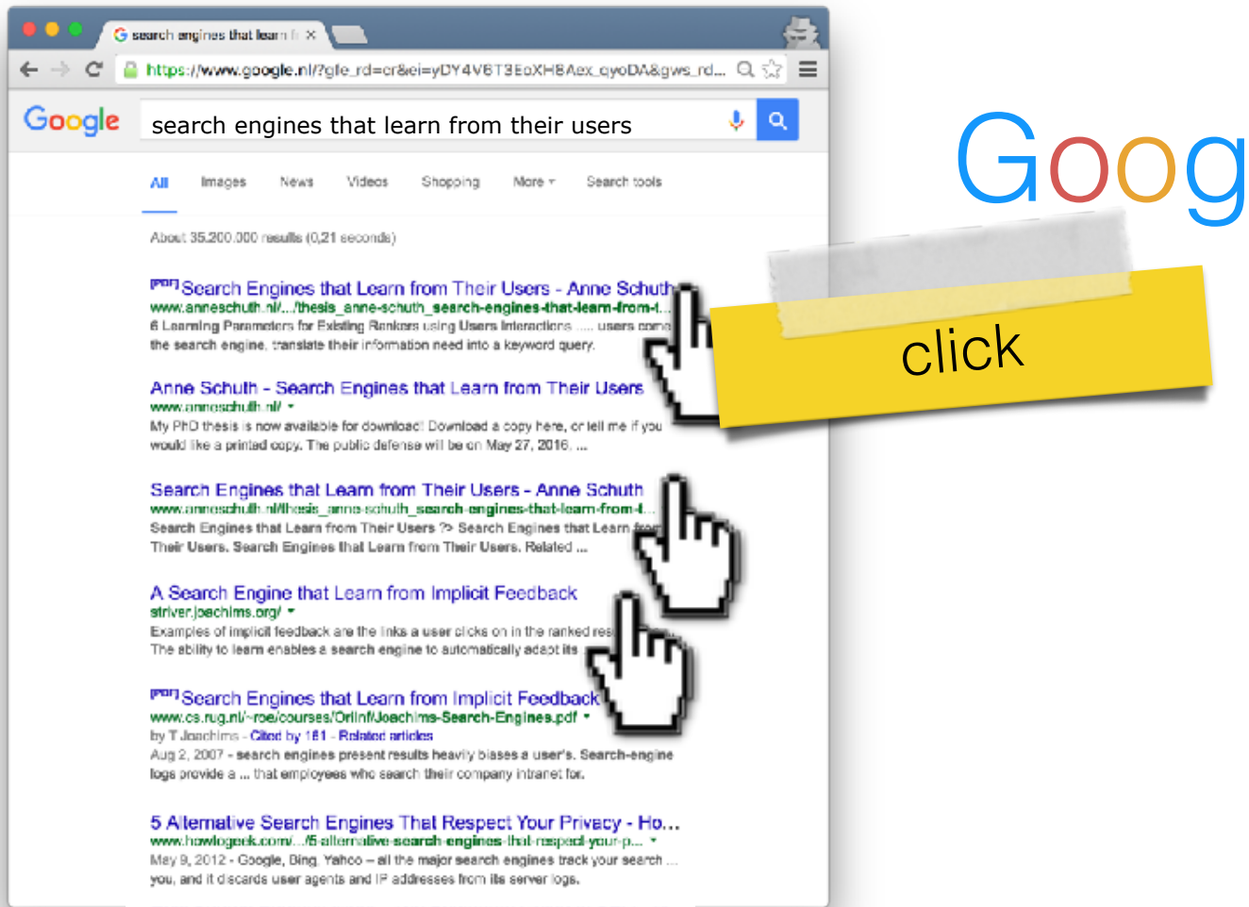
Google **A**



Google **B**

INTERLEAVING

Google **A**



Google **B**

INTERLEAVING

Google **A**

Google **B**

INTERLEAVING

Google **A**

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t
Search Engines that Learn from Their Users ?> Search Engines that Learn fr
Their Users. Search Engines that Learn from Their Users. Related ...

Google **B**

[PDF] Search Engines that Learn from Their Users - Anne Schu
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t
6 Learning Parameters for Existing Rankers using Users Interactions ..... users one
the search engine, translate their information need into a keyword query.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked l
The ability to learn enables a search engine to automatically adapt it

INTERLEAVING

Google **A** loses against Google **B**

INTERLEAVING

Google A

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions ..... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlInf/Joachims-Search-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Engines that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

MULTILEAVING

Google A

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/files/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/files/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

Google C

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/files/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

Google D

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/files/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

Google E

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

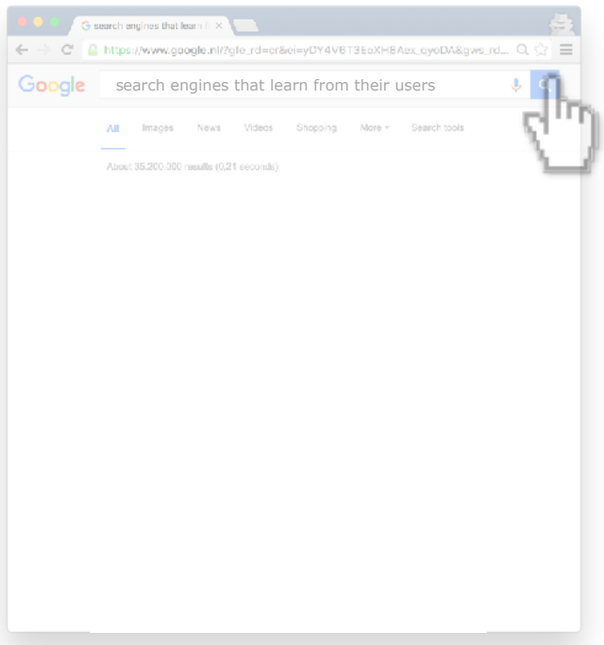
Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/files/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.



MULTILEAVING

Google A

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

Google B

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

Google C

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

Google D

[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

Google E

Anne Schuth - Search Enginas that Learn from Their Users
www.anneschuth.nl/
My PhD thesis is now available for download! Download a copy here, or tell me if you would like a printed copy. The public defense will be on May 27, 2016. ...

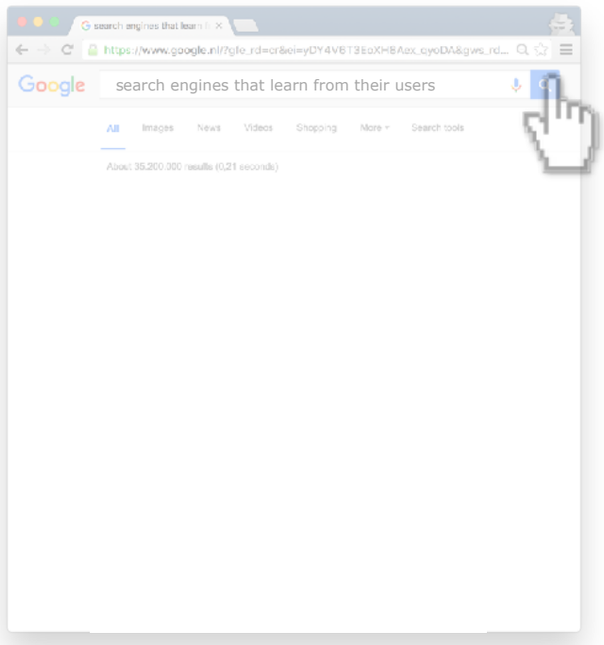
Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/thesis_anne-schuth_search-engines-that-learn-from-t...
Search Engines that Learn from Their Users ?> Search Engines that Learn from Their Users. Search Engines that Learn from Their Users. Related ...

A Search Engine that Learn from Implicit Feedback
striver.joachims.org/
Examples of implicit feedback are the links a user clicks on in the ranked results. the ... The ability to learn enables a search engine to automatically adapt its ...

5 Alternative Search Engines That Respect Your Privacy - Ho...
www.howtogeek.com/.../5-alternative-search-engines-that-respect-your-p...
May 9, 2012 - Google, Bing, Yahoo – all the major search engines track your search ... you, and it discards user agents and IP addresses from its server logs.

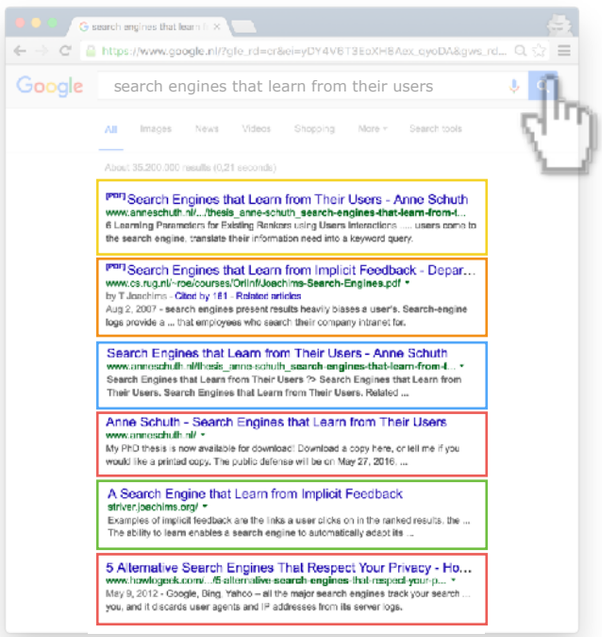
[PDF] Search Engines that Learn from Implicit Feedback - Depar...
www.cs.rug.nl/~roe/courses/OrlinI/JoachimsSearch-Engines.pdf
by T Joachims - Cited by 181 - Related articles
Aug 2, 2007 - search engines present results heavily biases a user's. Search-engine logs provide a ... that employees who search their company intranet for.

[PDF] Search Engines that Learn from Their Users - Anne Schuth
www.anneschuth.nl/.../thesis_anne-schuth_search-engines-that-learn-from-t...
6 Learning Parameters for Existing Rankers using Users Interactions .... users come to the search engine, translate their information need into a keyword query.



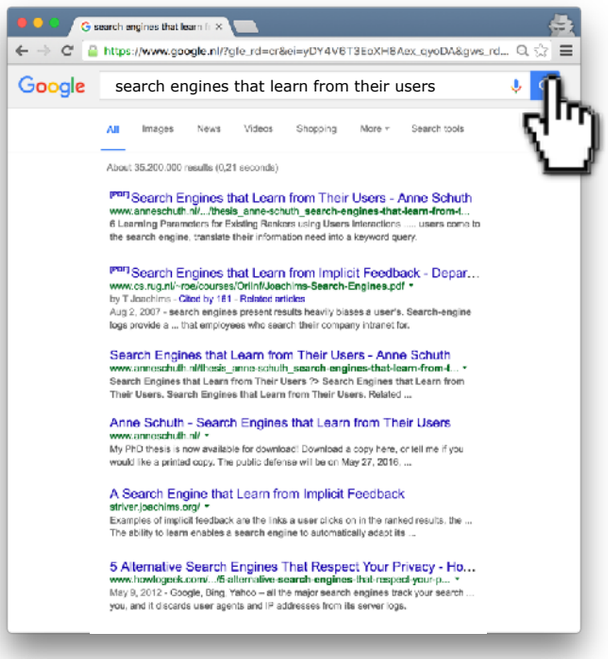
MULTILEAVING

Google **A** Google **B** Google **C** Google **D** Google **E**



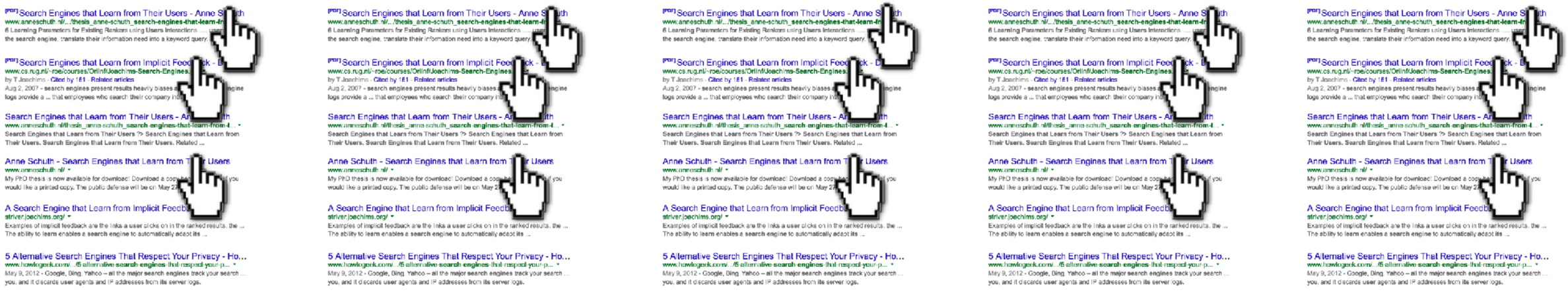
MULTILEAVING

Google **A** Google **B** Google **C** Google **D** Google **E**



MULTILEAVING

Google **A** Google **B** Google **C** Google **D** Google **E**

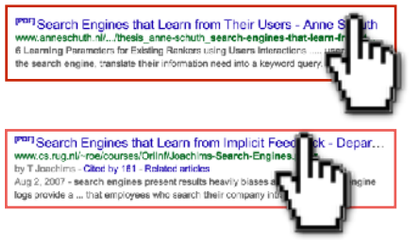


MULTILEAVING

Google **A** Google **B** Google **C** Google **D** Google **E**

MULTILEAVING

Google A



Google B

Google C

Google D

Google E

University of Amsterdam

Microsoft Research

DIR 2015 Amsterdam

CIKM2015
19–23 October 2015 Melbourne Australia

Struggling and Success in Web Search

Daan Odijk, University of Amsterdam, The Netherlands

Ryen W. White, Ahmed Hassan Awadallah & Susan T. Dumais
Microsoft Research, Redmond, WA, USA

Motivation

- Long web search sessions are common

Motivation

- Long web search sessions are common
  - Half of web search sessions contain multiple queries [Shokouhi et al., SIGIR'13]
  - 40% of sessions take 3+ minutes [Dumais, NSF'13]
    - Account for most of search time

Motivation

- Long web search sessions are common
  - Half of web search sessions contain multiple queries [Shokouhi et al., SIGIR'13]
    - 40% of sessions take 3+ minutes [Dumais, NSF'13]
      - Account for most of search time
- What are searchers doing? Struggling or Exploring. [Hassan et al., WSDM'14]

Motivation

- Long web search sessions are common
  - Half of web search sessions contain multiple queries [Shokouhi et al., SIGIR'13]
    - 40% of sessions take 3+ minutes [Dumais, NSF'13]
      - Account for most of search time
- What are searchers doing? Struggling or Exploring. [Hassan et al., WSDM'14]

Motivation

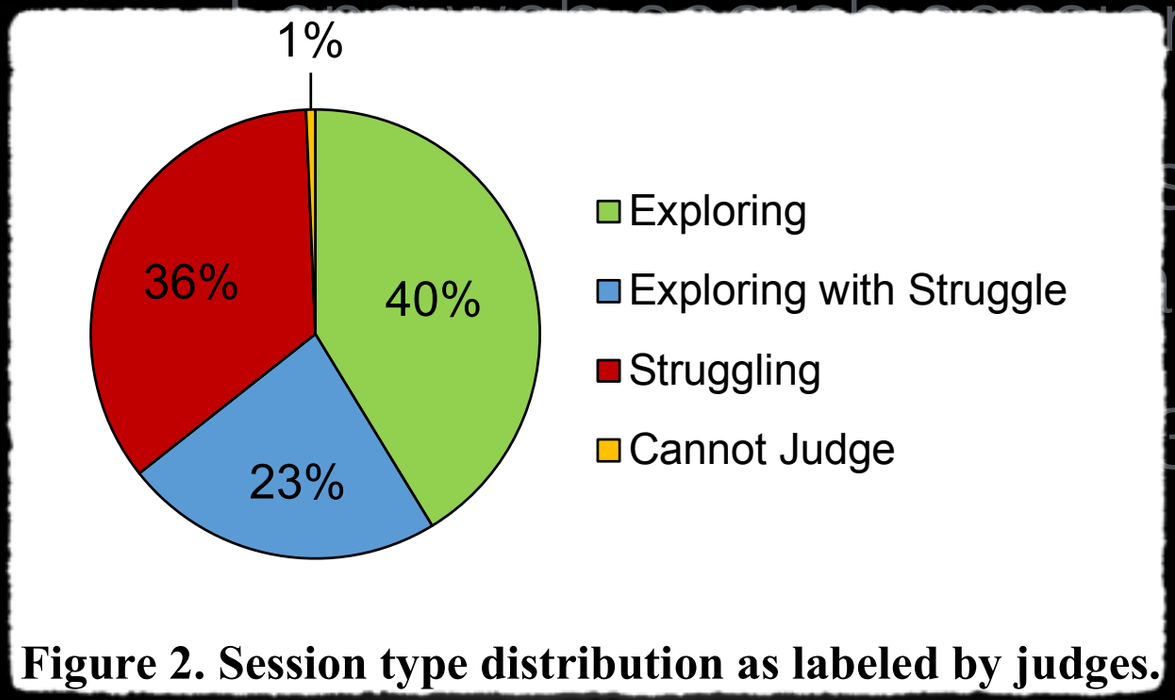


Figure 2. Session type distribution as labeled by judges.

ns are common

ssions contain multiple
l., SIGIR'13]

3+ minutes [Dumais, NSF'13]

search time

- What are searchers doing? Struggling or Exploring. [Hassan et al., WSDM'14]

Motivation

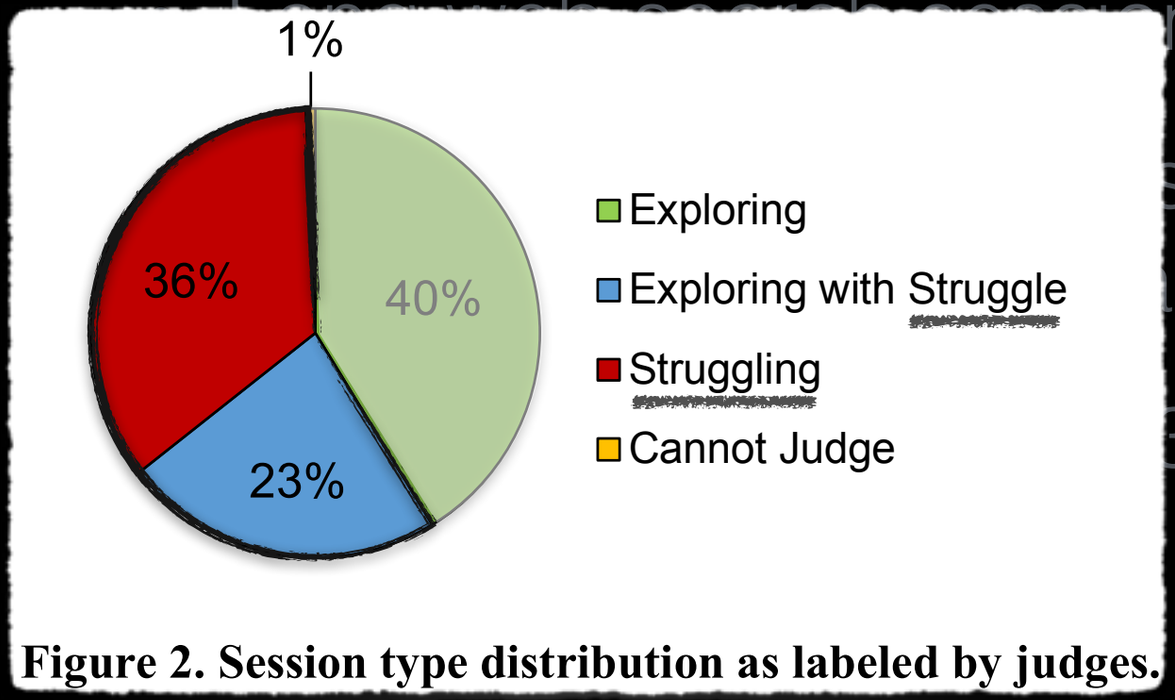


Figure 2. Session type distribution as labeled by judges.

ns are common

ssions contain multiple
l., SIGIR'13]

3+ minutes [Dumais, NSF'13]

search time

- What are searchers doing? Struggling or Exploring. [Hassan et al., WSDM'14]
  - Struggling in 60% of long sessions

Motivation

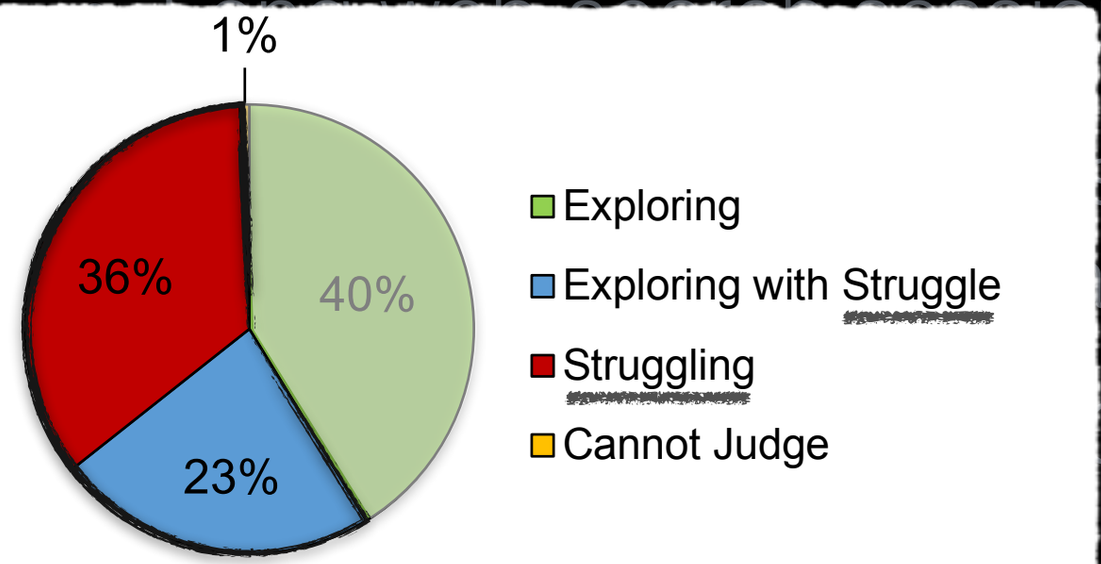


Figure 2. Session type distribution as labeled by judges.

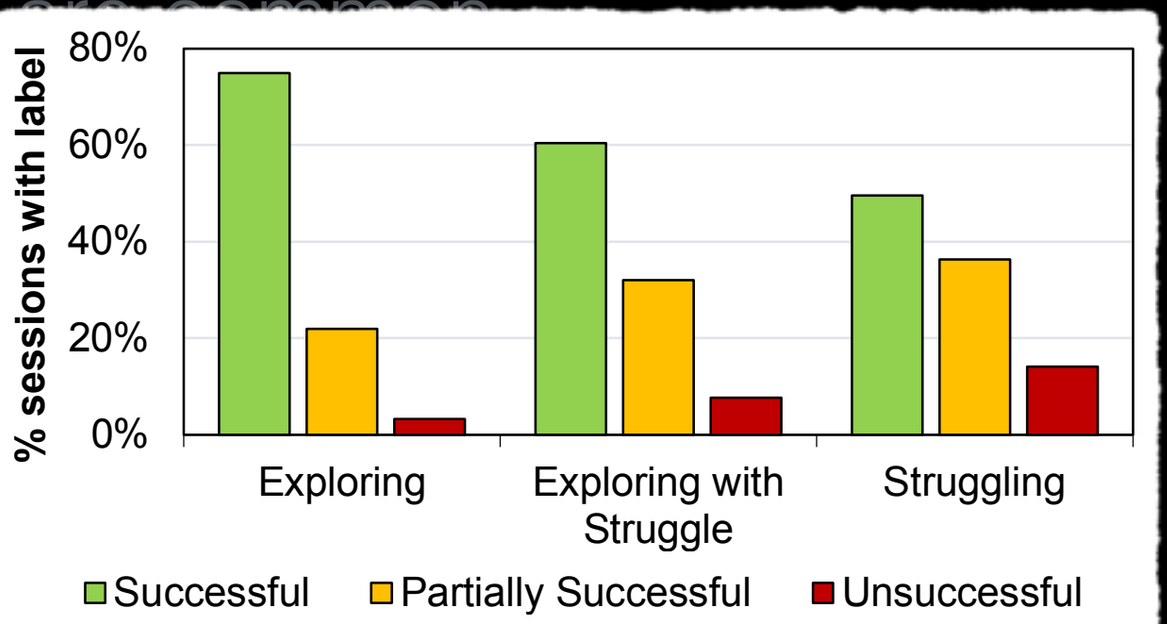


Figure 3. Session success distribution as labeled by judges.

- What are searchers doing? Struggling or Exploring. [Hassan et al., WSDM'14]
  - Struggling in 60% of long sessions, often successful

Motivation

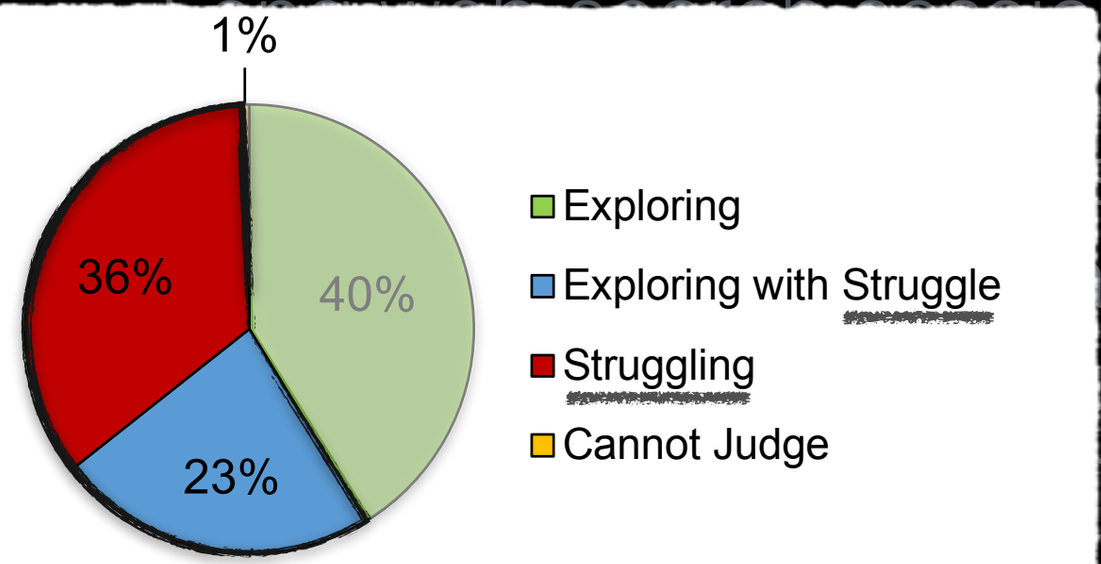


Figure 2. Session type distribution as labeled by judges.

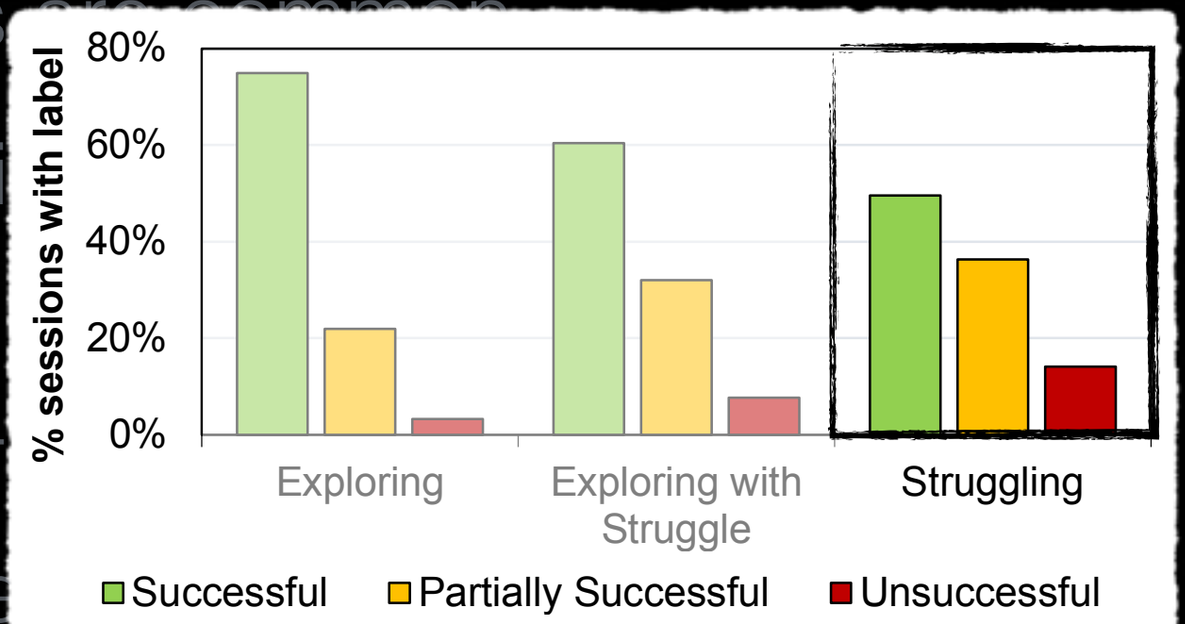


Figure 3. Session success distribution as labeled by judges.

- What are searchers doing? Struggling or Exploring. [Hassan et al., WSDM'14]
  - Struggling in 60% of long sessions, often successful

Example of Struggling

Logged session from June 2014

Example of Struggling

Logged session from June 2014

9:13:11 AM	**Query**	us open
9:13:24 AM	**Query**	us open golf
9:13:36 AM	**Query**	us open golf 2013 live
9:13:59 AM	**Query**	watch us open live streaming
9:14:02 AM	**Click**	http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video
9:31:55 AM	**END**	

Example of Struggling

Logged session from June 2014

9:13:11 AM	**Query**	us open
9:13:24 AM	**Query**	us open golf
9:13:36 AM	**Query**	us open golf 2013 live
9:13:59 AM	**Query**	watch us open live streaming
9:14:02 AM	**Click**	http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video
9:31:55 AM	**END**	

Example of Struggling

Logged session from June 2014

9:13:11 AM	**Query**	us open
9:13:24 AM	**Query**	us open golf
9:13:36 AM	**Query**	us open golf 2013 live
9:13:59 AM	**Query**	watch us open live streaming
9:14:02 AM	**Click**	http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video
9:31:55 AM	**END**	

Example of Struggling

Logged session from June 2014

9:13:11 AM	**Query**	us open
9:13:24 AM	**Query**	us open golf
9:13:36 AM	**Query**	us open golf 2013 live
9:13:59 AM	**Query**	watch us open live streaming
9:14:02 AM	**Click**	http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video
9:31:55 AM	**END**	

Example of Struggling

Logged session from June 2014

9:13:11 AM	**Query**	us open
9:13:24 AM	**Query**	us open golf
9:13:36 AM	**Query**	us open golf 2013 live
9:13:59 AM	**Query**	watch us open live streaming
9:14:02 AM	**Click**	http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video
9:31:55 AM	**END**	

Example of Struggling

Logged session from June 2014

9:13:11 AM **Query** us open

9:13:24 AM **Query** us open golf

9:13:36 AM **Query** us open golf 2013 live

Pivotal Query **Query** watch us open live streaming

9:14:02 AM **Click** http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video

9:31:55 AM **END**

Example of Struggling

Logged session from June 2014

9:13:11 AM **Query** us open

9:13:24 AM **Query** us open golf

9:13:36 AM **Query** us open golf 2013 live

Pivotal Query

Query watch us open live streaming

9:14:02 AM **Click** http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video

9:31:55 AM **END**

How & why do searchers struggle?

Example of Struggling

Logged session from June 2014

9:13:11 AM **Query** us open

9:13:24 AM **Query** us open golf

9:13:36 AM **Query** us open golf 2013 live

Pivotal Query

Query watch us open live streaming

9:14:02 AM **Click** http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video

9:31:55 AM **END**

How & why do searchers struggle?

How do searchers go from struggle to success?

Example of Struggling

Logged session from June 2014

9:13:11 AM **Query** us open

9:13:24 AM **Query** us open golf

9:13:36 AM **Query** us open golf 2013 live

Pivotal Query **Query** watch us open live streaming

9:14:02 AM **Click** http://inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video

9:31:55 AM **END**

How & why do searchers struggle?

How do searchers go from struggle to success?

Can we help them struggle less?

Outline

- Characterizing Struggling
- Annotating Query Transitions
- Predicting Future Actions

Outline

- Characterizing Struggling

  Struggling searchers behave differently given different outcomes on many search aspects, including: queries, reformulations, clicks, dwell time & topic.

- Annotating Query Transitions

- Predicting Future Actions

Identifying Struggling Sessions

June 1–7, 2014

1. **Filter sessions:** US & English, start with a typed query
2. **Segment into tasks:** Topically coherent sub-sessions with at least 3 queries
3. **Filter struggling tasks**

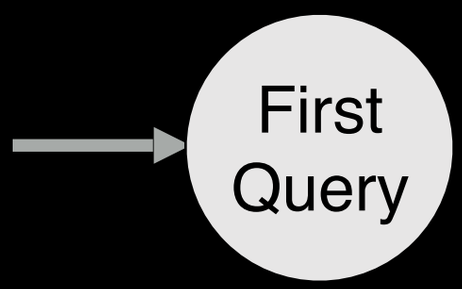
[Hassan et al., WSDM'14]

Identifying Struggling Sessions

June 1–7, 2014

1. **Filter sessions:** US & English, start with a typed query
2. **Segment into tasks:** Topically coherent sub-sessions with at least 3 queries
3. **Filter struggling tasks**

[Hassan et al., WSDM'14]



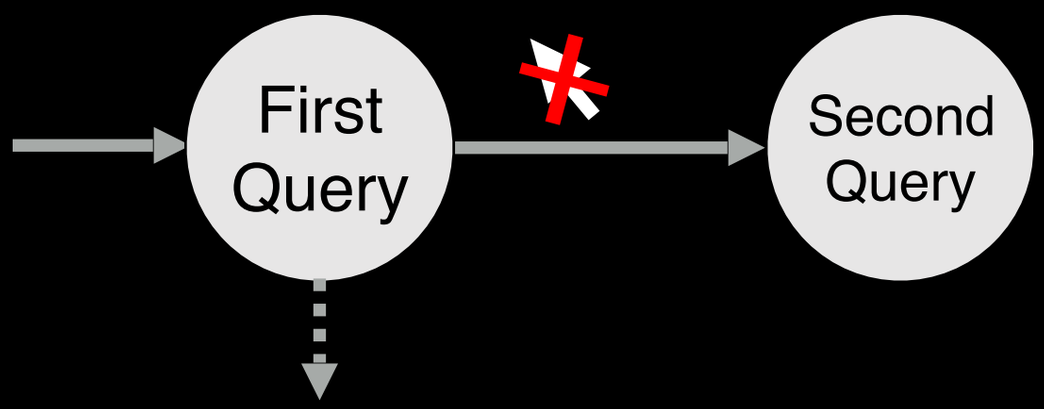
Identifying Struggling Sessions

bing

June 1–7, 2014

1. **Filter sessions:** US & English, start with a typed query
2. **Segment into tasks:** Topically coherent sub-sessions with at least 3 queries
3. **Filter struggling tasks**

[Hassan et al., WSDM'14]



No click or dwell time < 10s

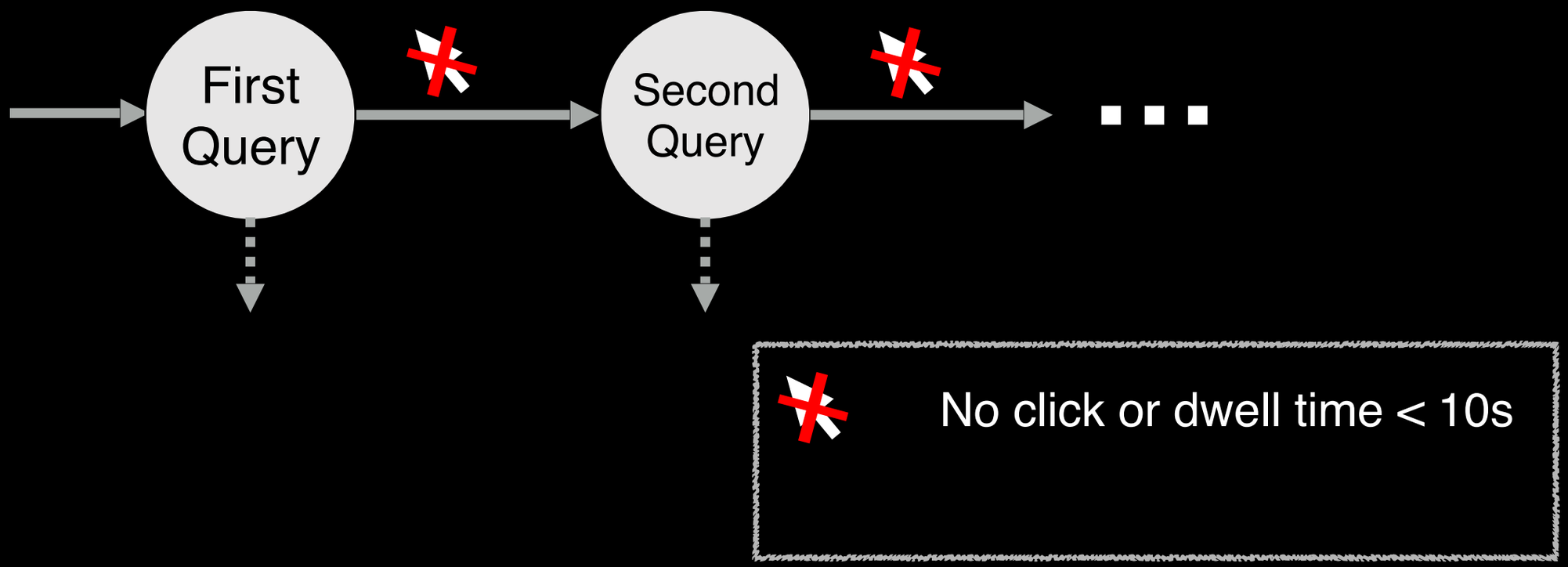
Identifying Struggling Sessions

bing

June 1–7, 2014

1. **Filter sessions:** US & English, start with a typed query
2. **Segment into tasks:** Topically coherent sub-sessions with at least 3 queries
3. **Filter struggling tasks**

[Hassan et al., WSDM'14]



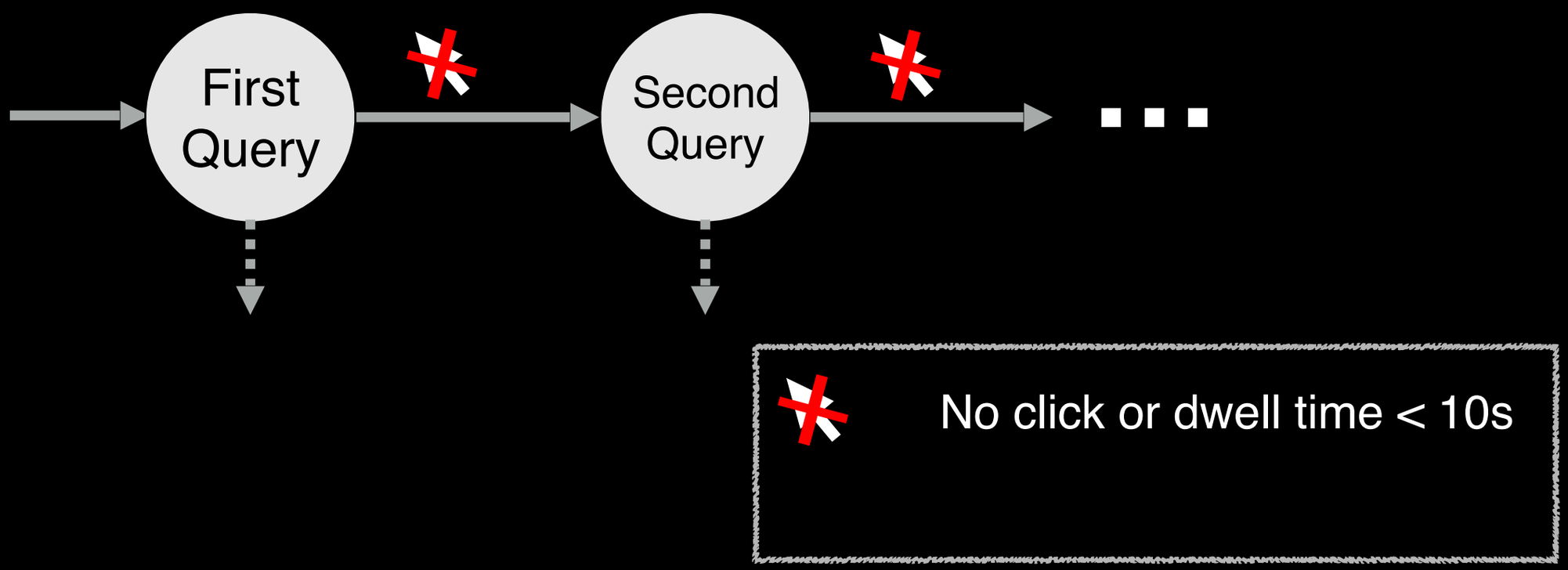
Identifying Struggling Sessions

bing

June 1–7, 2014

1. **Filter sessions:** US & English, start with a typed query
2. **Segment into tasks:** Topically coherent sub-sessions with at least 3 queries
3. **Filter struggling tasks** [Hassan et al., WSDM'14]
4. **Partition based on final clicks**

Commonly used proxy for success.



Identifying Struggling Sessions

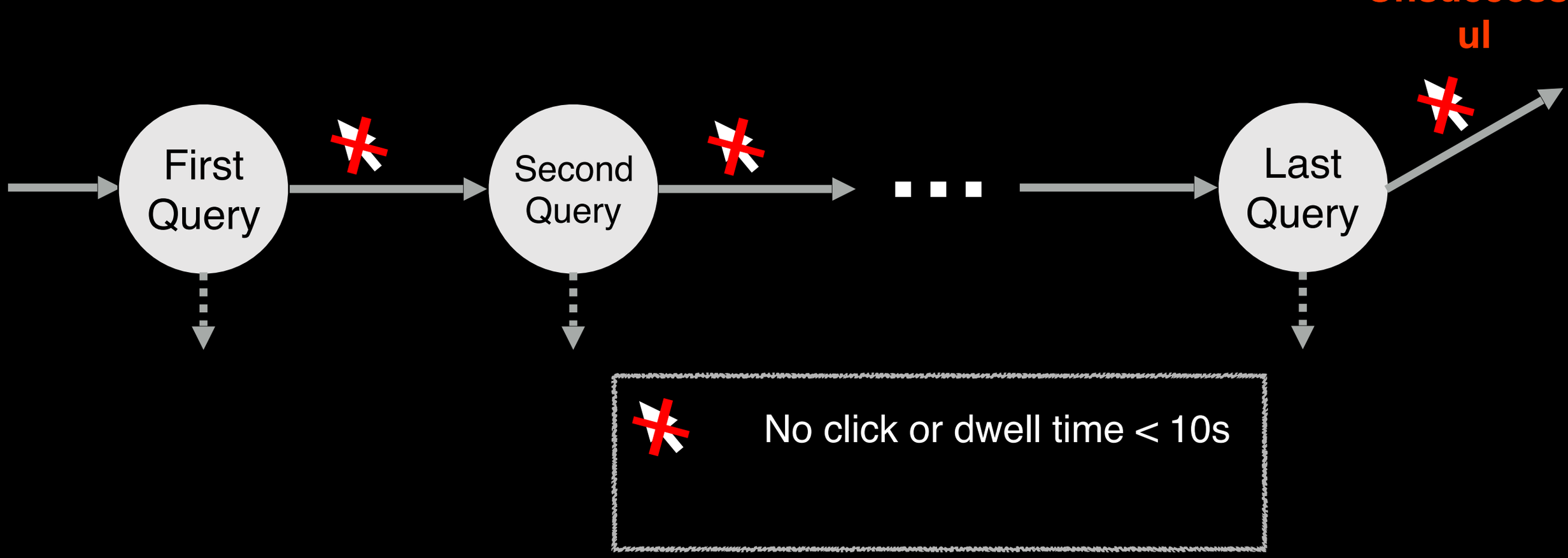
bing

June 1–7, 2014

1. **Filter sessions:** US & English, start with a typed query
2. **Segment into tasks:** Topically coherent sub-sessions with at least 3 queries
3. **Filter struggling tasks** [Hassan et al., WSDM'14]
4. **Partition based on final clicks**

Commonly used proxy for success.

4,508,821 Unsuccessful



Identifying Struggling Sessions

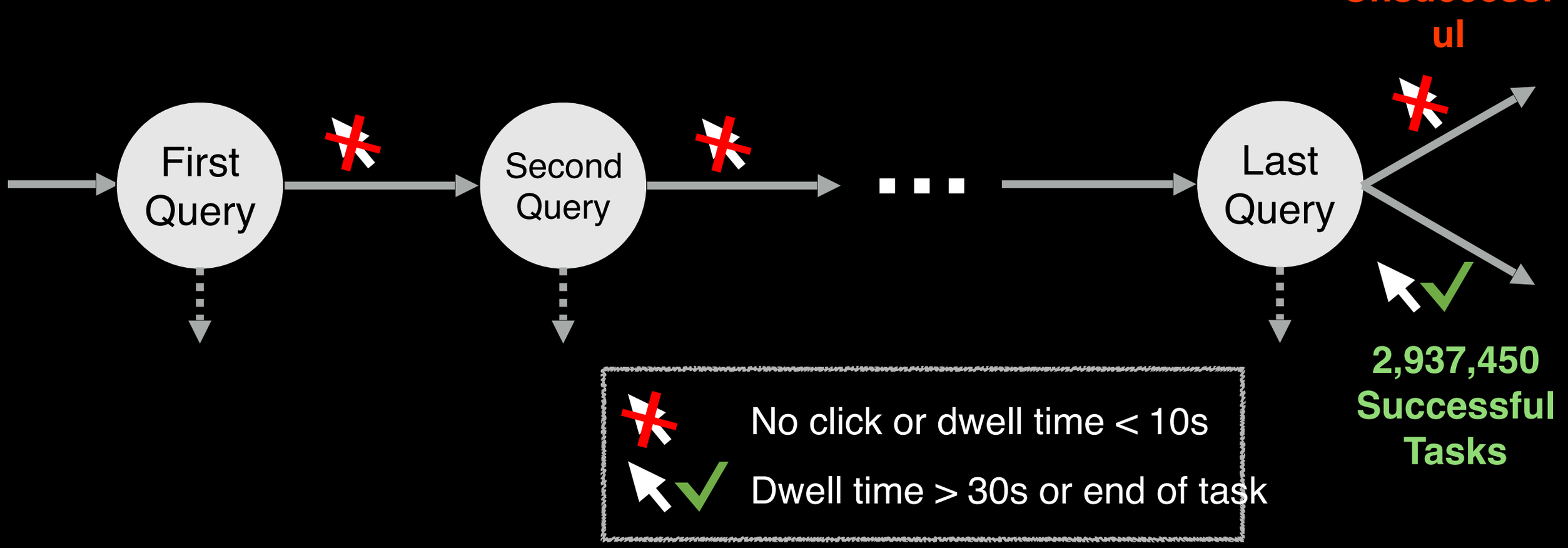
bing

June 1–7, 2014

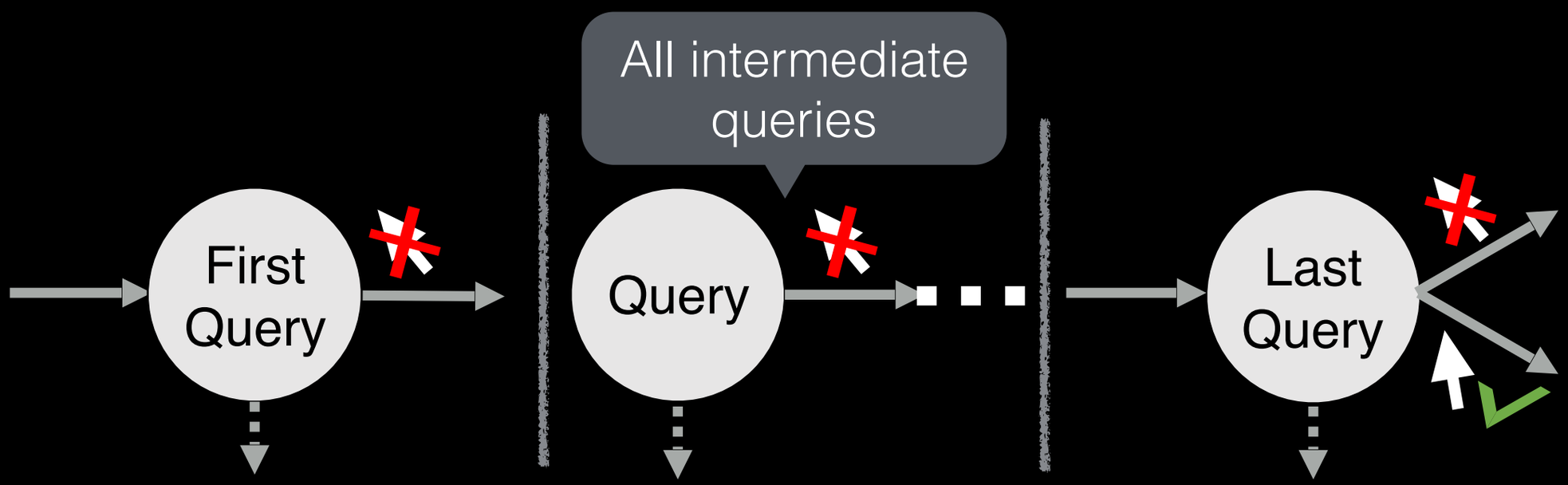
1. **Filter sessions:** US & English, start with a typed query
2. **Segment into tasks:** Topically coherent sub-sessions with at least 3 queries
3. **Filter struggling tasks** [Hassan et al., WSDM'14]
4. **Partition based on final clicks**

Commonly used proxy for success.

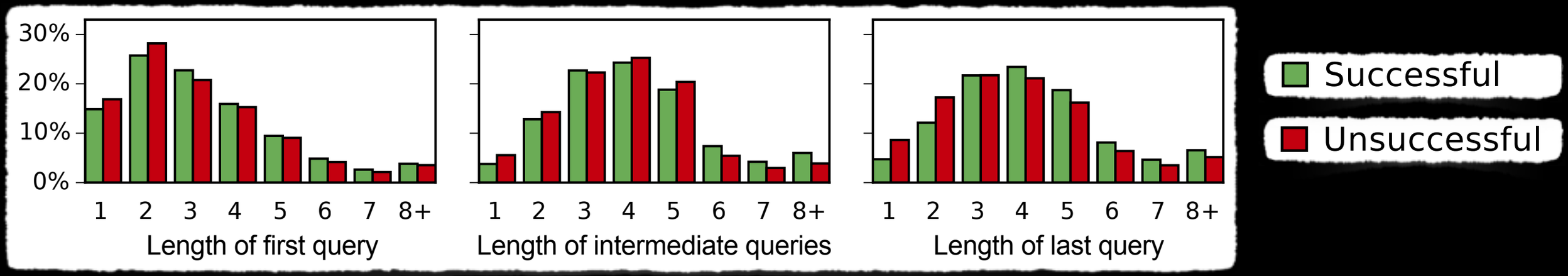
4,508,821 Unsuccessful Tasks



Query Grouping

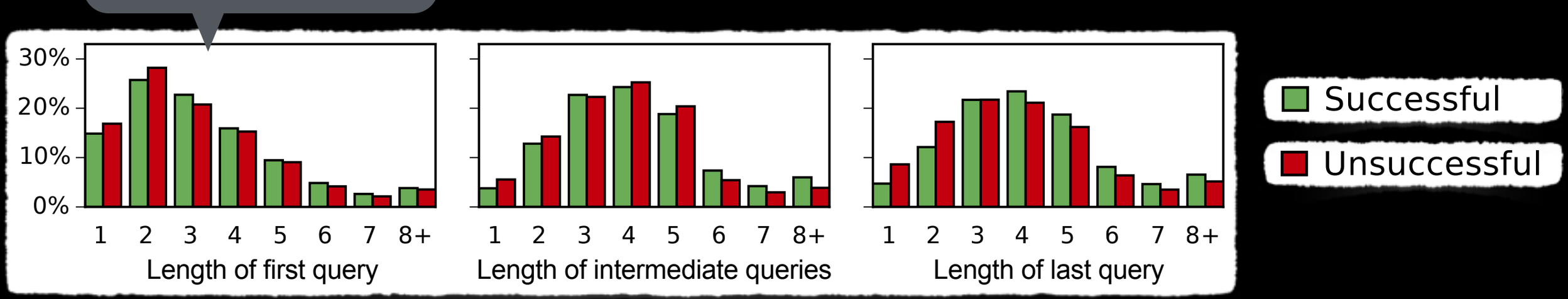


Queries become longer



Queries become longer

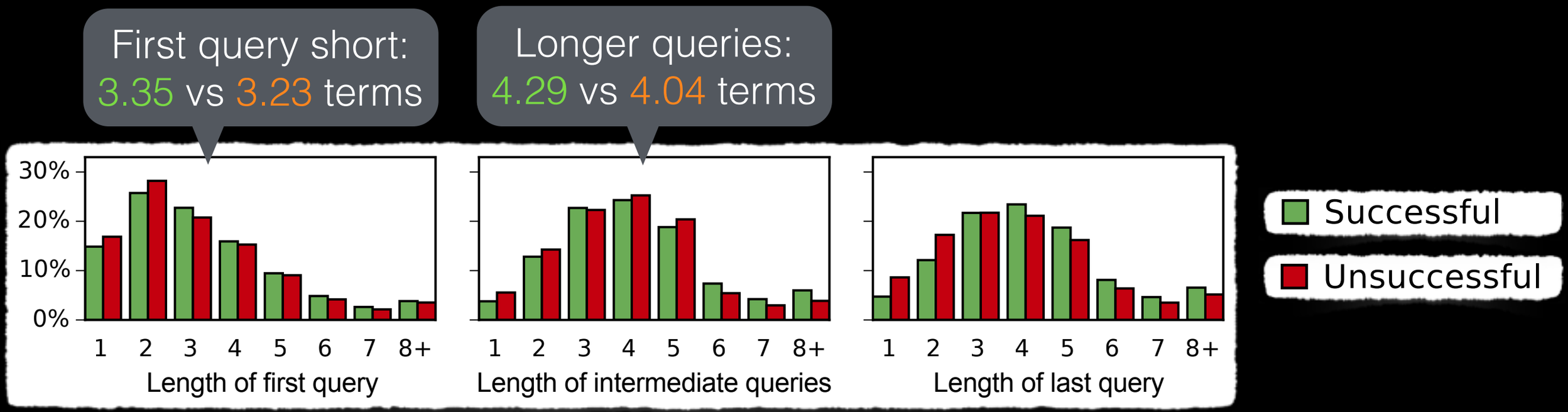
First query short:
3.35 vs 3.23 terms



Queries become longer

First query short:
3.35 vs 3.23 terms

Longer queries:
4.29 vs 4.04 terms

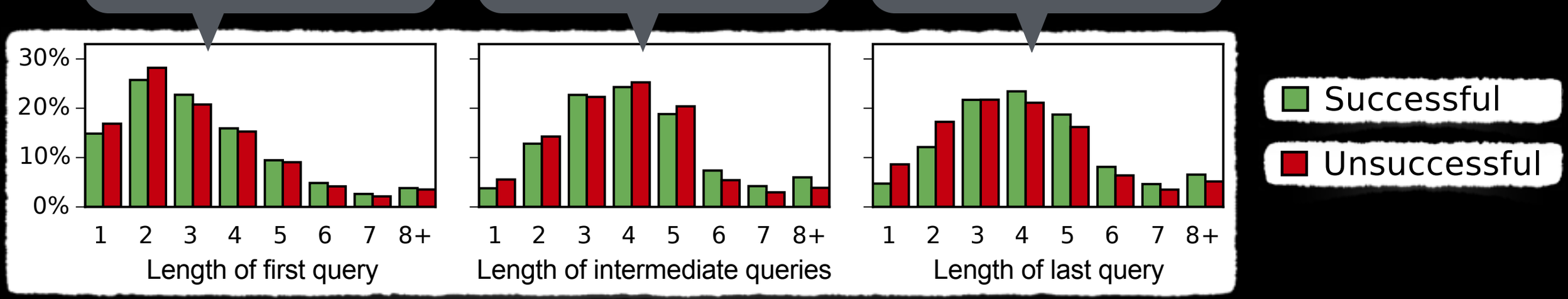


Queries become longer

First query short:
3.35 vs 3.23 terms

Longer queries:
4.29 vs 4.04 terms

Larger difference:
4.29 vs 3.93 terms

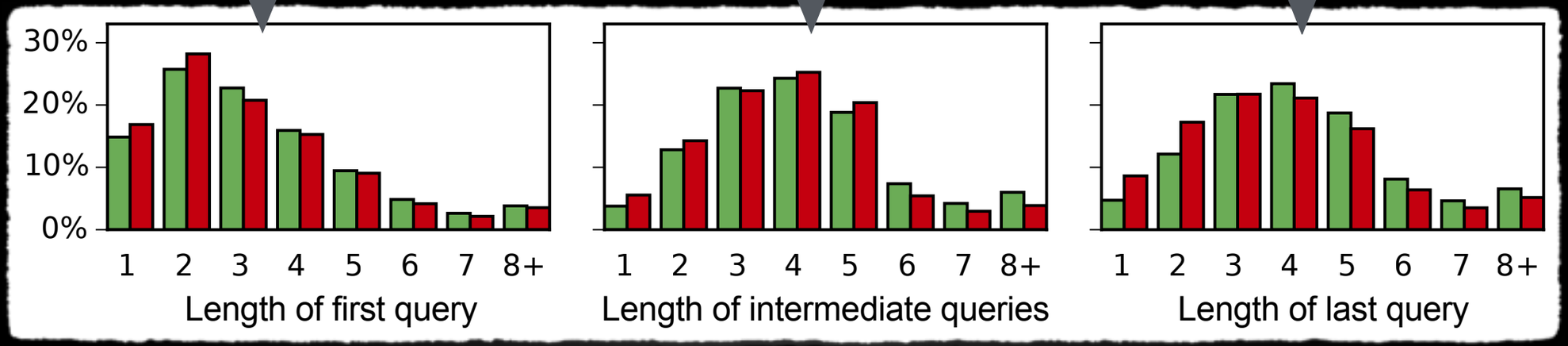


Queries become longer

First query short:
3.35 vs 3.23 terms

Longer queries:
4.29 vs 4.04 terms

Larger difference:
4.29 vs 3.93 terms



Successful

Unsuccessful

Shorter

Outline

- Characterizing Struggling

  Struggling searchers in successful vs unsuccessful tasks issue fewer and shorter queries. Query reformulations indicate more trouble choosing correct vocabulary.

- Annotating Query Transitions

- Predicting Future Actions

Outline

Characterizing Struggling

Struggling searchers in successful vs unsuccessful tasks issue fewer and shorter queries. Query reformulations indicate more trouble choosing correct vocabulary.

Annotating Query Transitions

Crowd-sourcing to better understand: connection between struggling and success & query transitions.
First an exploratory pilot, then main annotations.

Predicting Future Actions

Intent-based query reformulation taxonomy

Based on the answers to the open-ended questions

Added, removed □ an **action** (e.g., download, contact)
or substituted □ an **attribute** (e.g., printable, free,)

Specified □ a **particular instance**
(e.g. added a brand name or version number)

Rephrased □ Corrected a spelling error or **typo**

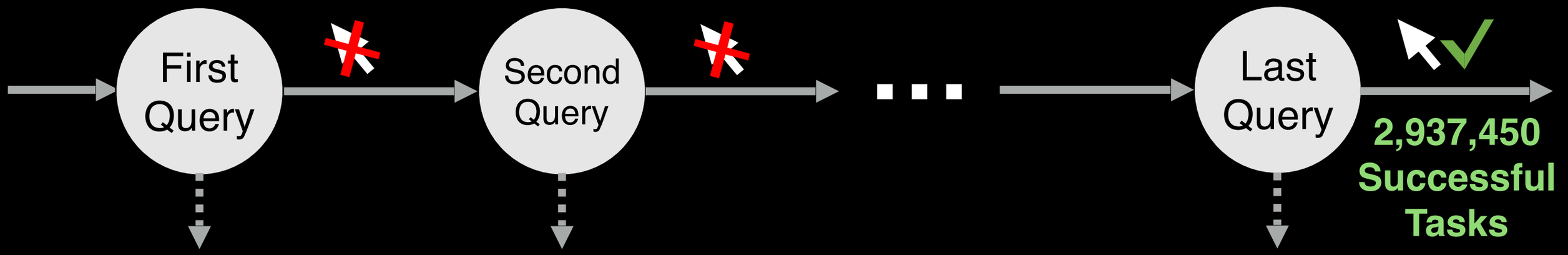
□ Used a **synonym** or related term

Switched □ to a **related** task (changed main focus)

□ to a **new task**

Main Annotations:

Transitions in 659 successful tasks



- Annotated 659 new successful tasks
  - Single tasks, no pairs
  - Annotated transitions between each query pair

Understanding Struggling Behavior

Struggling describes a situation whereby a user experiences difficulty in finding the information they are looking for.

Below is a session in which we believe that a Bing user could not immediately find the information they were looking for. The session is presented as a series of Bing Queries ( ) and Result Clicks ( ) with their timestamps. You can click on a query to repeat a search or click on a result click to inspect the page that the user viewed (the page will open in a new window).

Please examine the session shown below and then answer the questions about it.

Timestamp	Query/URL
9:13:24 AM	us open golf
9:13:36 AM	us open golf 2013 live
9:13:59 AM	watch us open live streaming
9:14:02 AM	inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video
9:31:55 AM	END

How successful do you think the user was in this session?

○ Not at all successful ○ Somewhat successful ○ Mostly successful ○ Completely successful

Understanding Struggling Behavior

Struggling describes a situation whereby a user experiences difficulty in finding the information they are looking for.

Below is a session in which we believe that a Bing user could not immediately find the information they were looking for. The session is presented as a series of Bing Queries ( ) and Result Clicks ( ) with their timestamps. You can click on a query to repeat a search or click on a result click to inspect the page that the user viewed (the page will open in a new window).

Please examine the session shown below and then answer the questions about it.

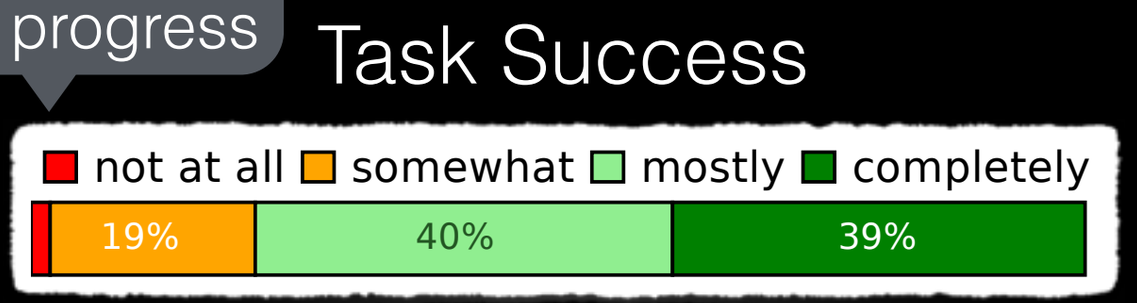
Timestamp	Query/URL
9:13:24 AM	us open golf
9:13:36 AM	us open golf 2013 live
9:13:59 AM	watch us open live streaming
9:14:02 AM	inquisitr.com/1300340/watch-2014-u-s-open-live-online-final-round-free-streaming-video
9:31:55 AM	END

How successful do you think the user was in this session?

○ Not at all successful ○ Somewhat successful ○ Mostly successful ○ Completely successful

Success & Pivotal Query

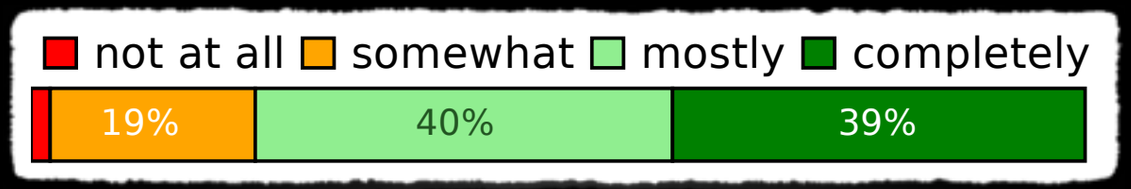
Made at least some progress



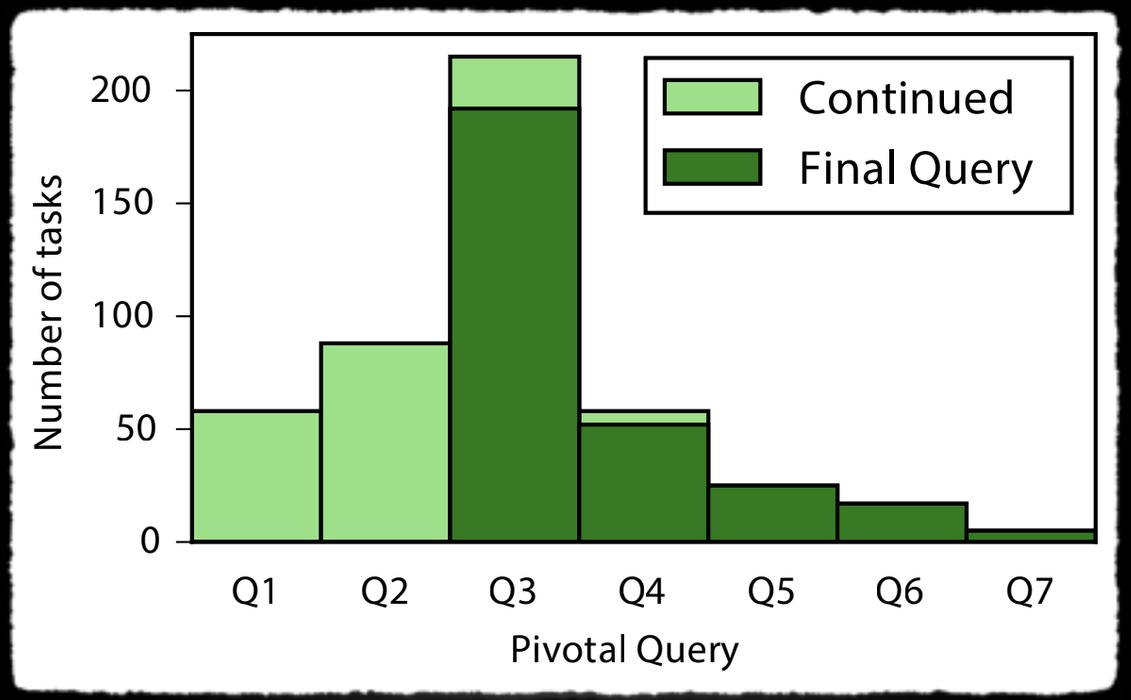
Success & Pivotal Query

Made at least some progress

Task Success



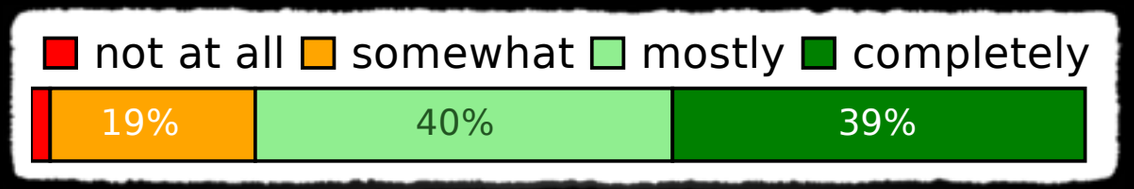
Location of the pivotal query



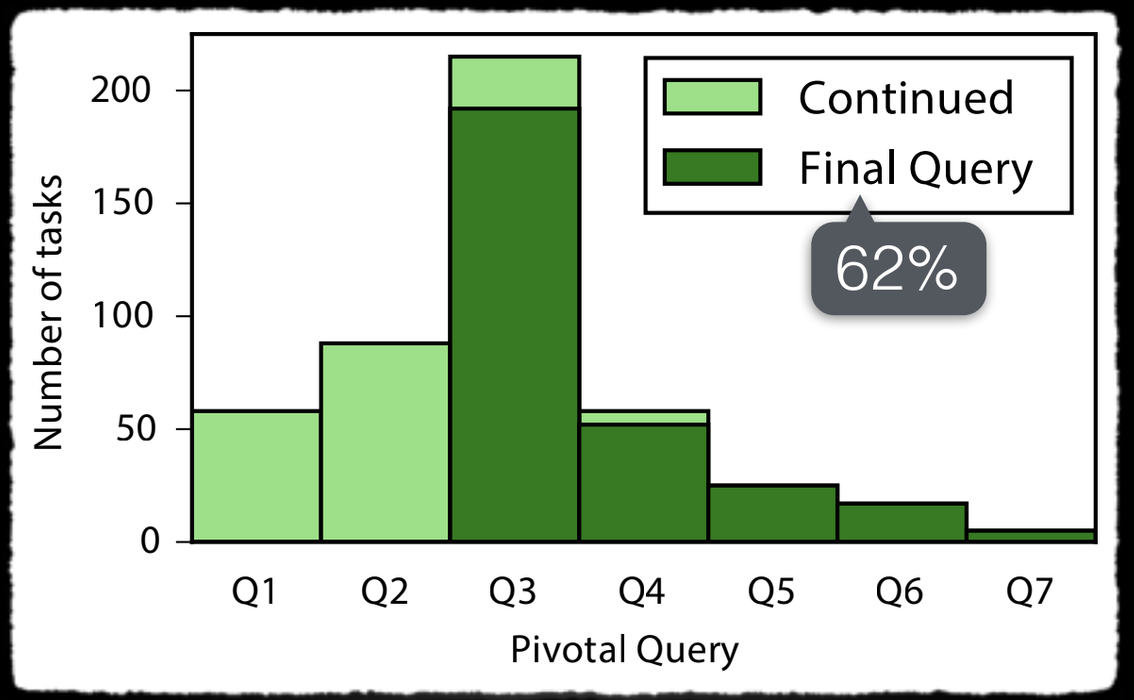
Success & Pivotal Query

Made at least some progress

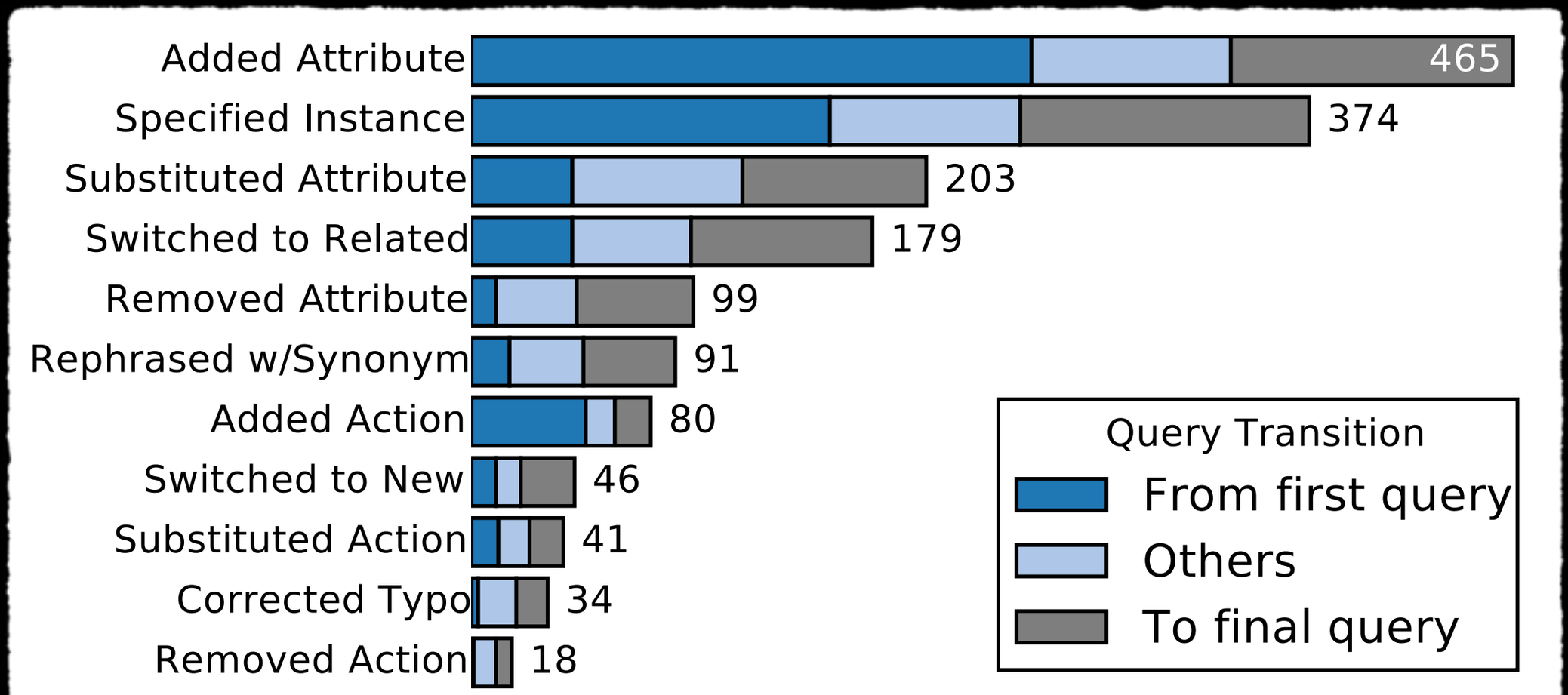
Task Success



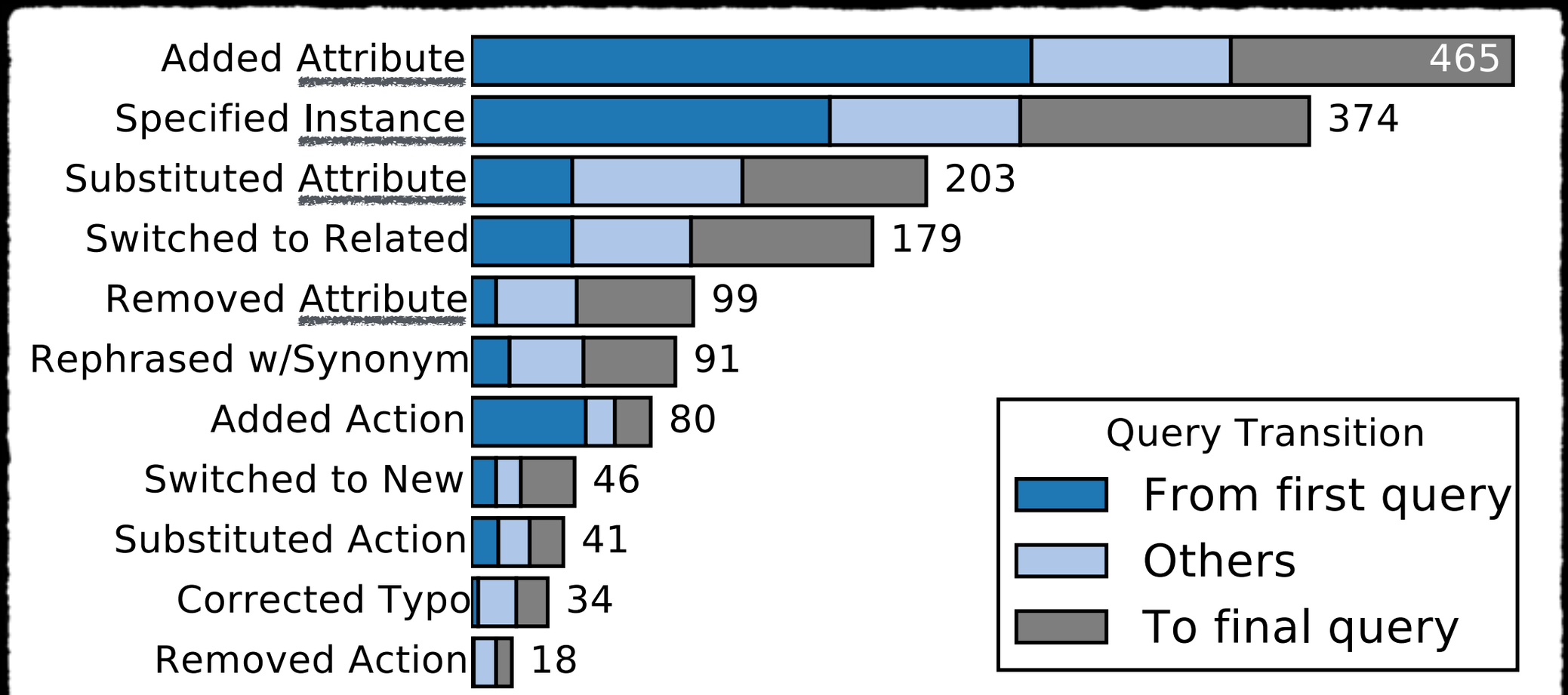
Location of the pivotal query



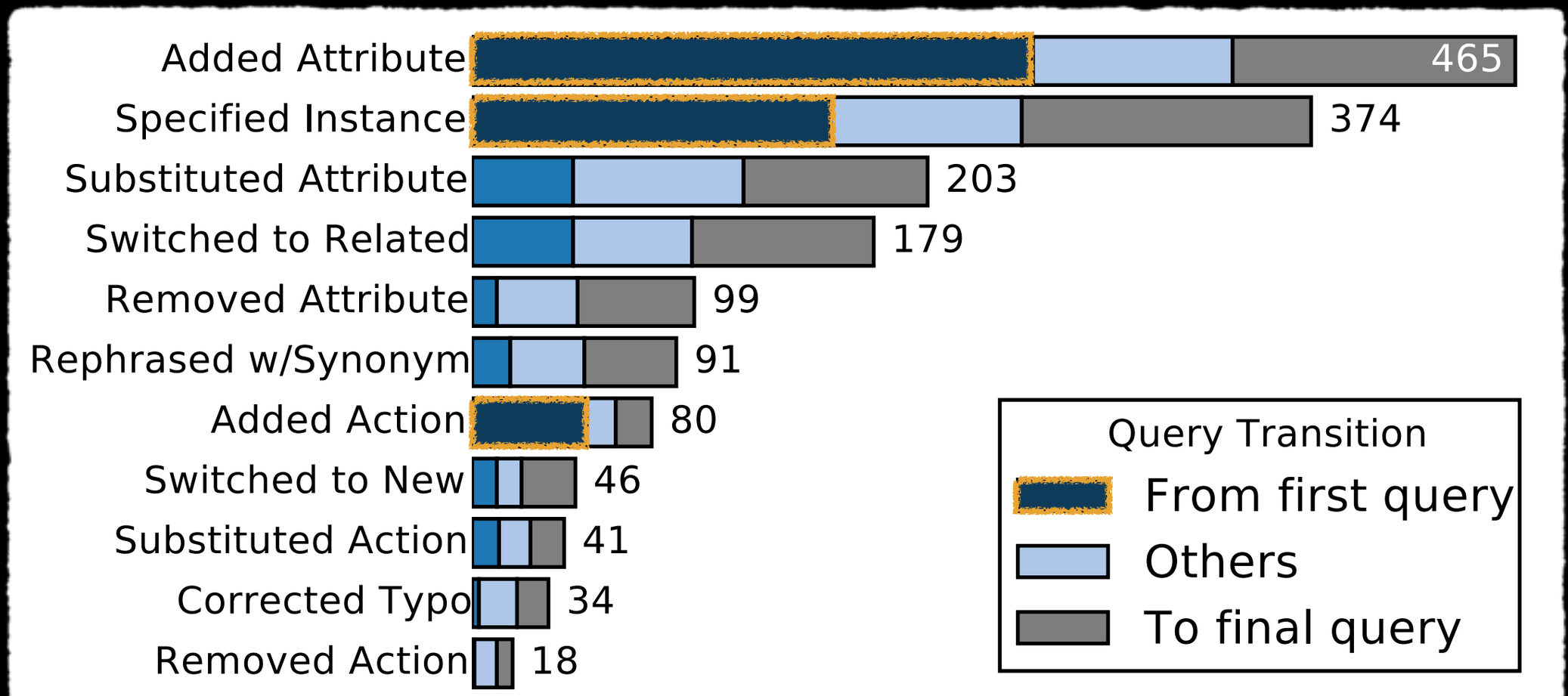
Query Transitions



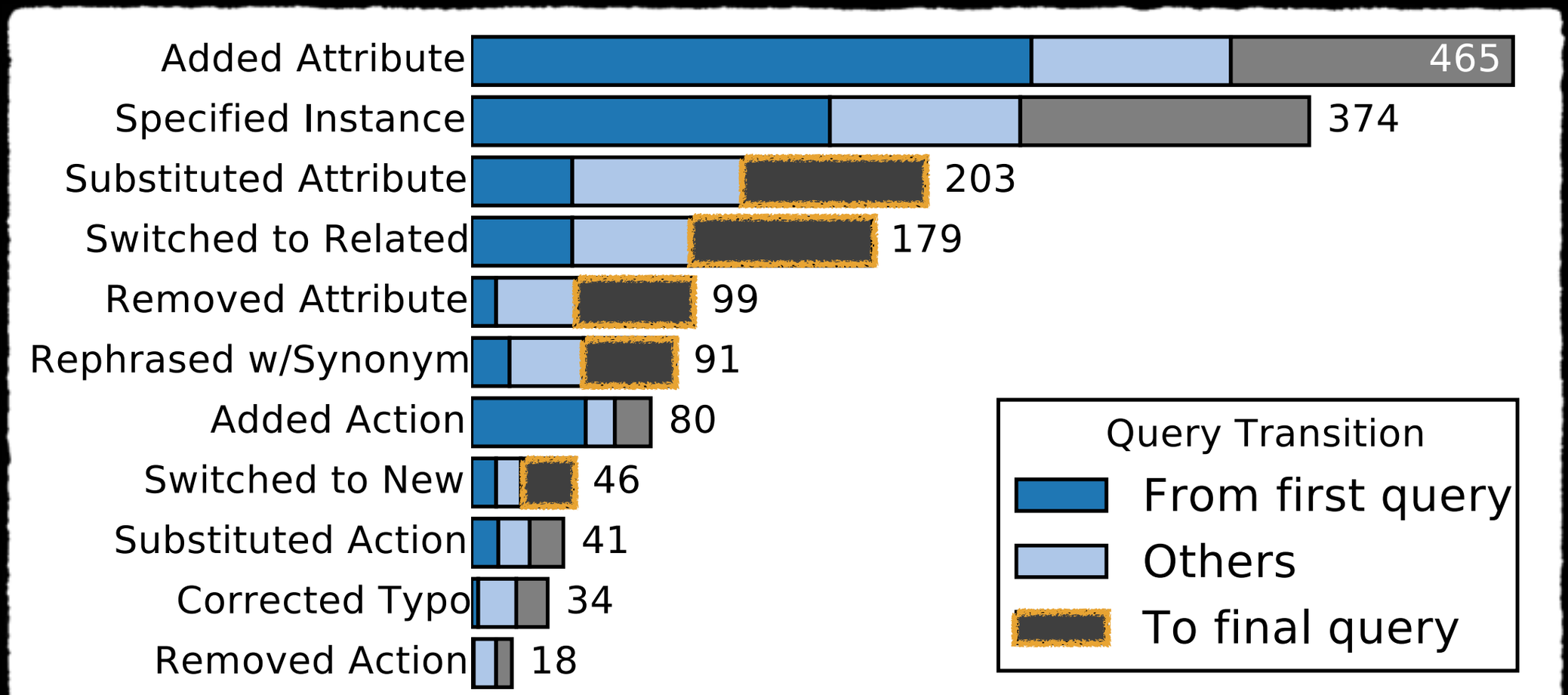
Query Transitions



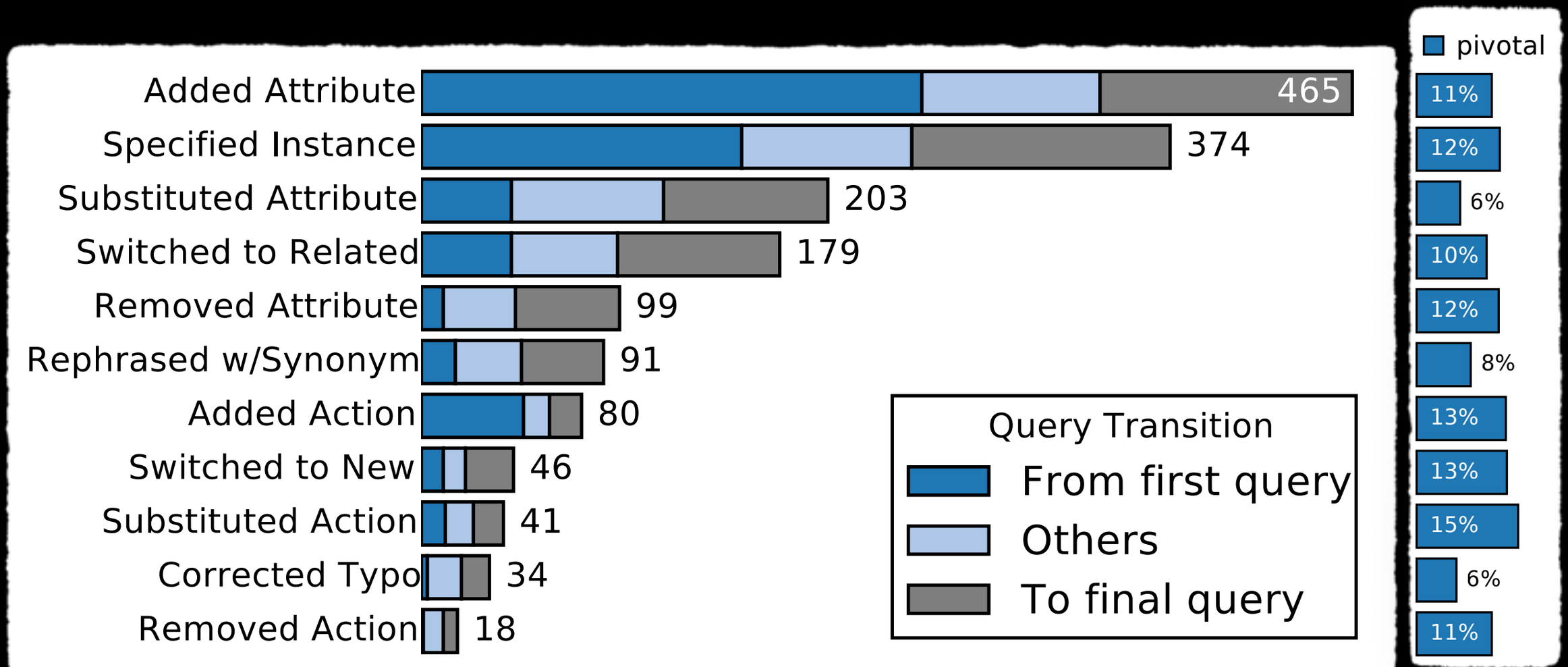
Query Transitions



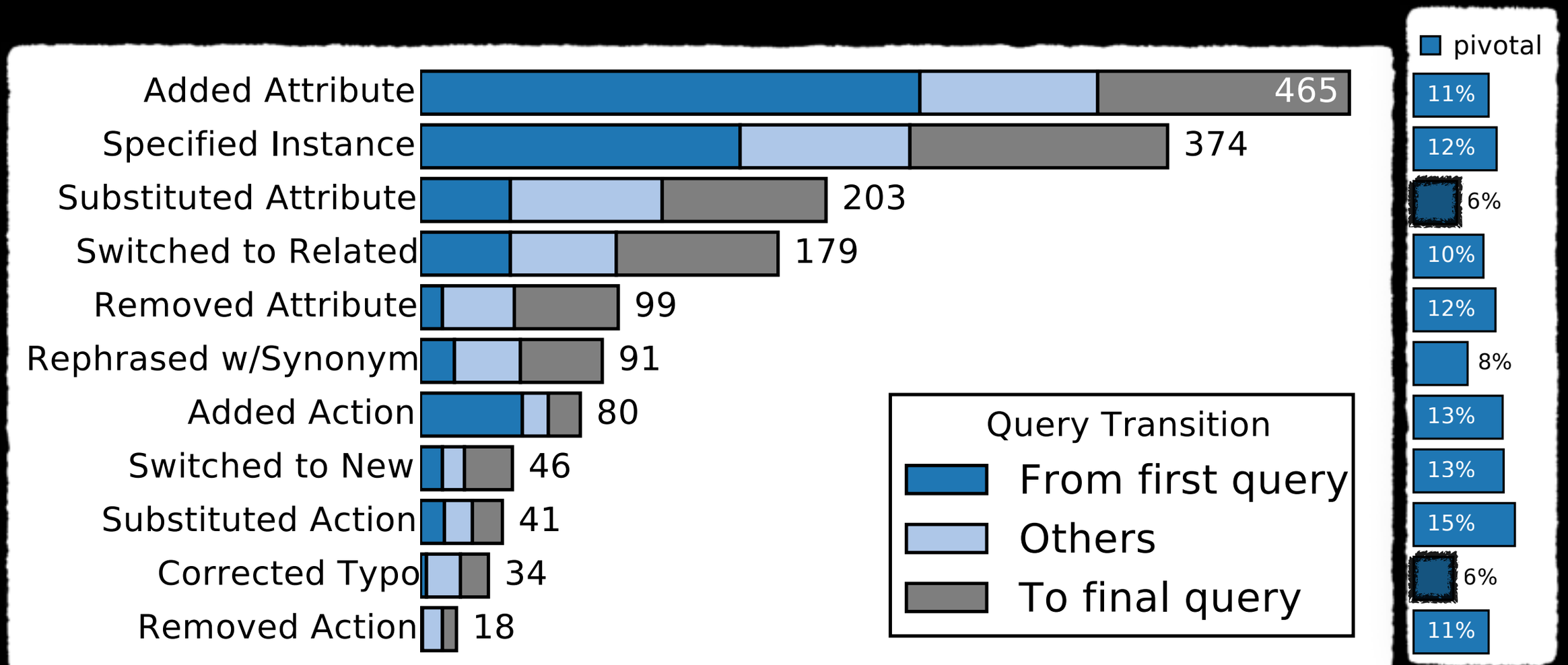
Query Transitions



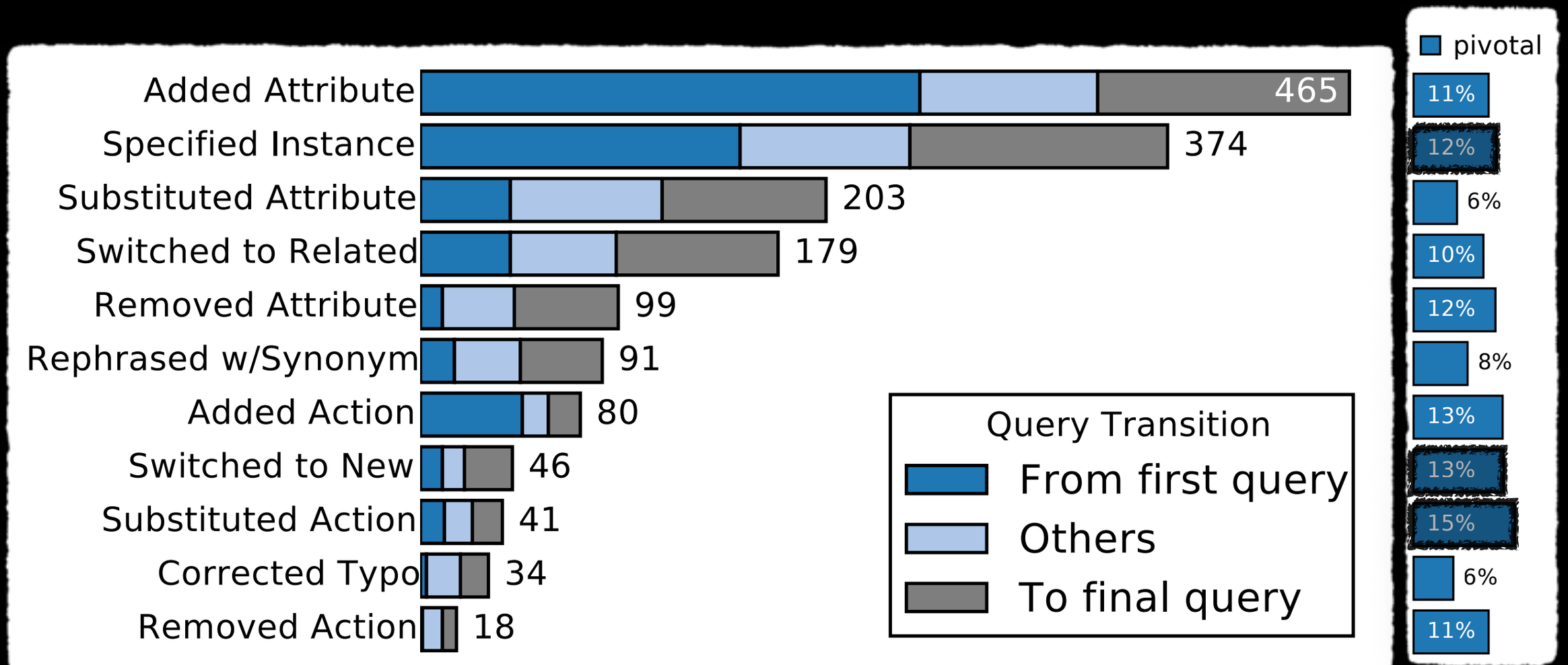
Query Transitions



Query Transitions



Query Transitions



Outline

1. Characterizing Struggling

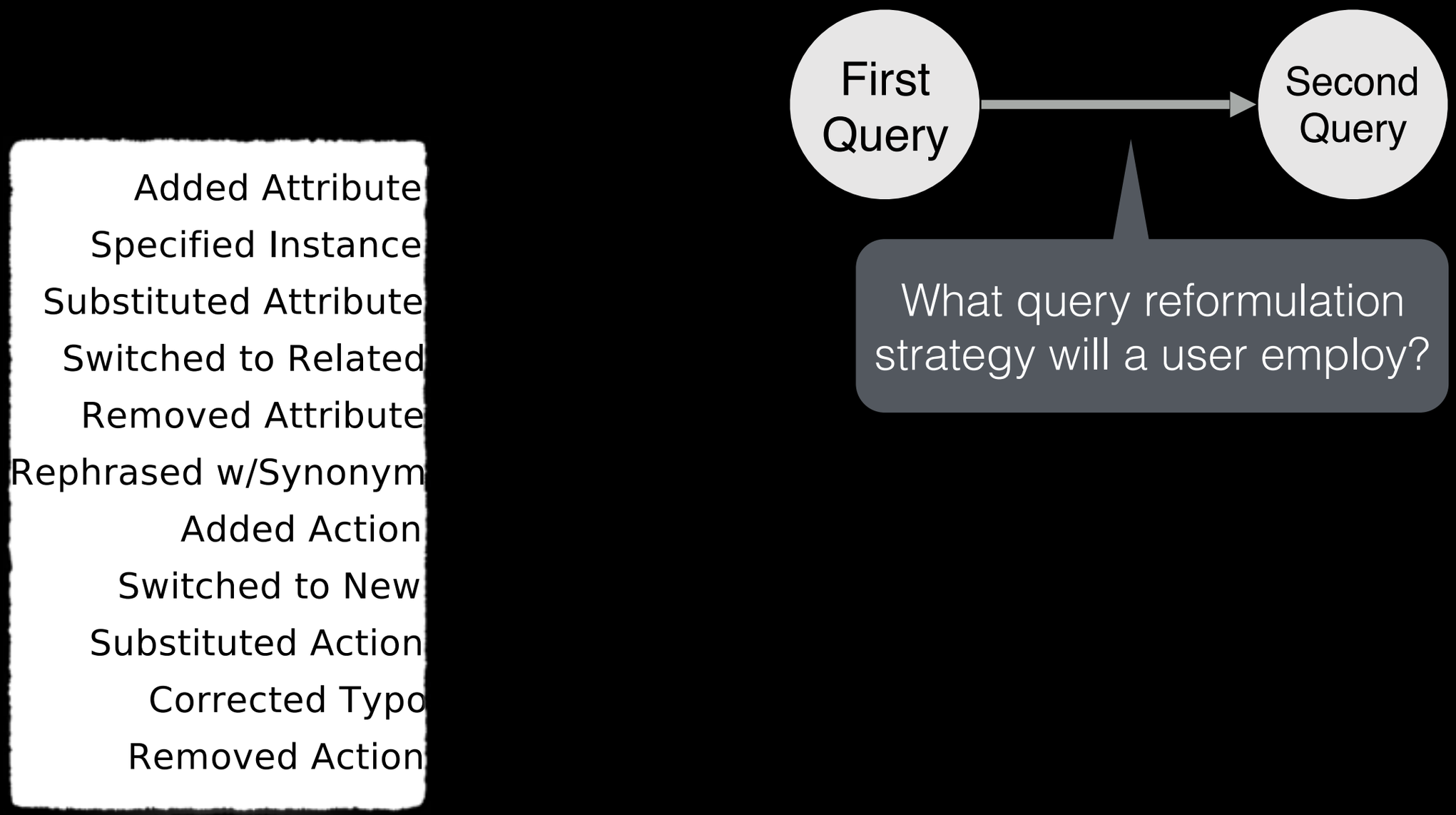
   Struggling searchers, successful or not, behave differently on many search aspects, including: queries, reformulations, clicks, dwell time & topic.

2. Annotating Query Transitions

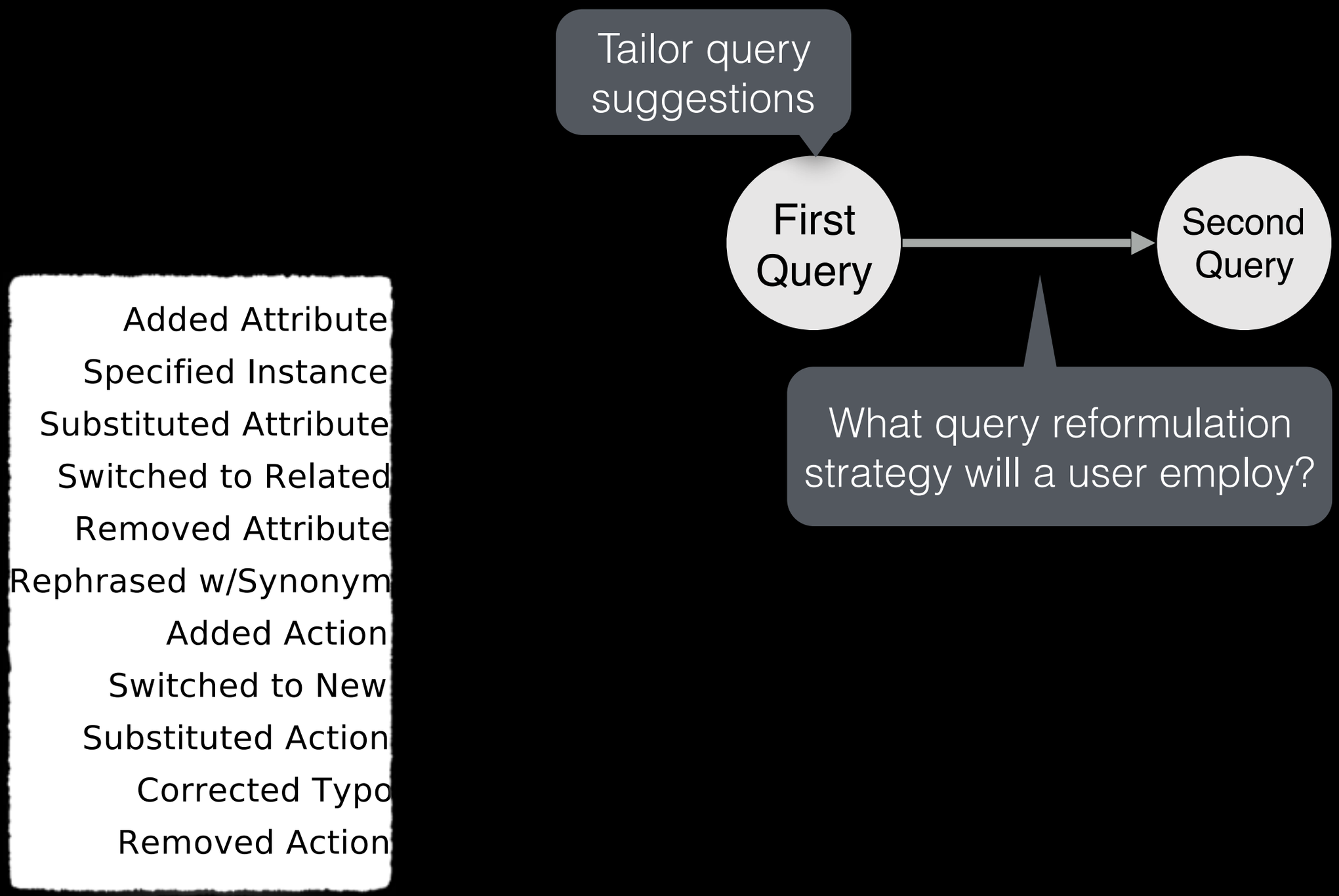
   Substantial differences in how searchers refine queries in different stages in a struggling task. Strong connections with task outcomes, and particular pivotal queries.

3. Predicting Future Actions

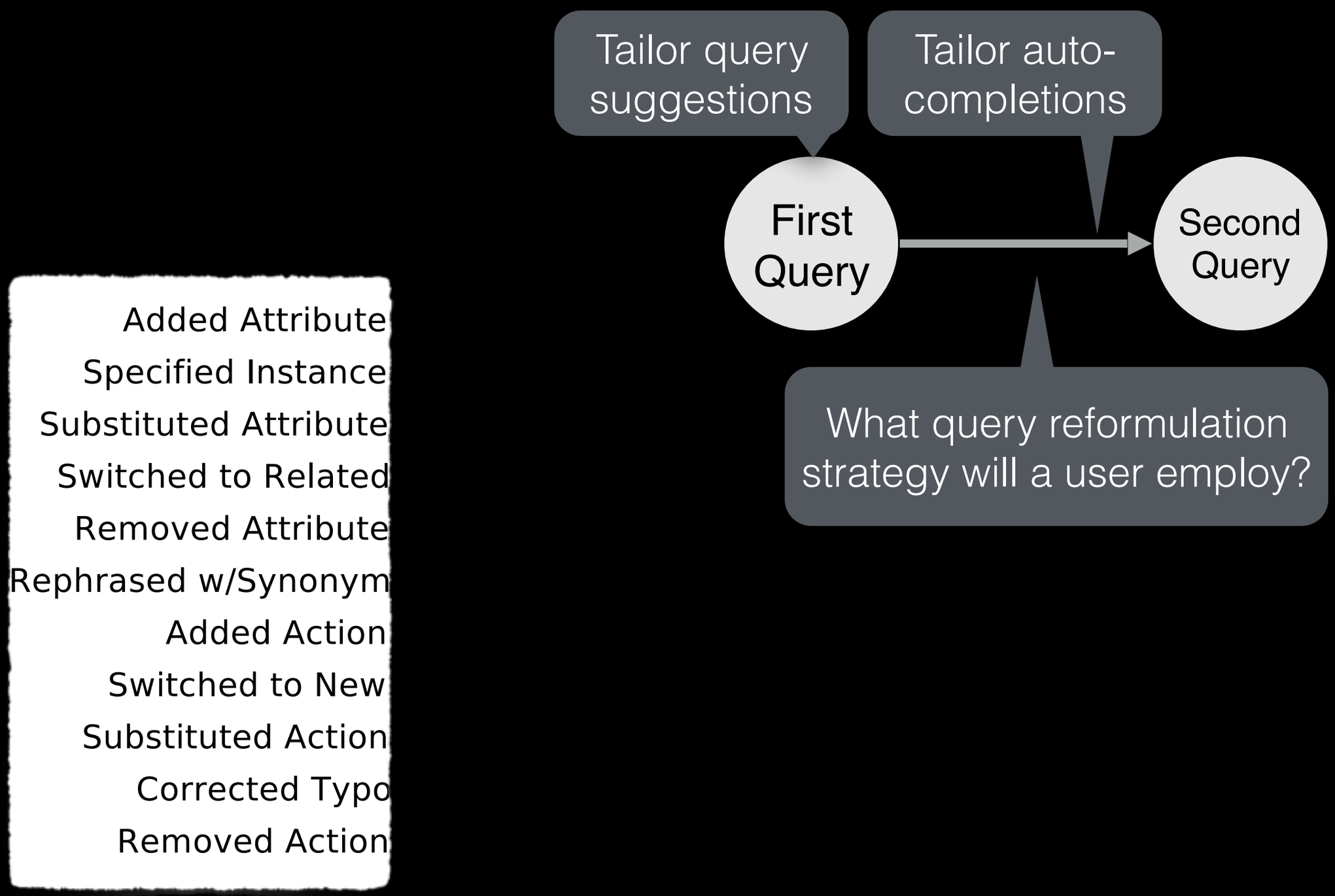
Predicting Reformulation



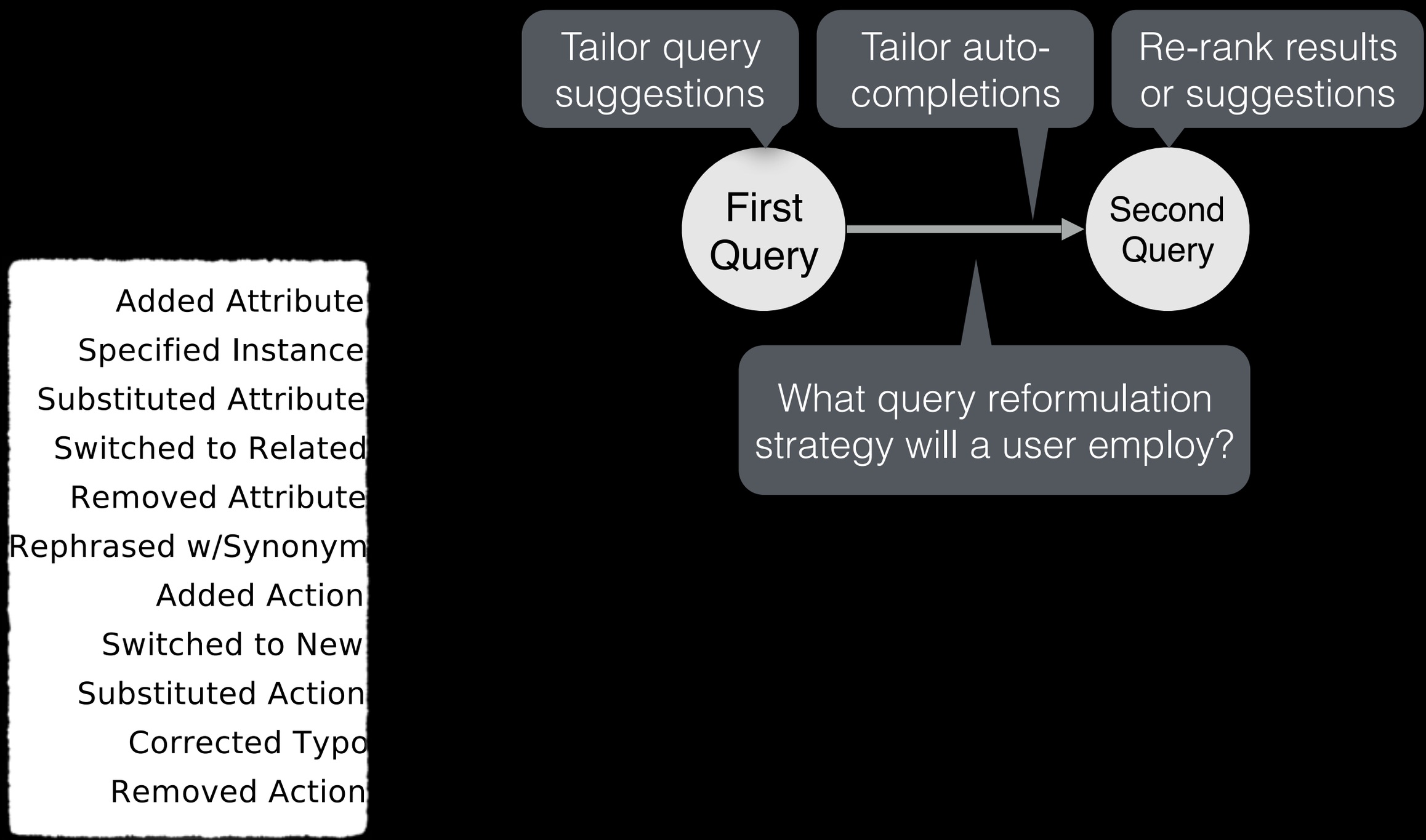
Predicting Reformulation



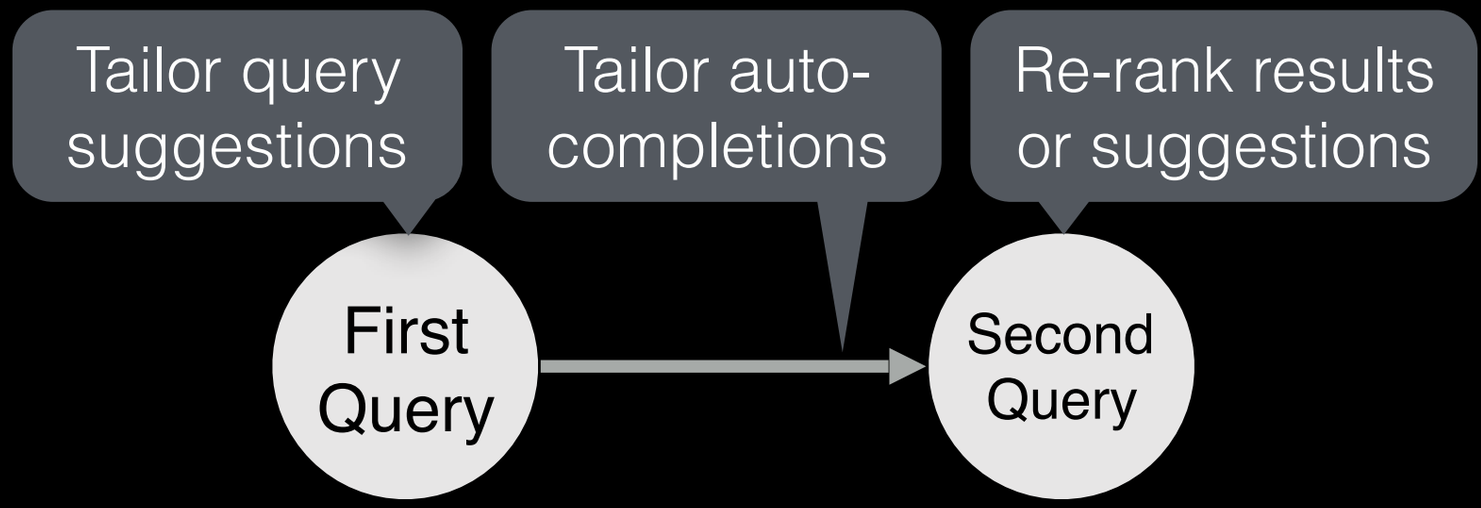
Predicting Reformulation



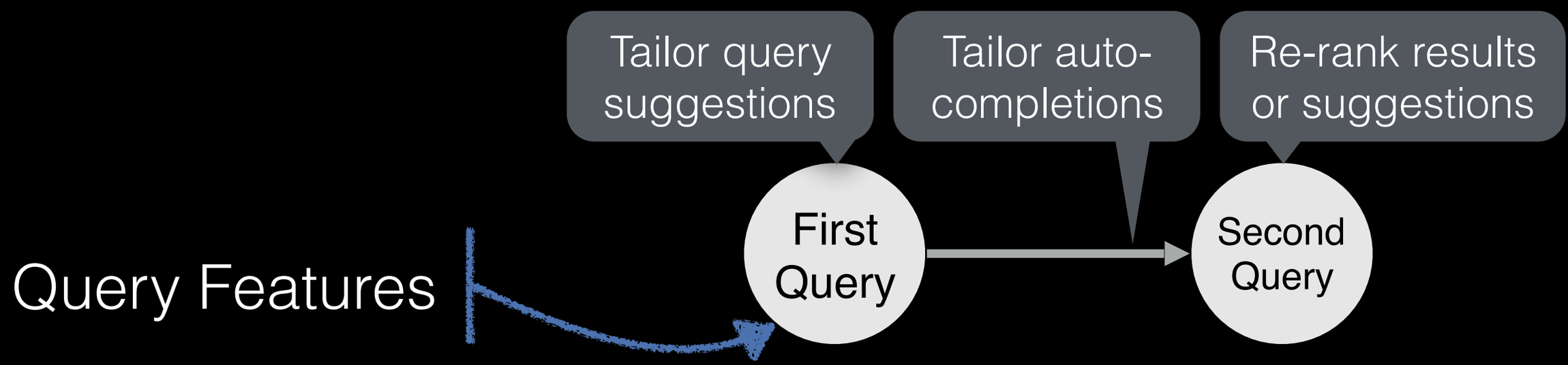
Predicting Reformulation



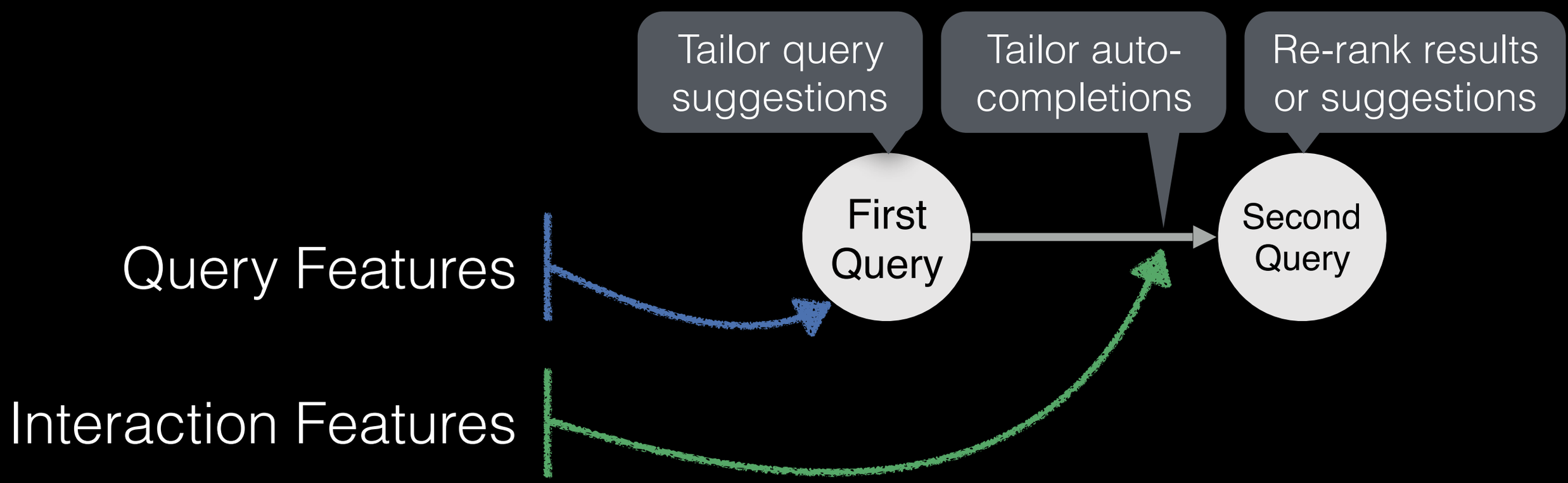
Feature Sets



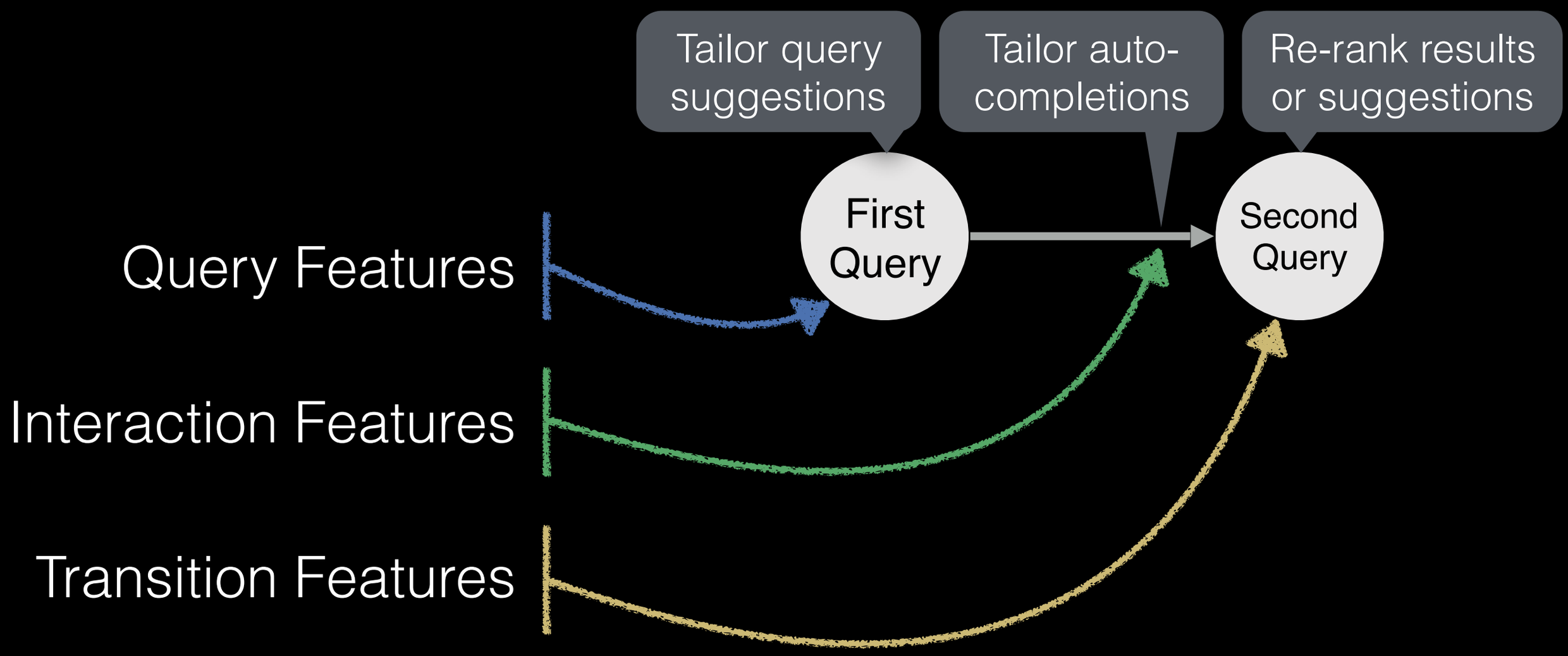
Feature Sets



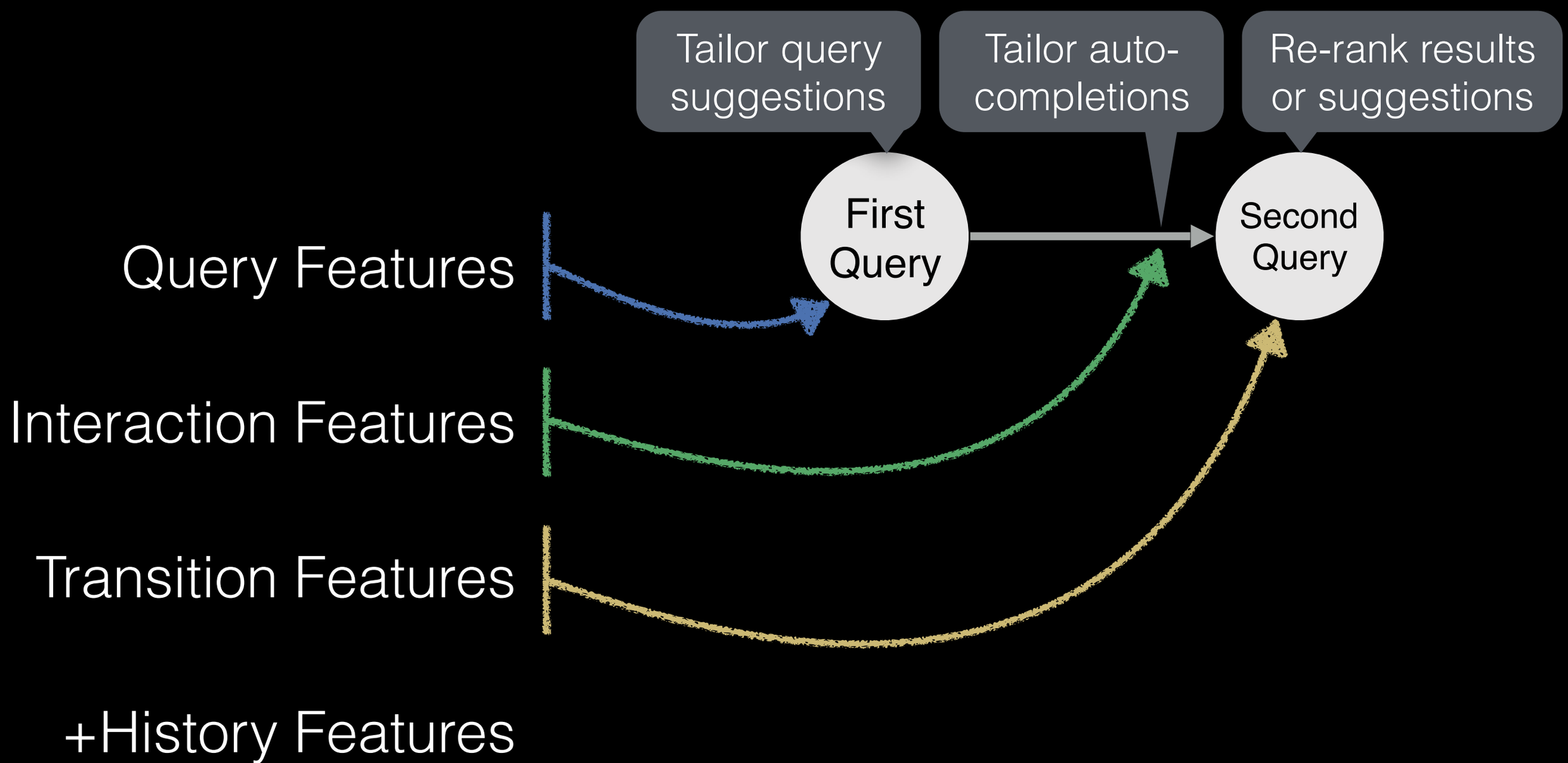
Feature Sets



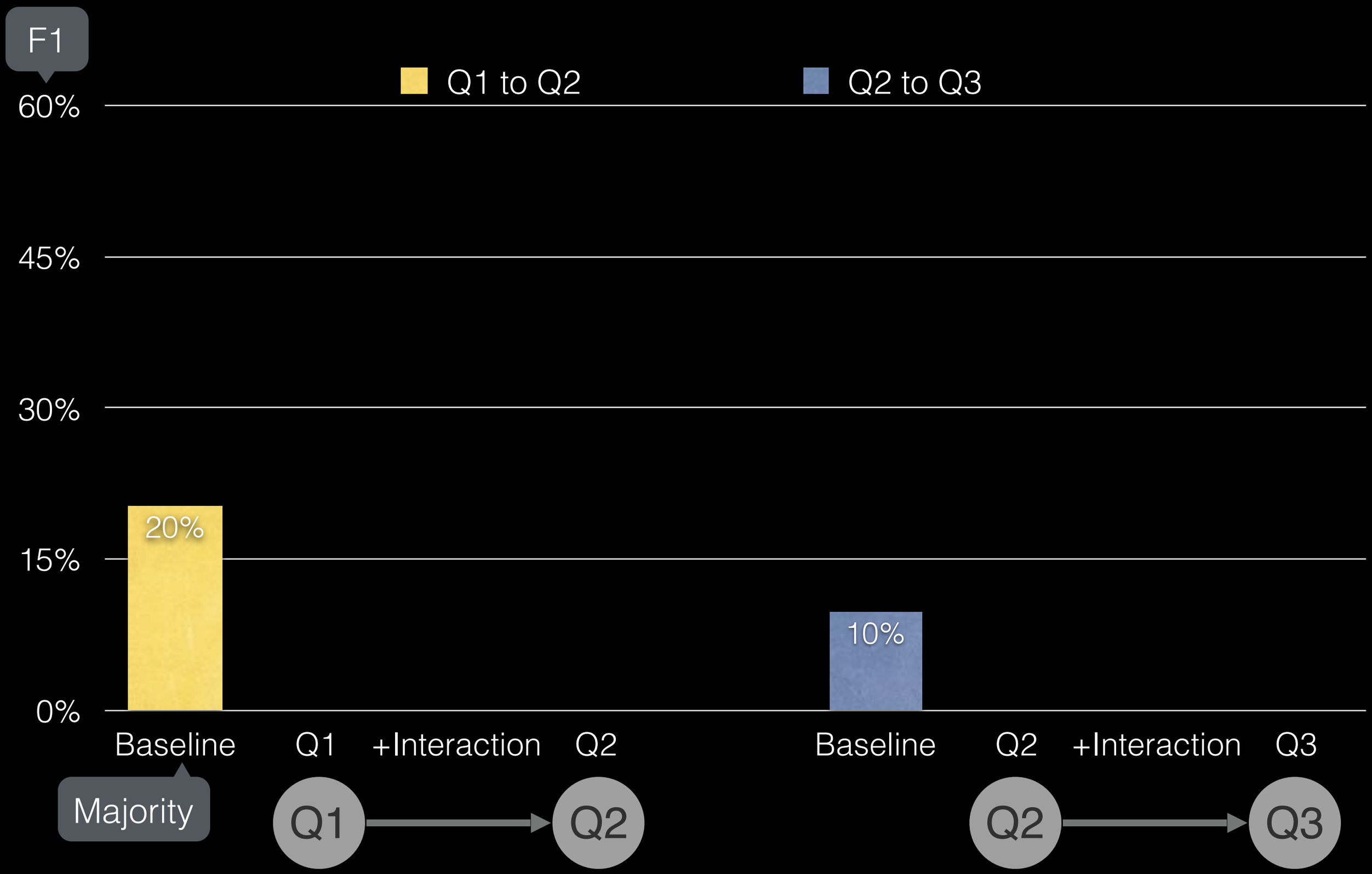
Feature Sets



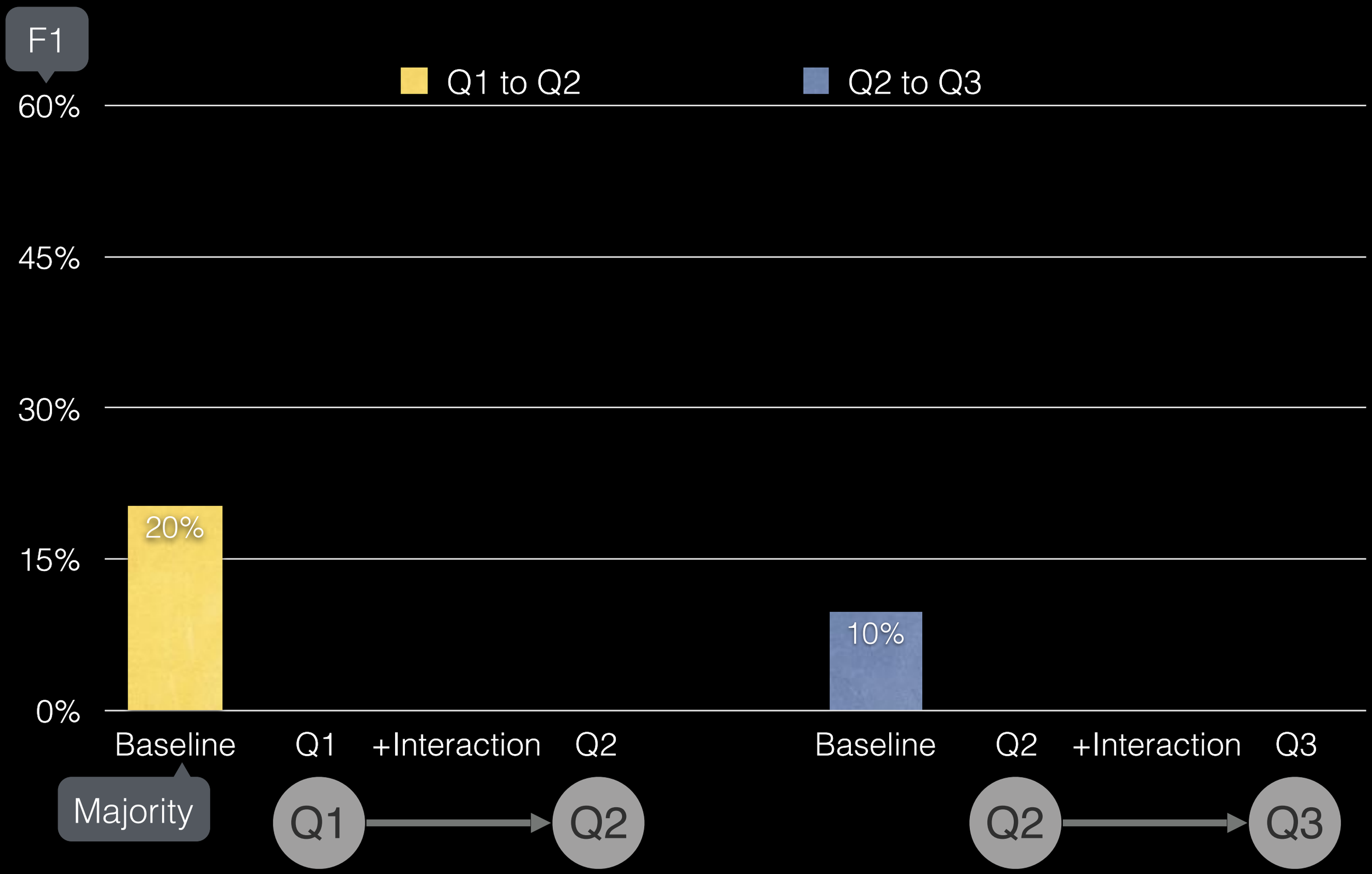
Feature Sets



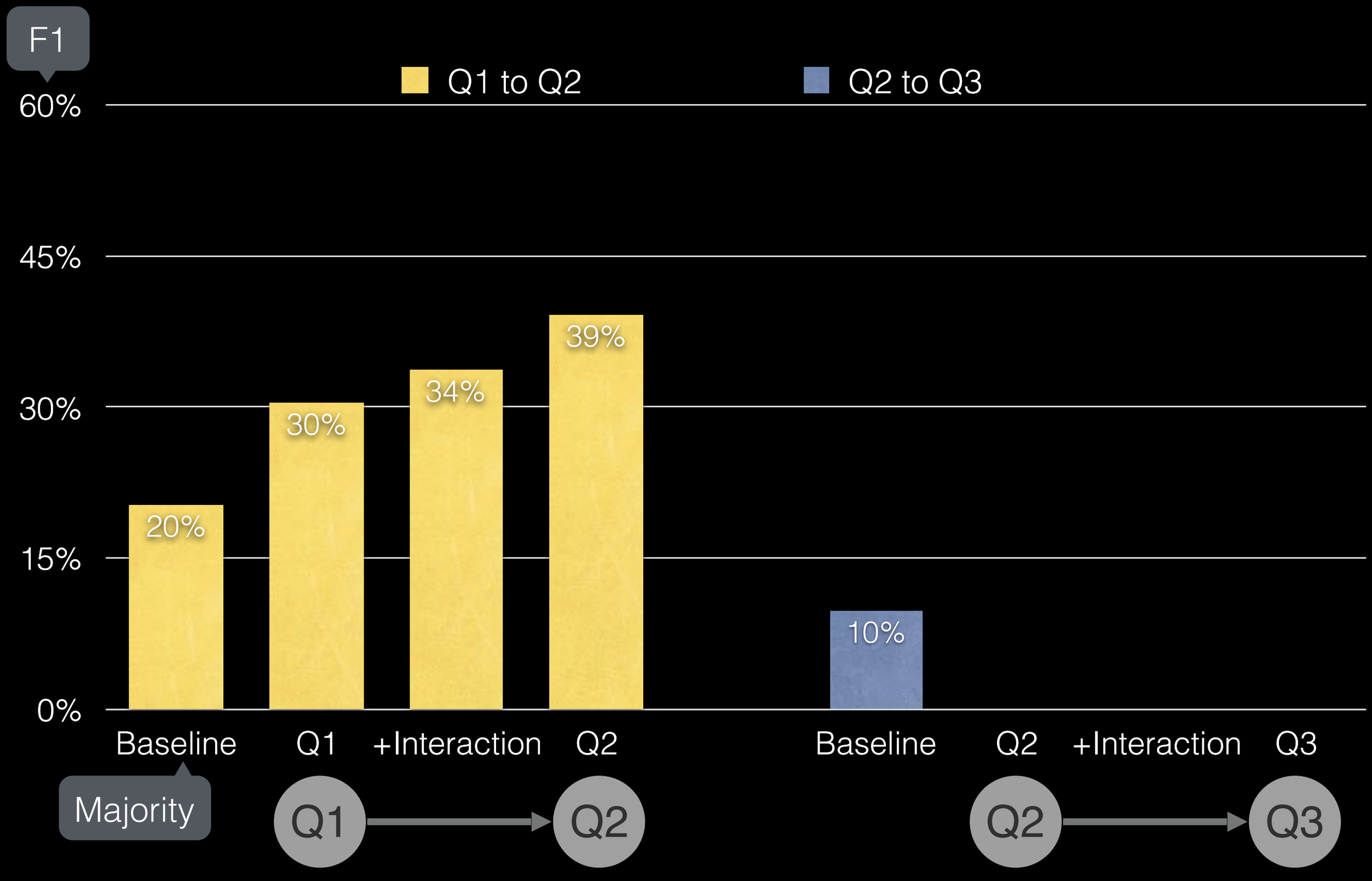
Prediction Results



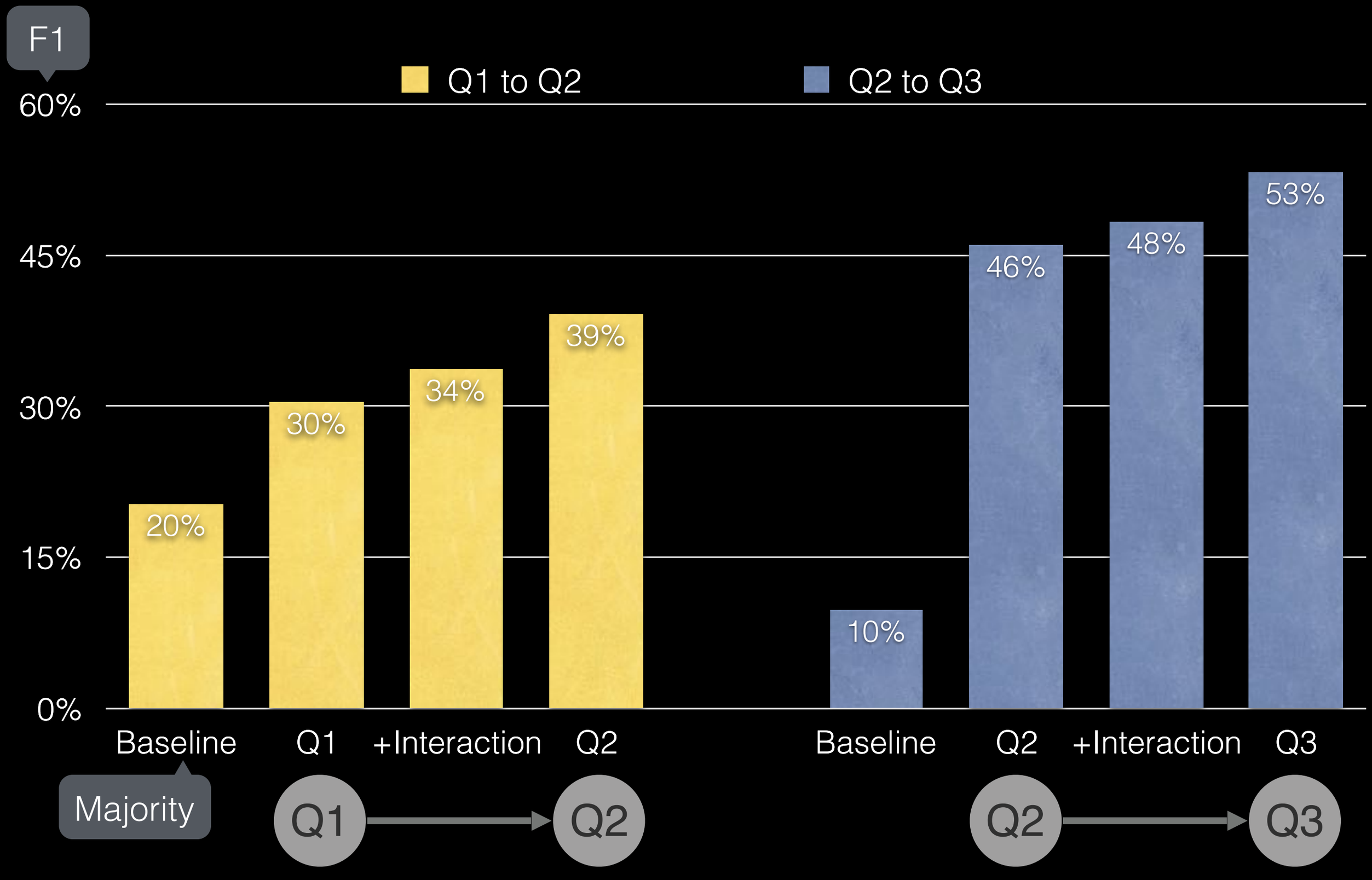
Prediction Results



Prediction Results



Prediction Results



Outline

- Characterizing Struggling

  Substantial differences in how searchers behave in a struggling task, depending on task outcomes.

- Annotating Query Transitions

  Accurately predict reformulation strategy to provide situation-specific support at a higher level.

- Predicting Future Actions

Limitations

- Very specific and apparent type of struggling

Limitations

- Very specific and apparent type of struggling
- Determination of search success
  - Final click as a proxy
  - Judgements by third-party, not by searchers themselves

Future Work

- Directly applying reformulation strategies
  - Query suggestions & auto-completions

Future Work

- Directly applying reformulation strategies
  - Query suggestions & auto-completions
  - Mining query ➢ pivotal query pairs

Future Work

- Directly applying reformulation strategies
  - Query suggestions & auto-completions
- Mining query ➢ pivotal query pairs
  - Can identify automatically: F1 59%
    (final query baseline: 51%)

Future Work

- Directly applying reformulation strategies
  - Query suggestions & auto-completions
    - Mining query ➢ pivotal query pairs
      - Can identify automatically: F1 59%
        (final query baseline: 51%)
- Hints and tips on reformulation strategies

Real-time Recommendations at Blendle

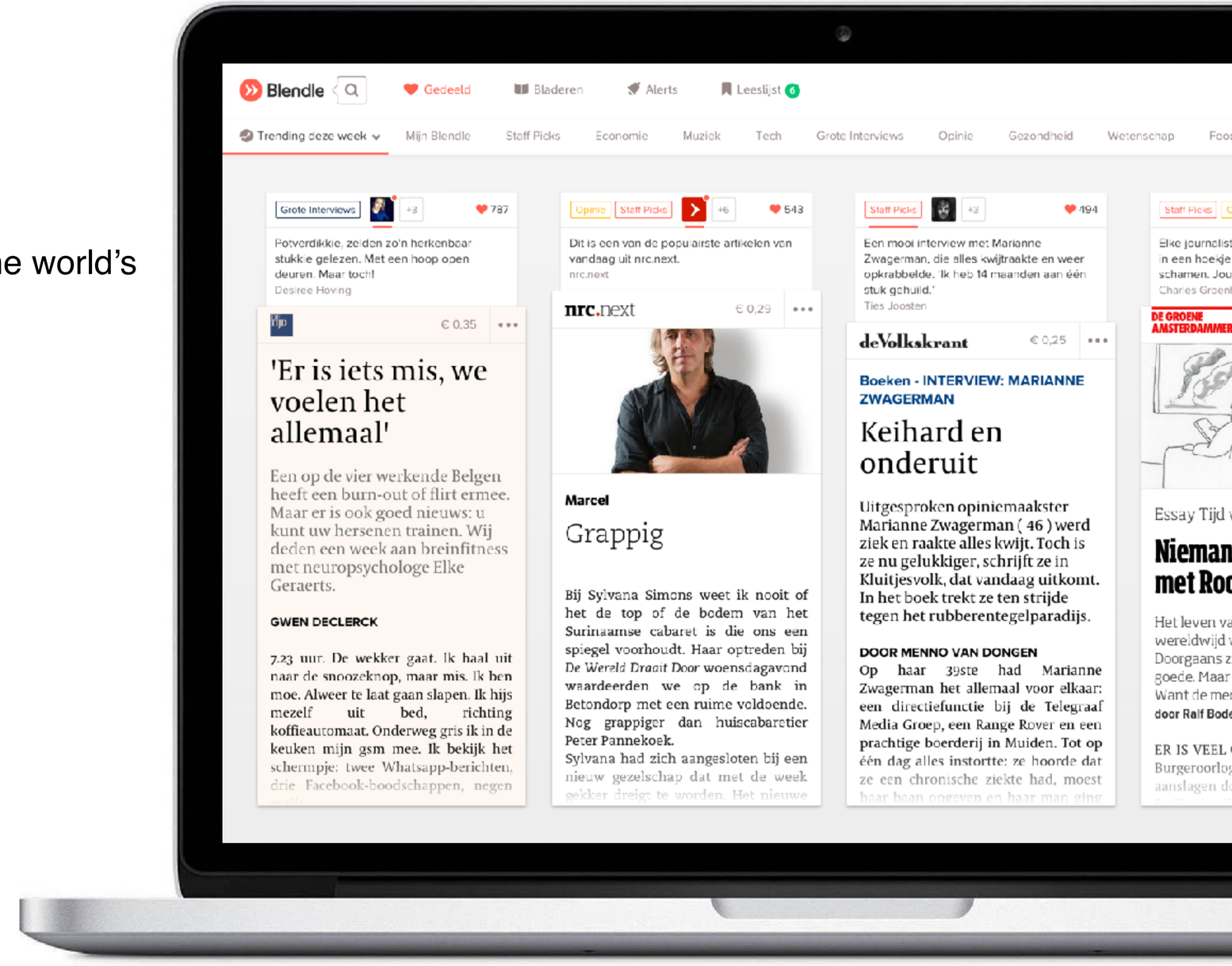
Daan Odijk & Anne Schuth
data scientists

Blendle

Blendle

Mission

Help you discover and support the world's best journalism

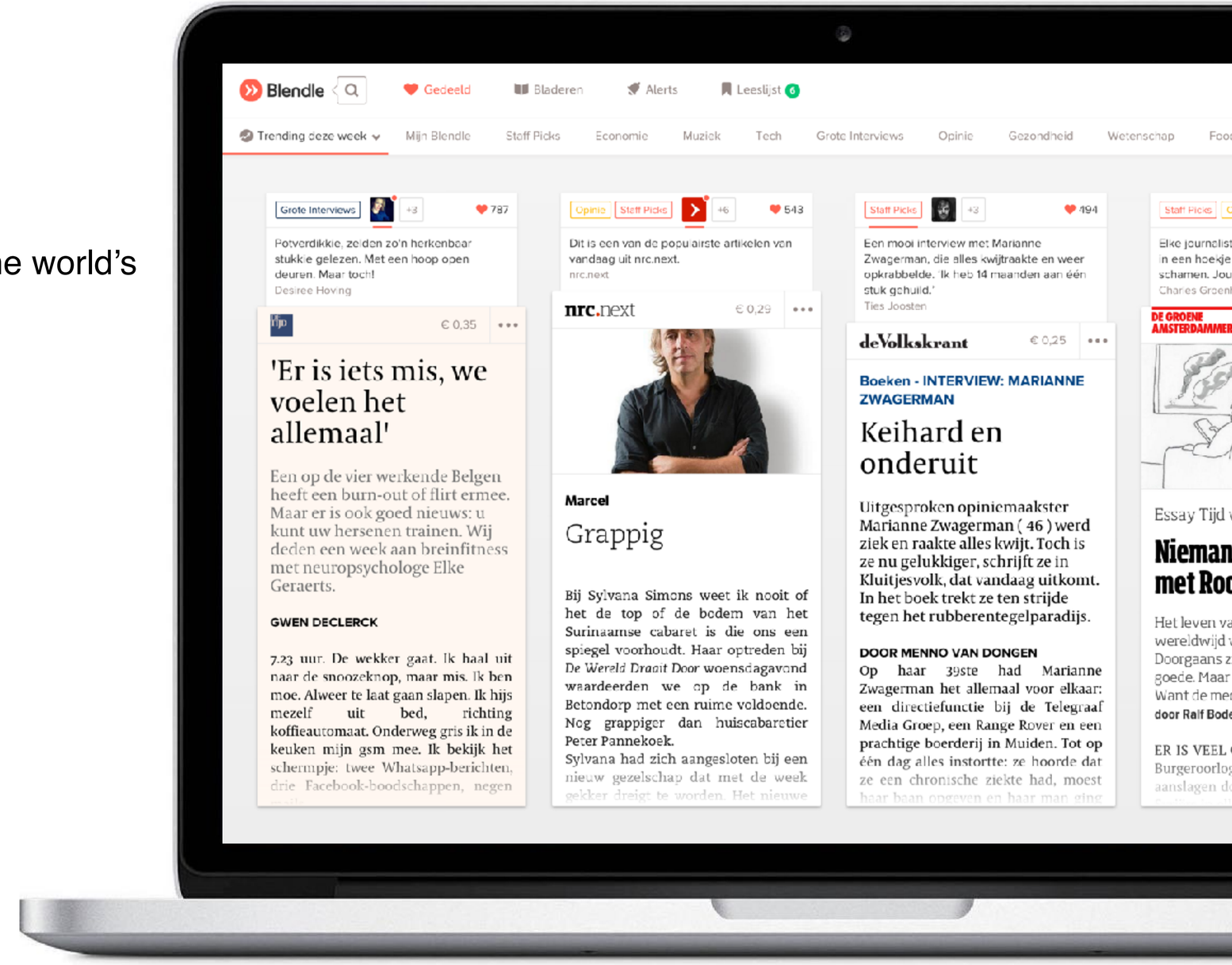


Blendle @anneschuth

Blendle

- **Mission**
  Help you discover and support the world's best journalism

- **Micropayment platform**
  articles, issues, subscriptions



Blendle @anneschuth

Blendle

Mission

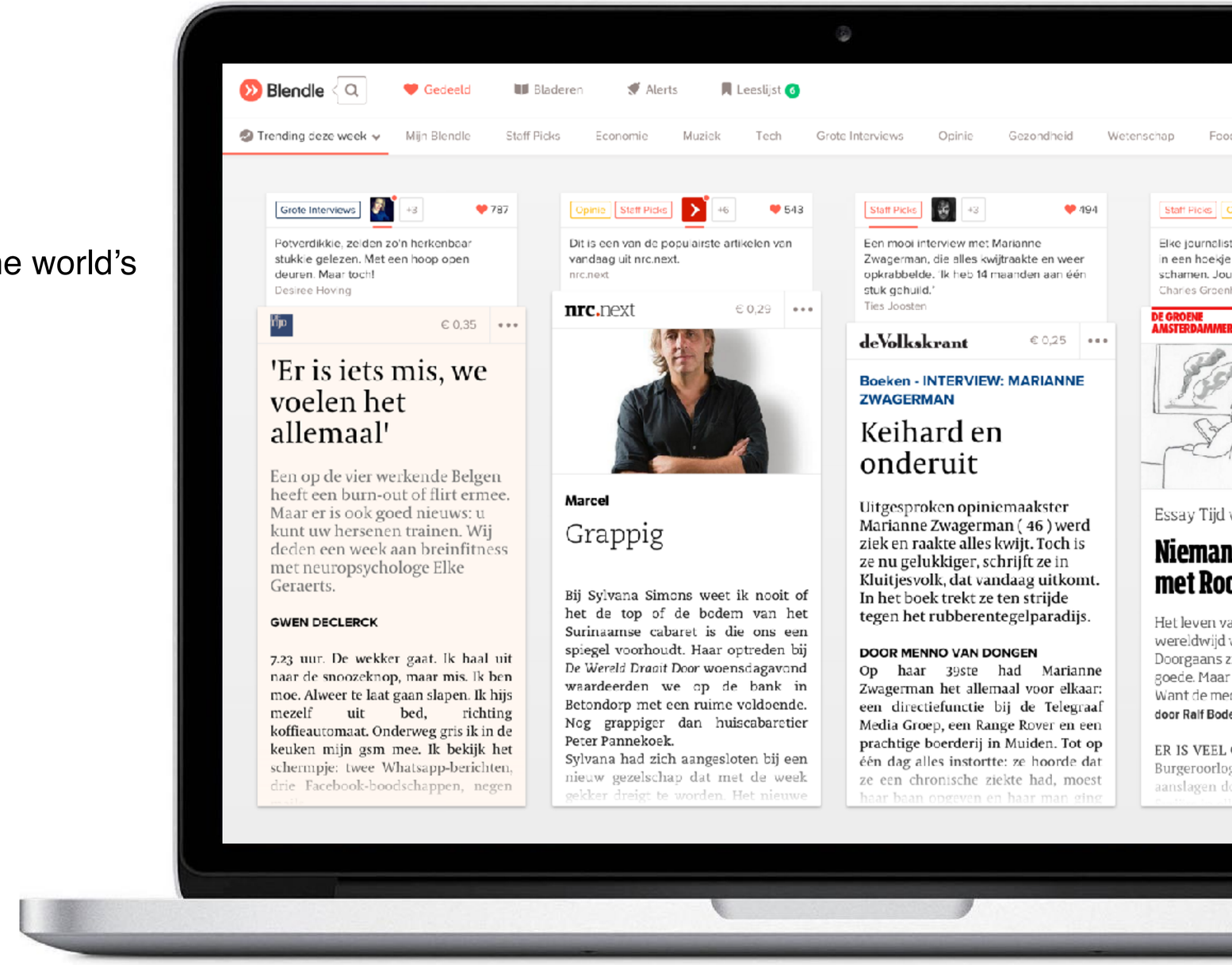
Help you discover and support the world's best journalism

Micropayment platform

articles, issues, subscriptions

International

May 2014 : The Netherlands
Sept 2015 : Germany
March 2016 : United States



Blendle @anneschuth

Blendle

Mission

Help you discover and support the world's best journalism

Micropayment platform

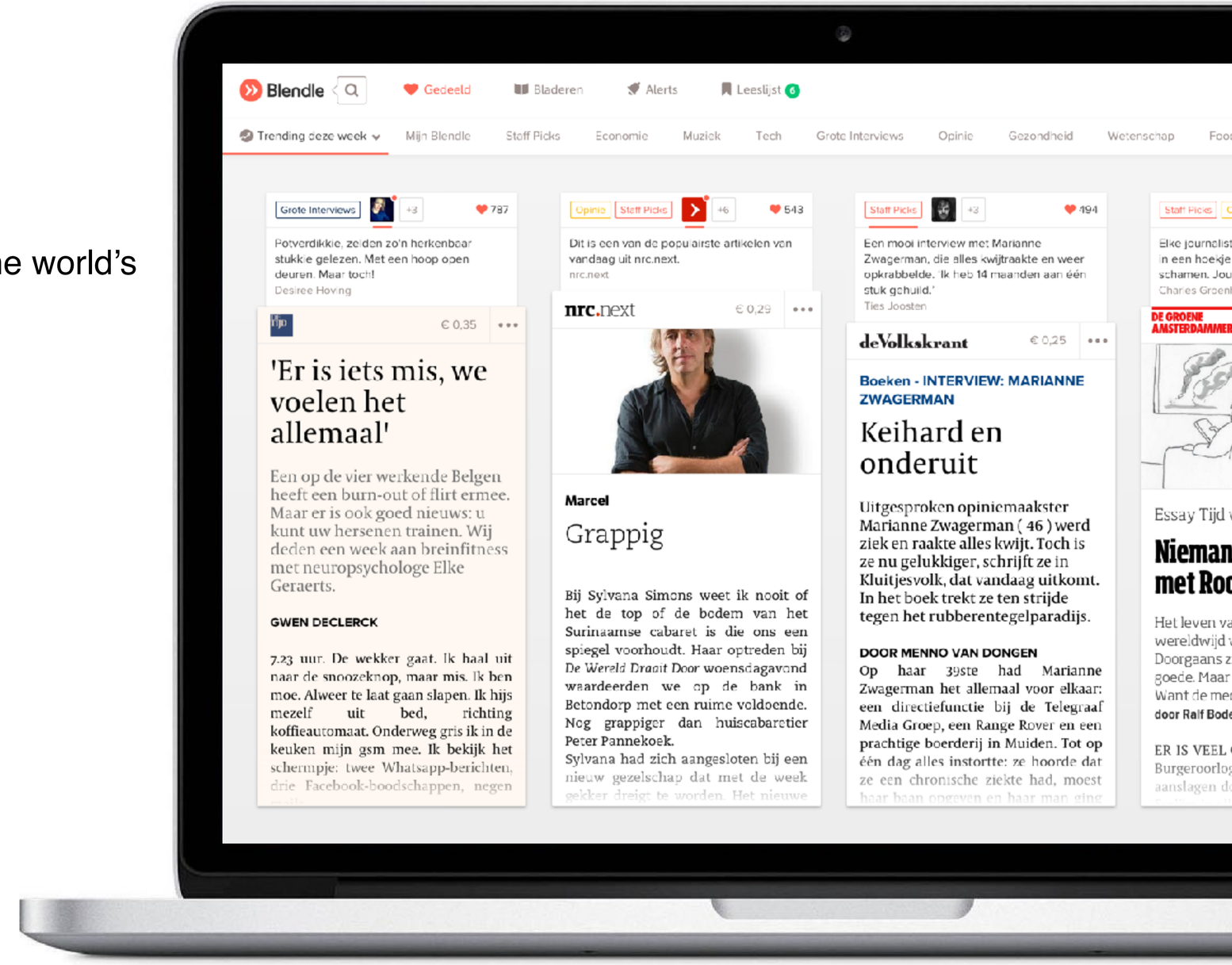
articles, issues, subscriptions

International

May 2014 : The Netherlands
Sept 2015 : Germany
March 2016 : United States

Numbers

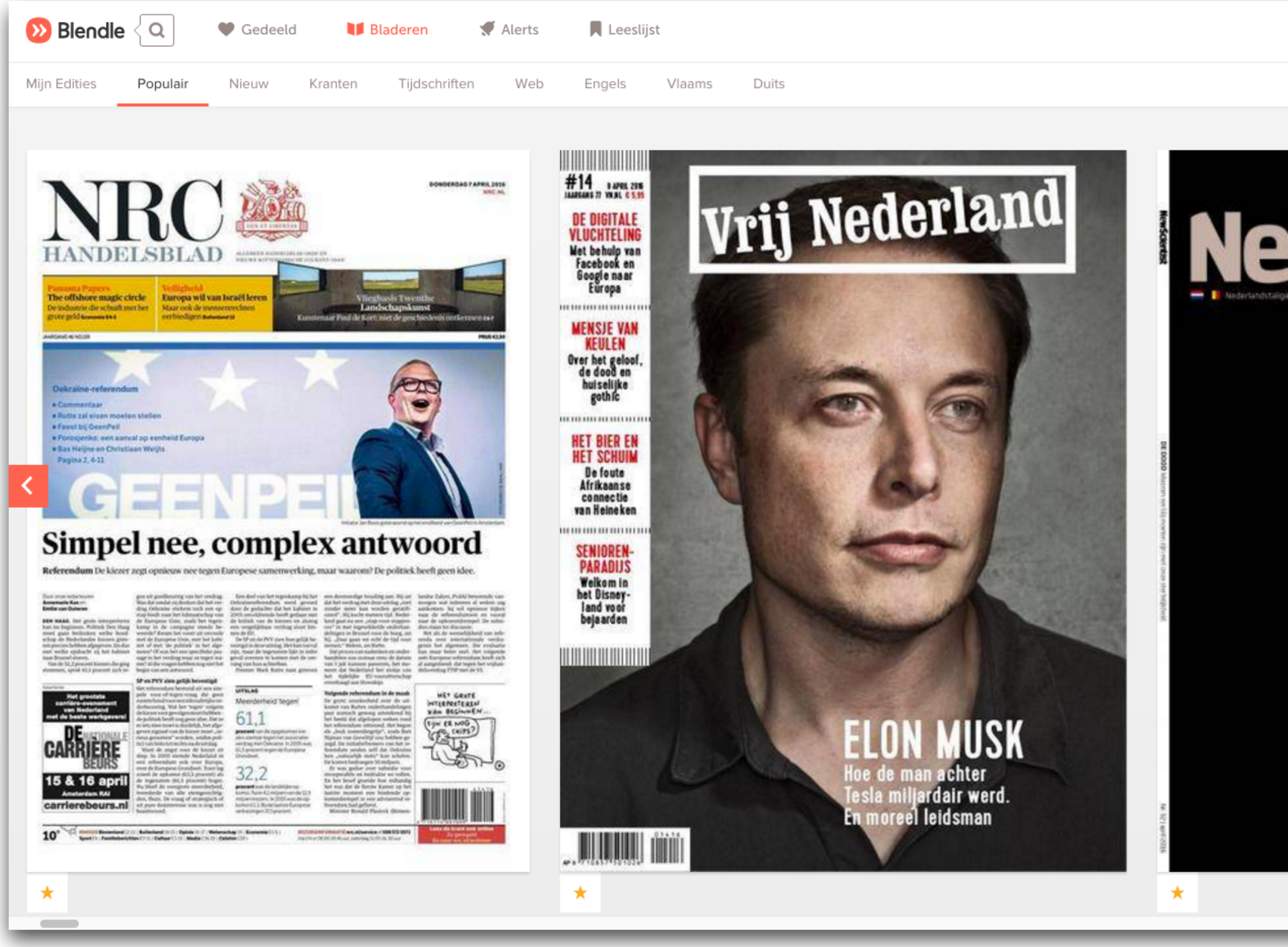
> 1M users
1 in 5 is paying
70 employees
(15 journalists, 50 developers,
+ 30 clippers)



Blendle

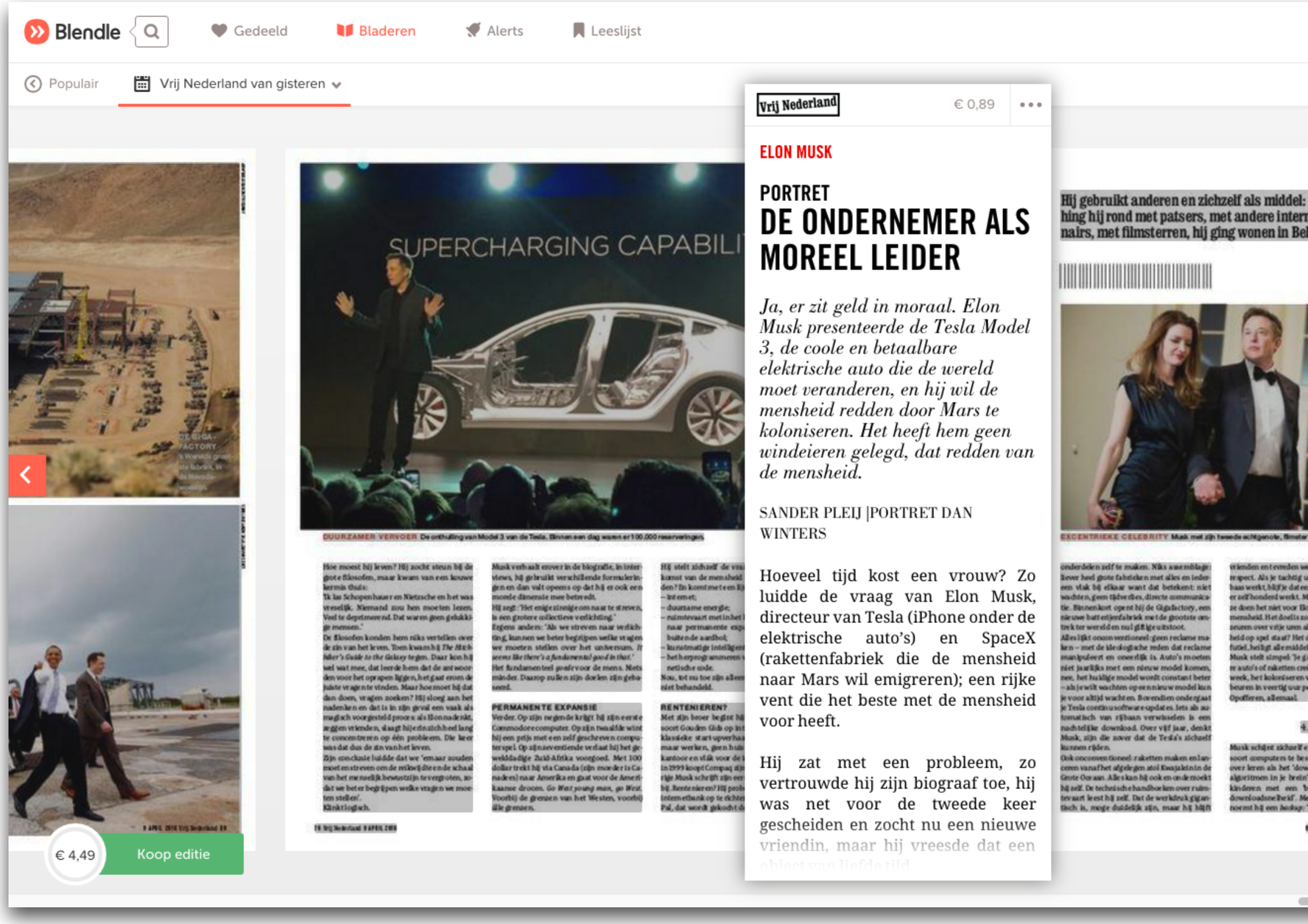
@anneschuth

In Blendle you can browse through all **quality newspapers** and **magazines**



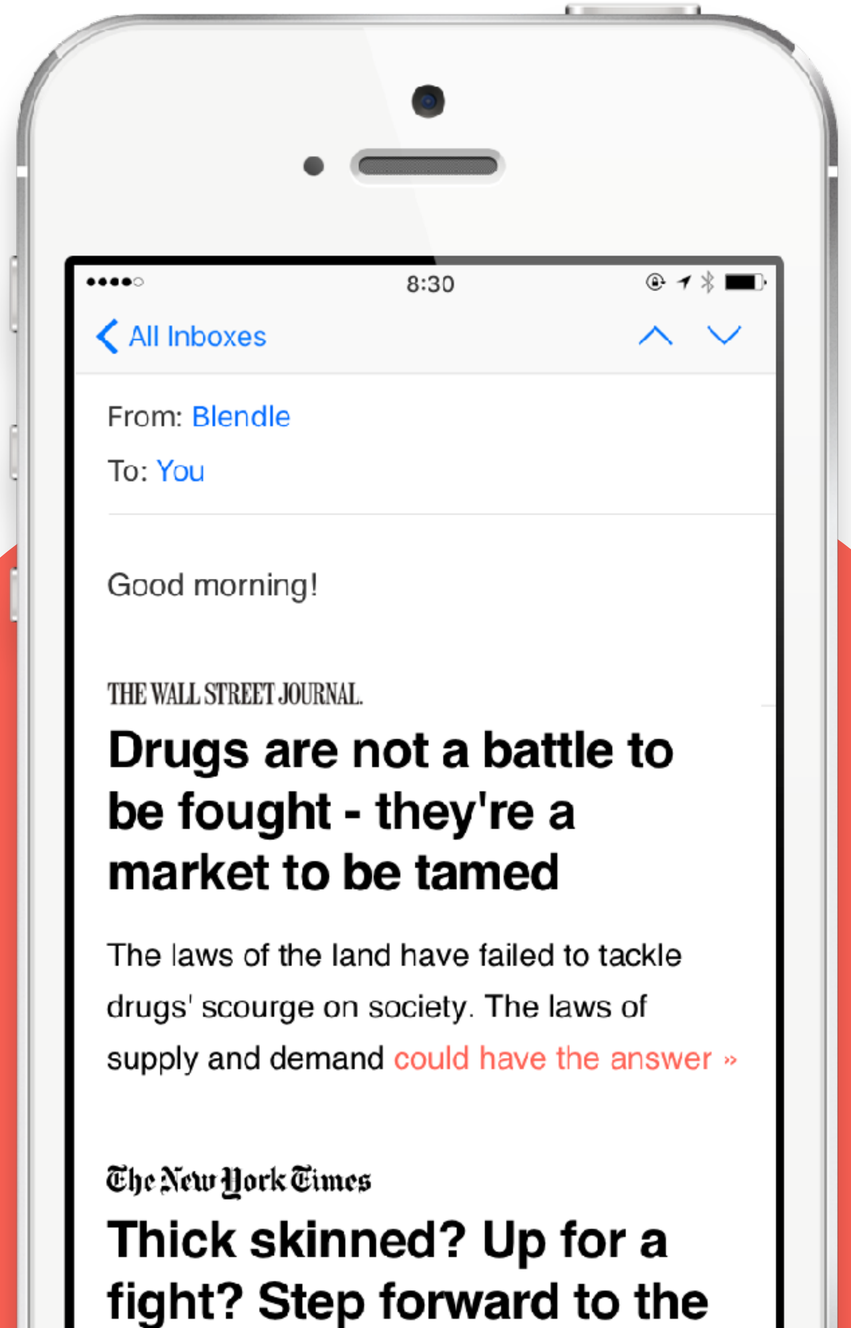
Blendle @anneschuth

You only pay for what you read, with a **single click**



Blendle @anneschuth

Our **editors** select the best articles for our email newsletter **every day**



Blendle @anneschuth

Data at Blendle

- Articles
  - > 3.5M in total
  - > 6K new every day
  - > 30% is read

Data at Blendle

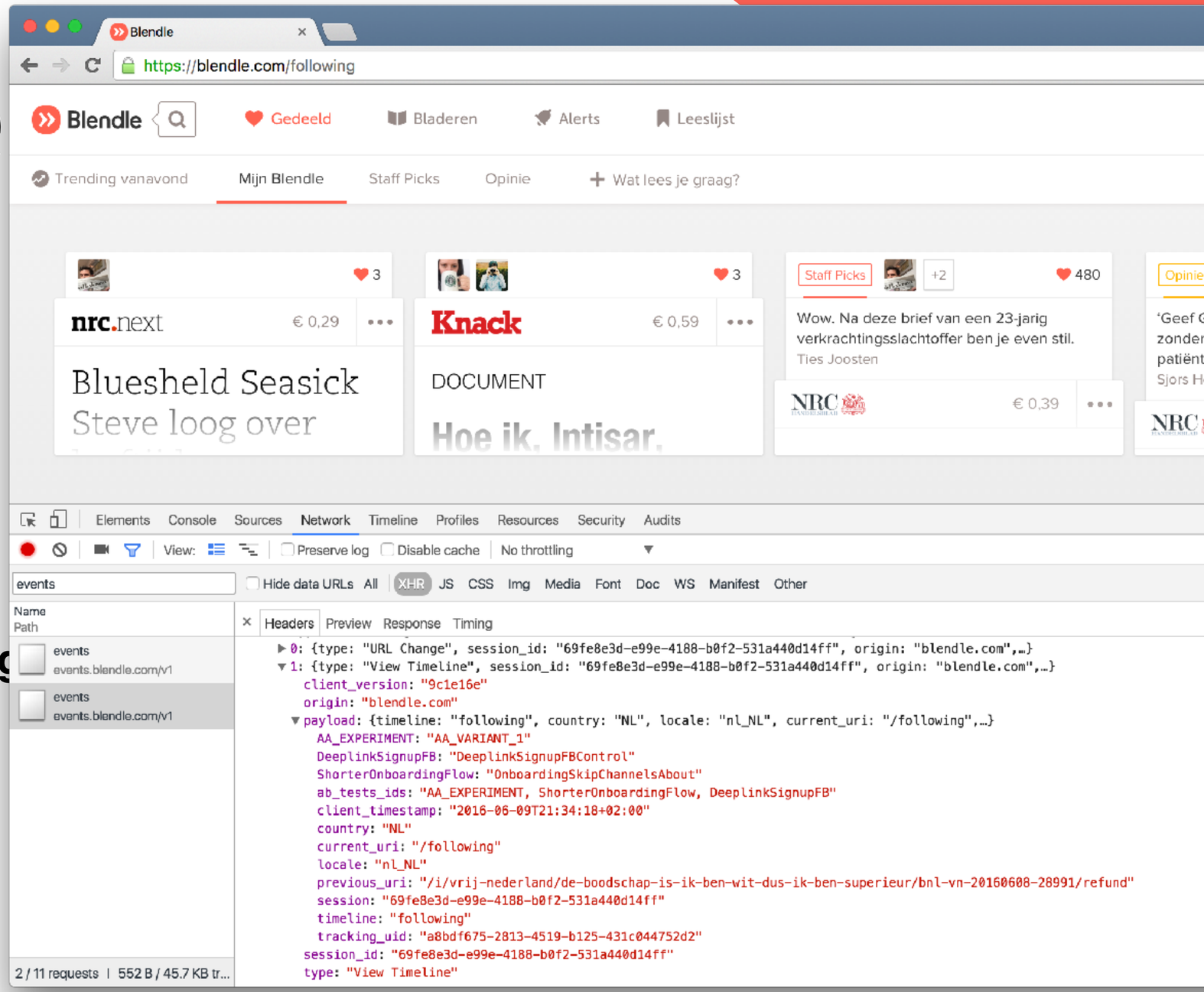
- Articles
  - > 3.5M in total
  - > 6K new every day
  - > 30% is read
- Users
  - > 1M users
  - ~ 1 in 5 converts to paying user

Data at Blendle

- Articles
  - > 3.5M in total
  - > 6K new every day
  - > 30% is read
- Users
  - > 1M users
  - ~ 1 in 5 converts to paying user
- Events
  - > 1B in total
  - > 2M new every day

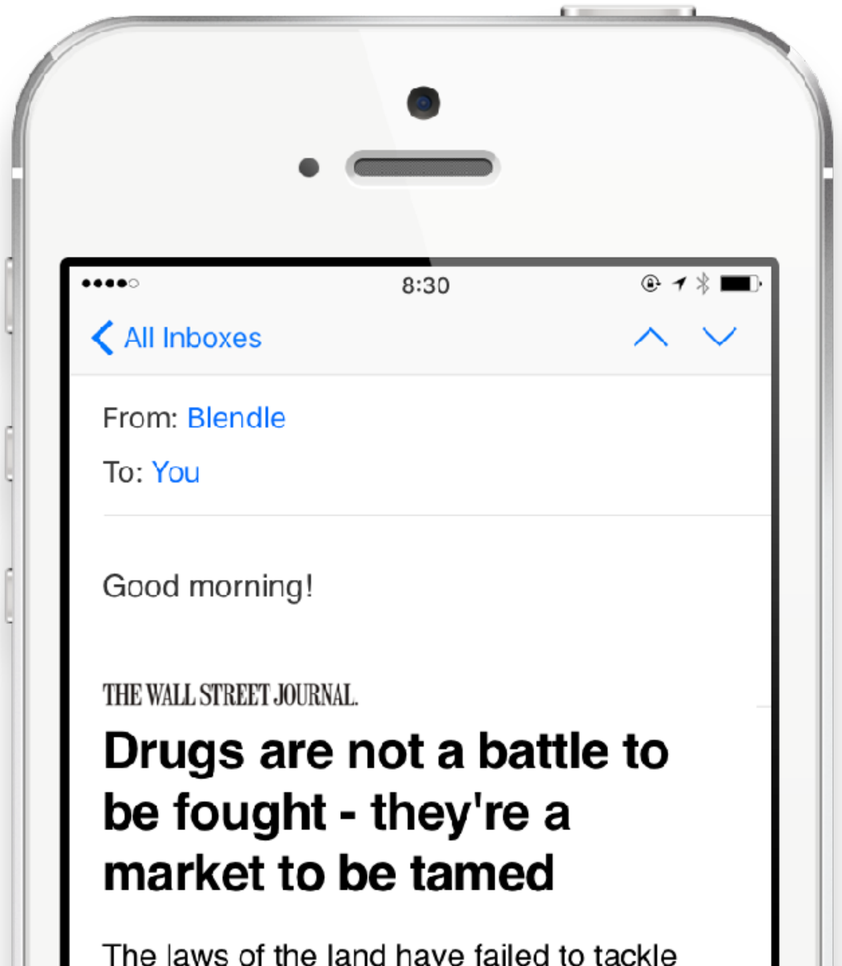
Data at Ble

- Articles
  - > 3.5M in total
  - > 6K new every day
  - > 30% is read
- Users
  - > 1M users
  - ~ 1 in 5 converts to paying
- Events
  - > 1B in total
  - > 2M new every day



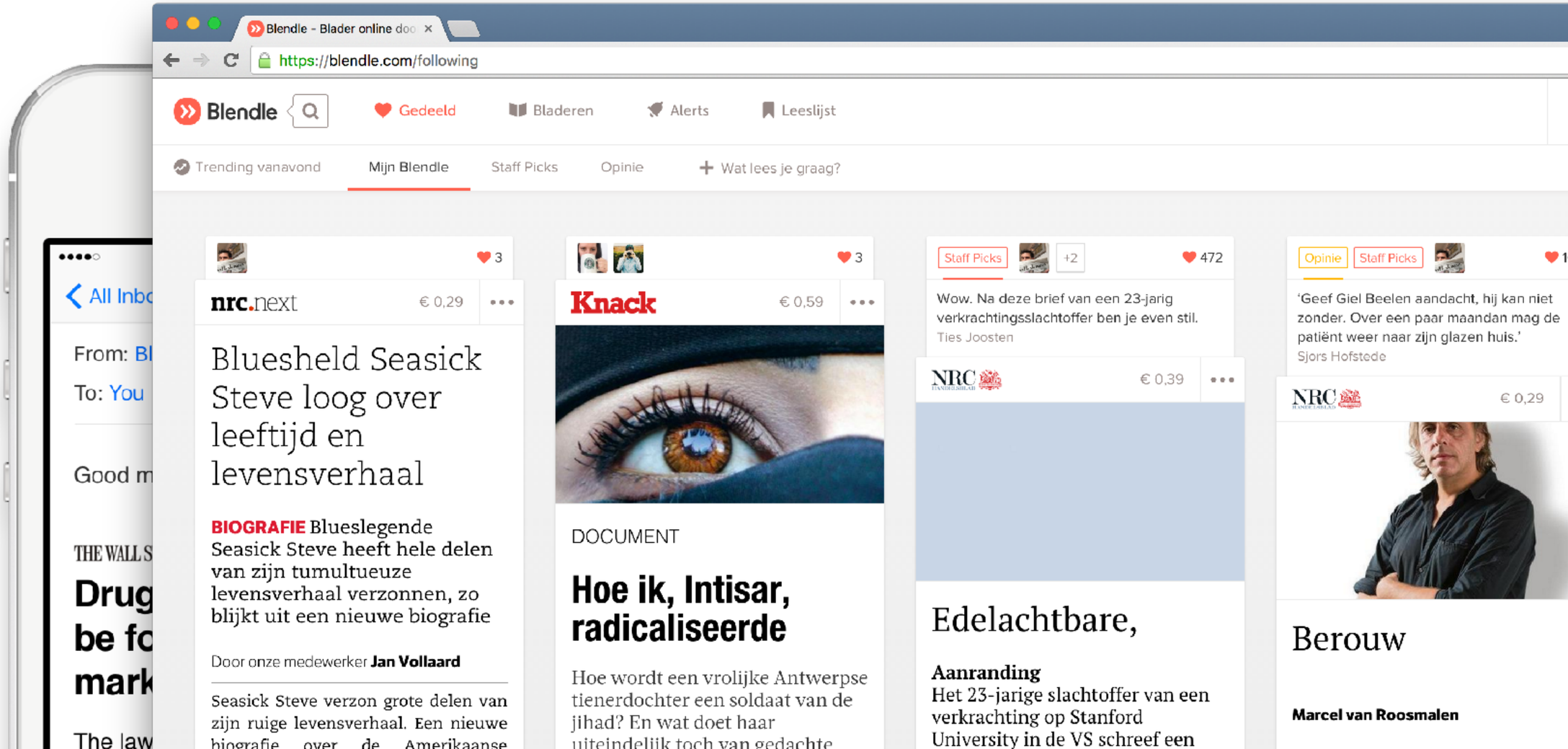
Blendle @anneschuth

Rankings at Blendle



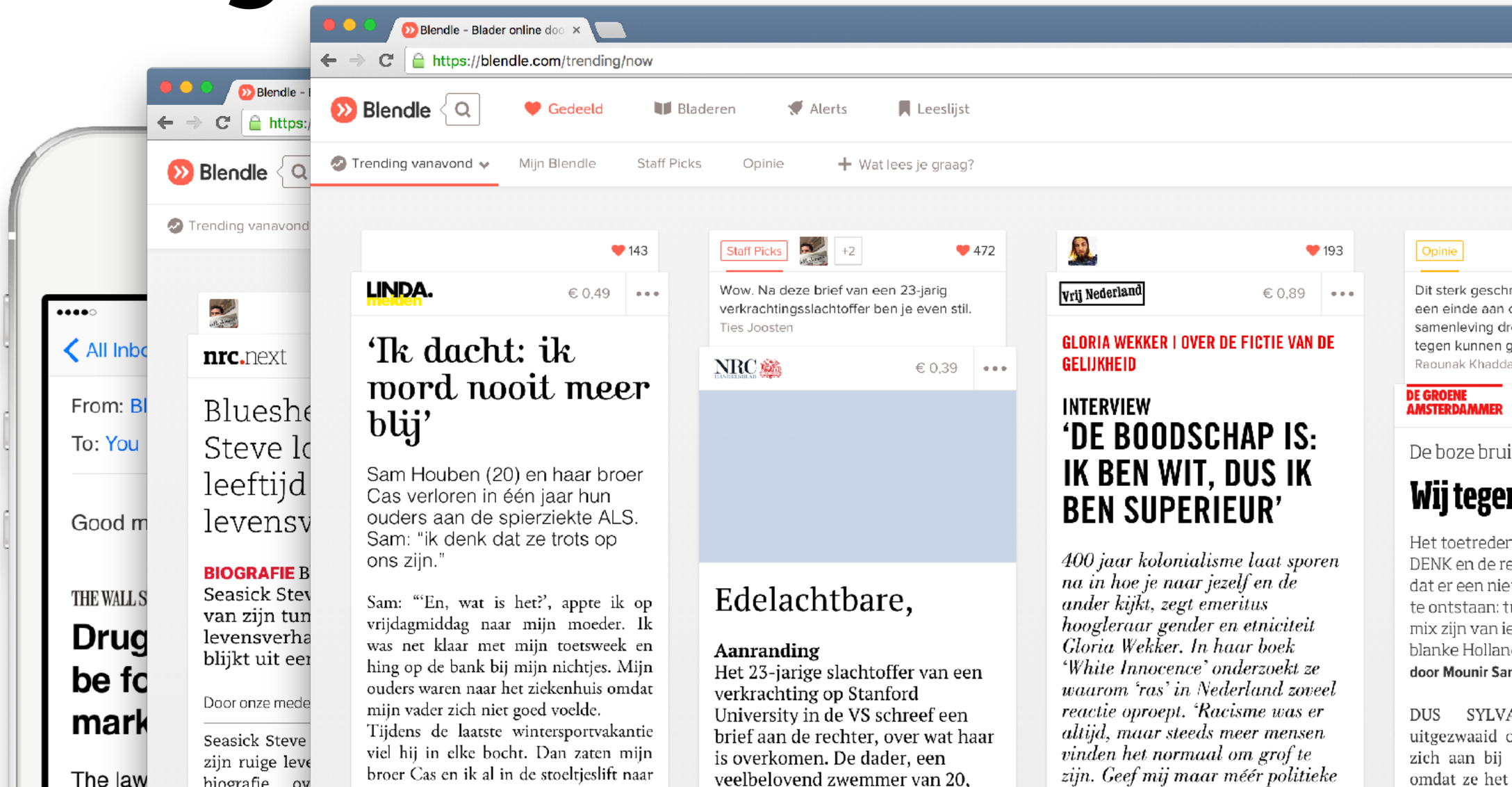
Blendle

Rankings at Blendle



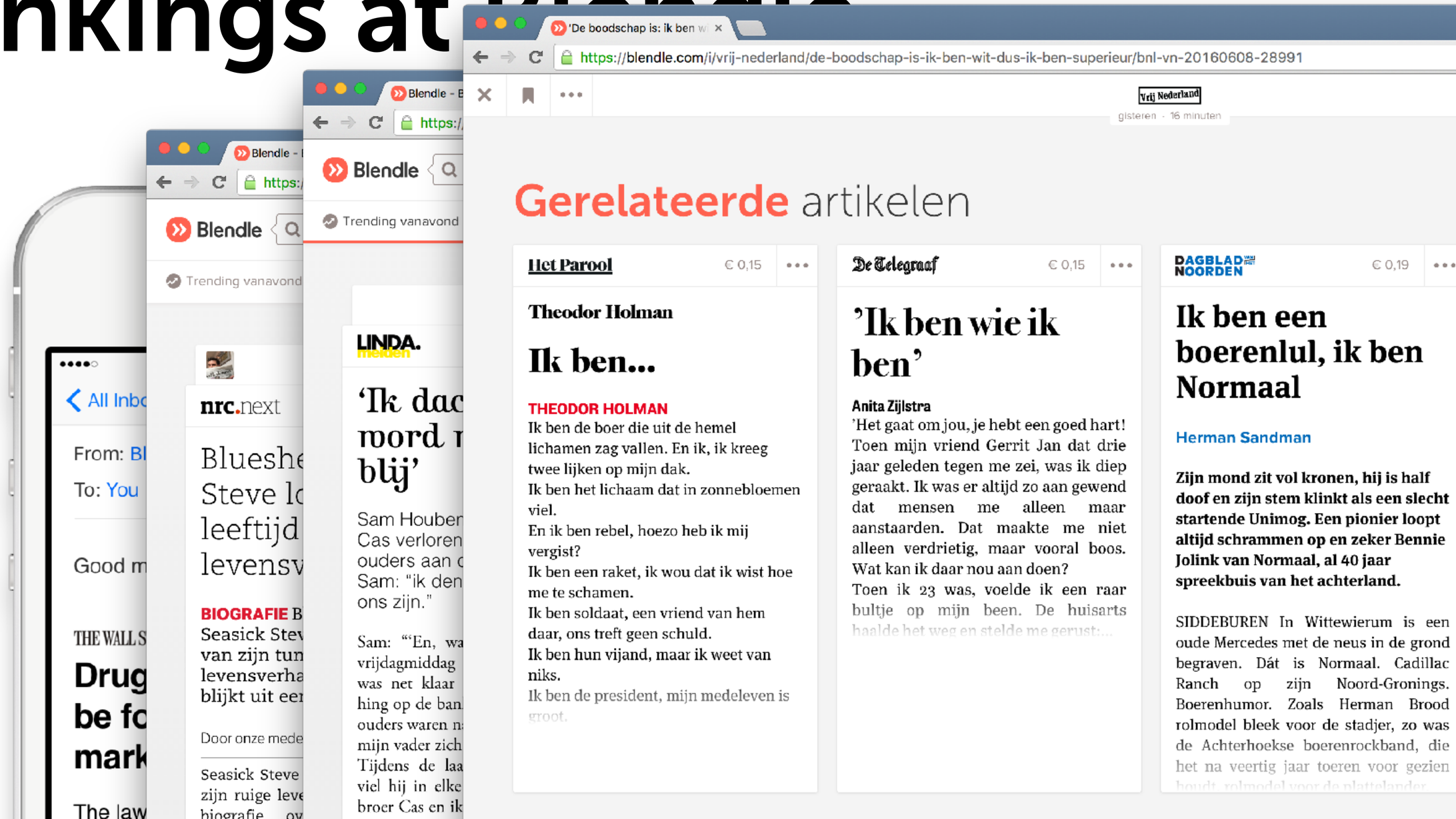
Blendle

Rankings at Blendle



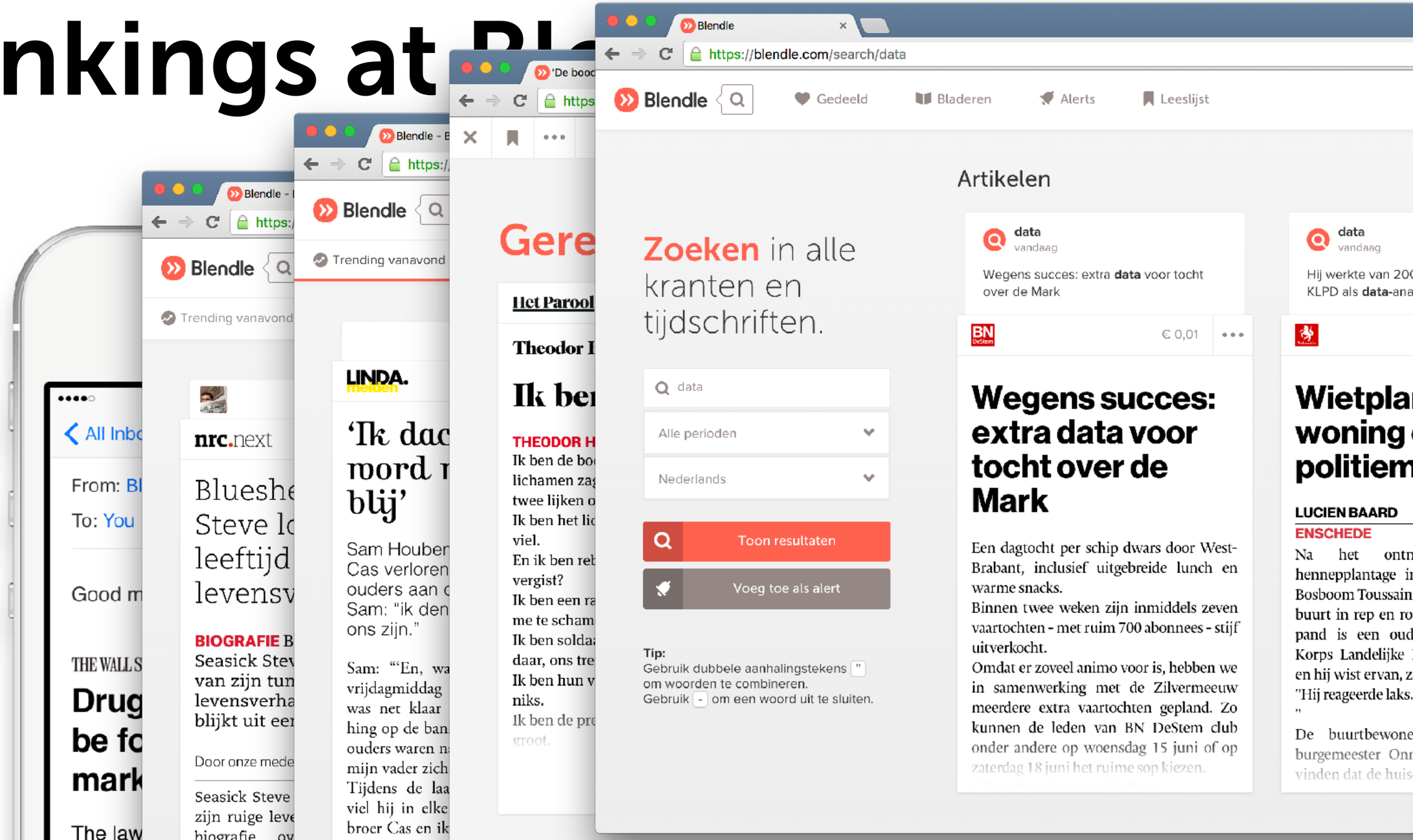
Blendle

Rankings at Blendle



Blendle

Rankings at Bl



Blendle

Rankings at Blendle

Rankings at Blendle

- Should be personalised

Rankings at Blendle

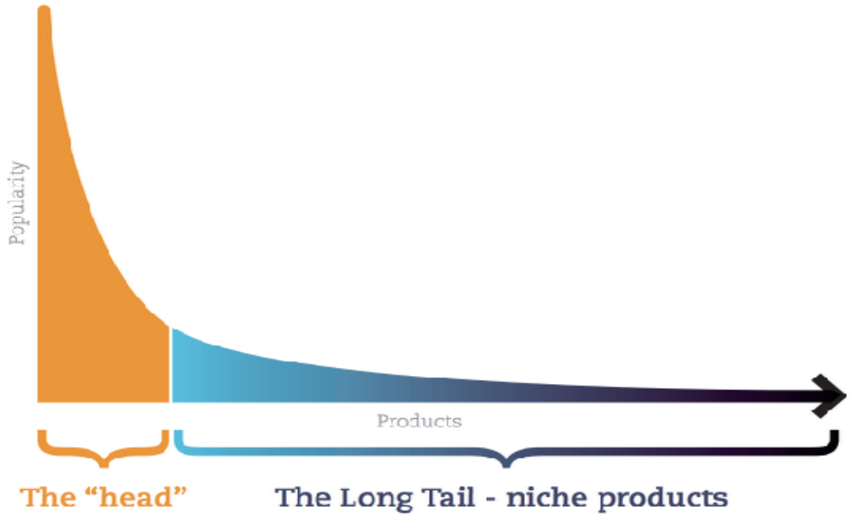
- Should be personalised
    - Lots of content

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes

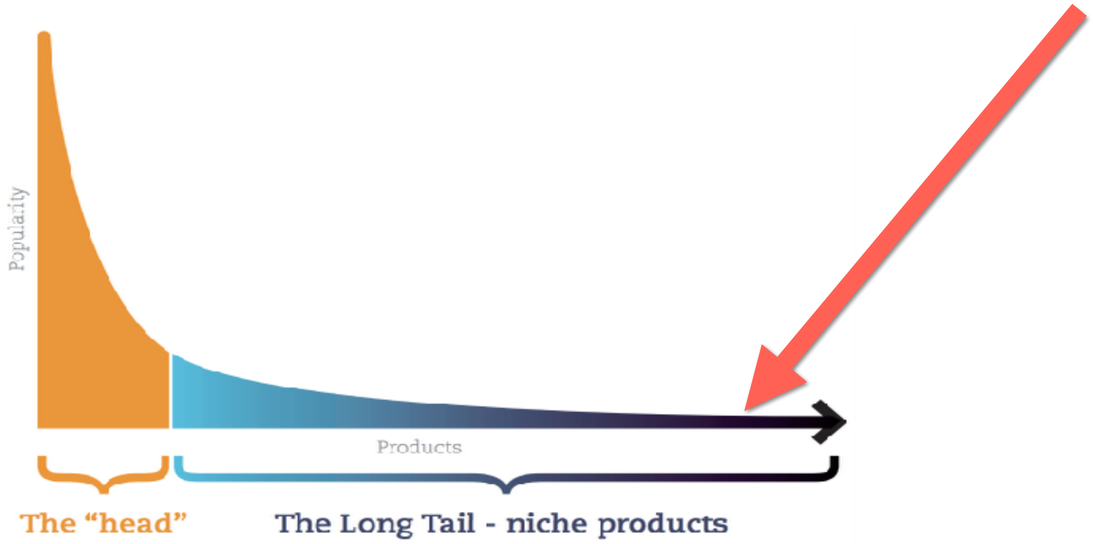
Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes



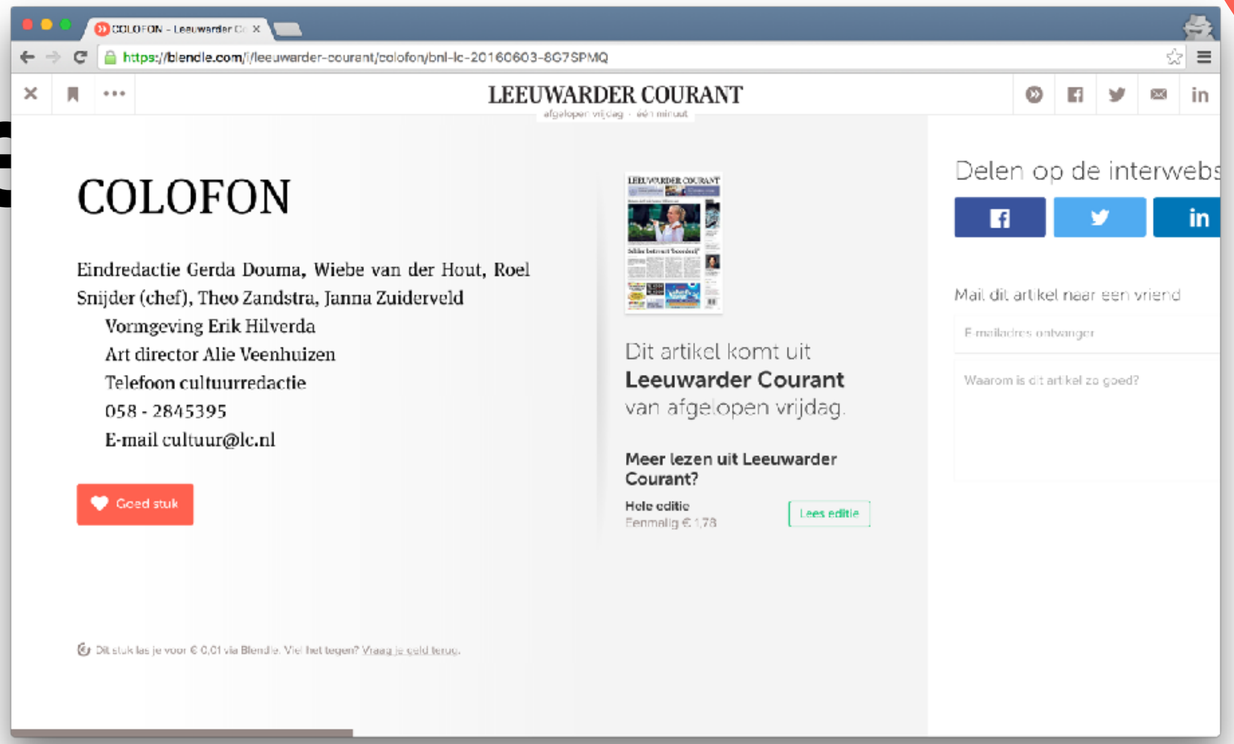
Rankings at Blendle

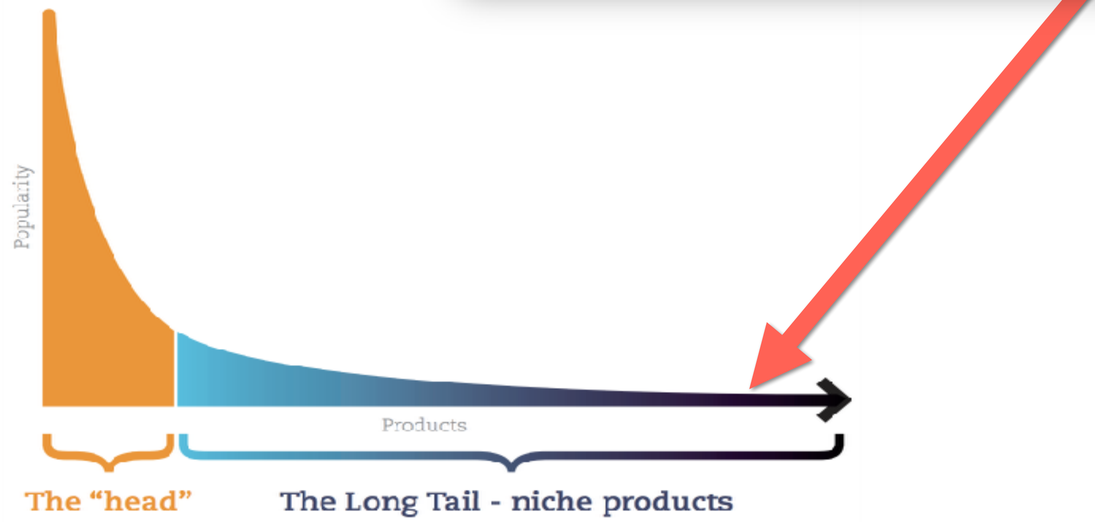
- Should be personalised
  - Lots of content
  - Lots of different tastes



Rankings at Ble

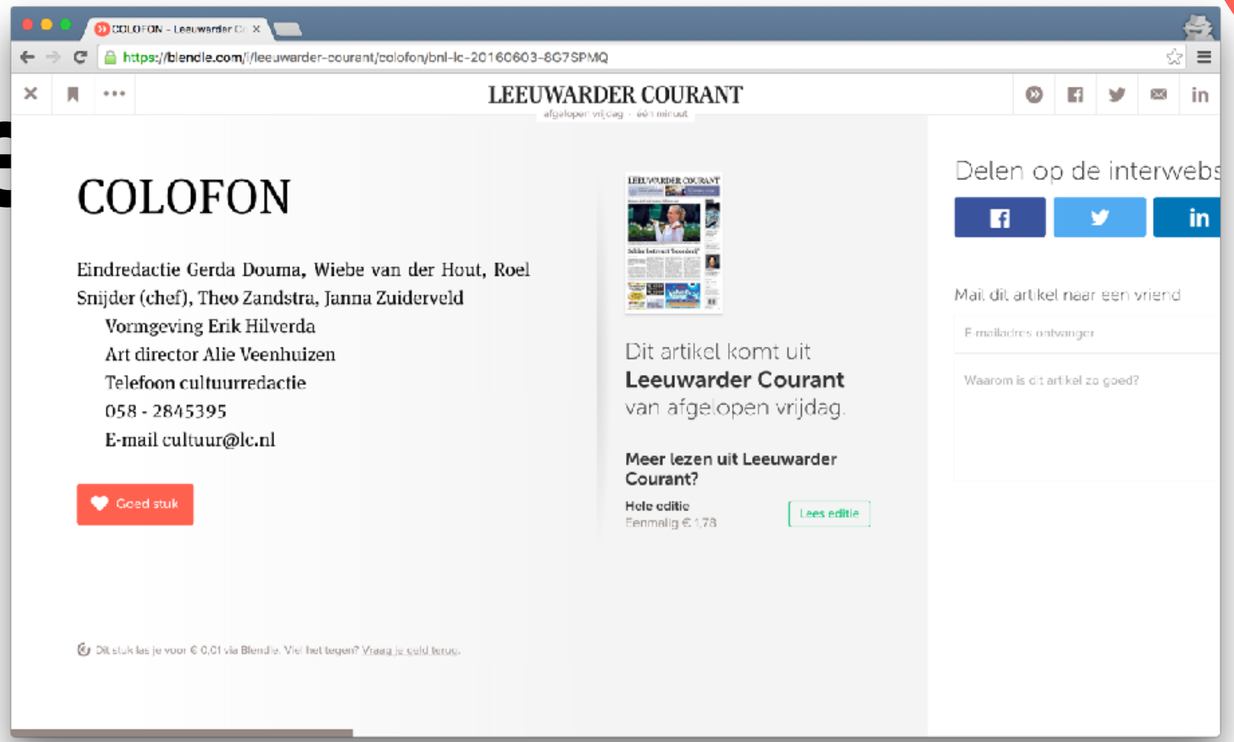
- Should be personalised
  - Lots of content
  - Lots of different tastes

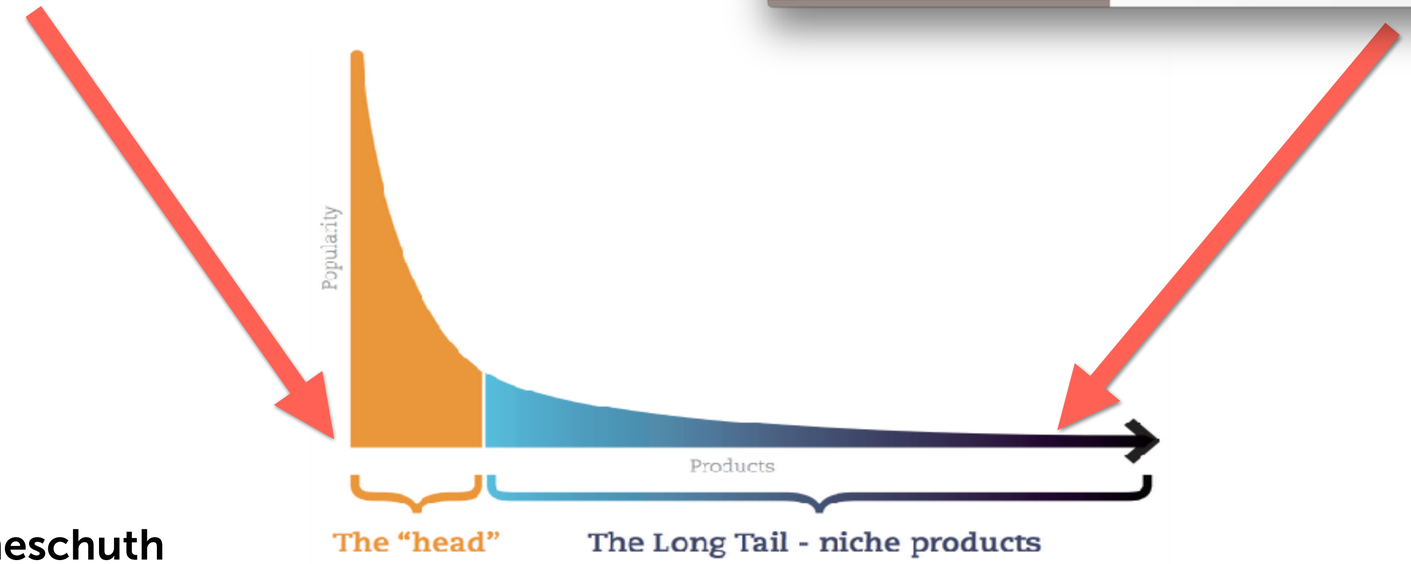


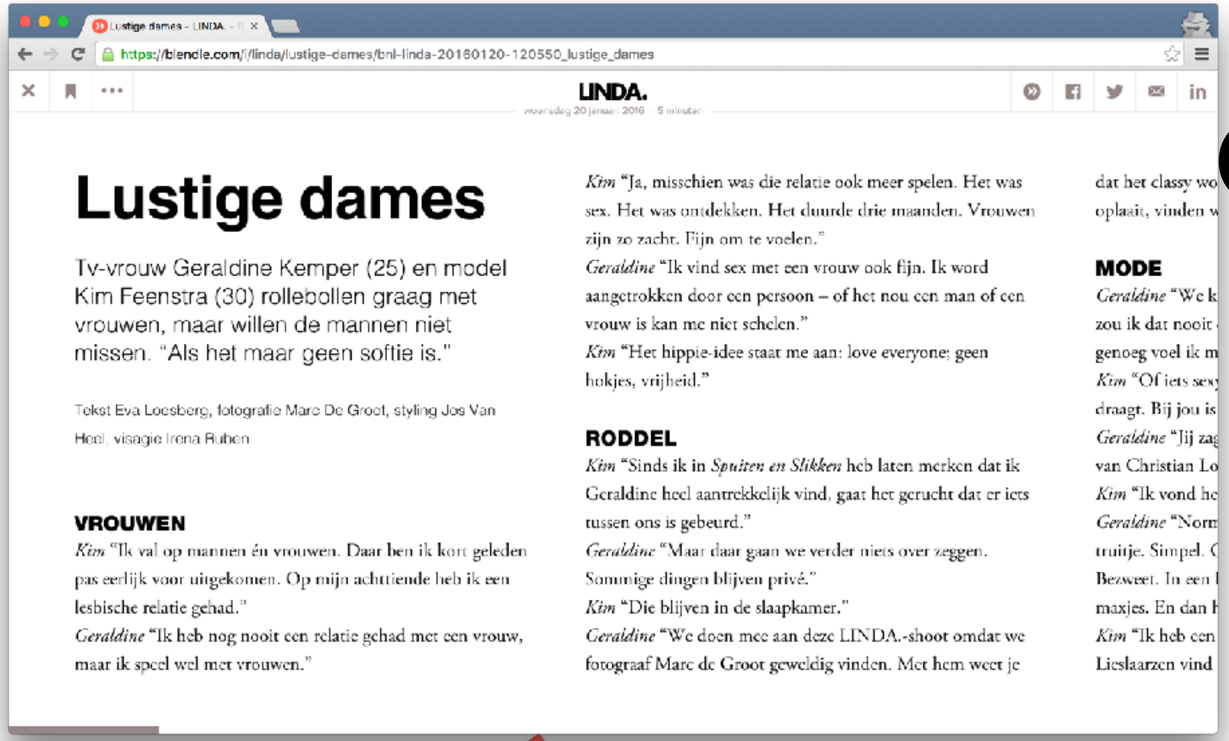


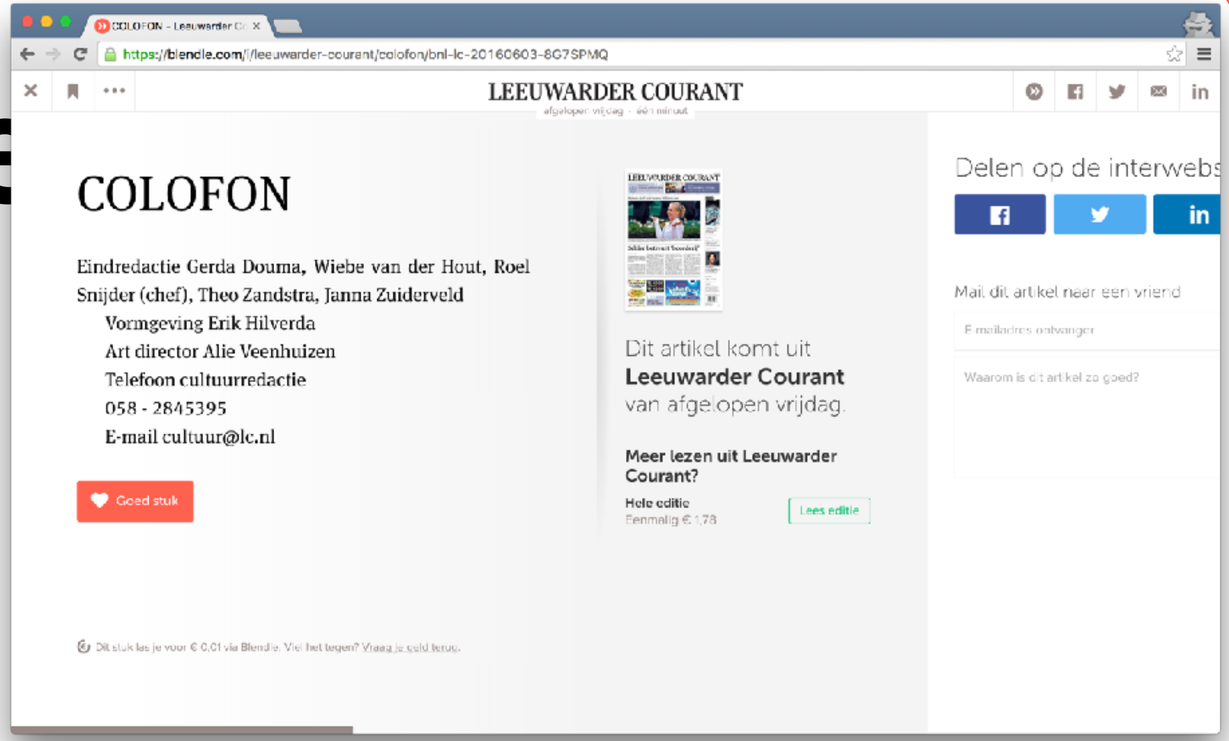
Rankings at Ble

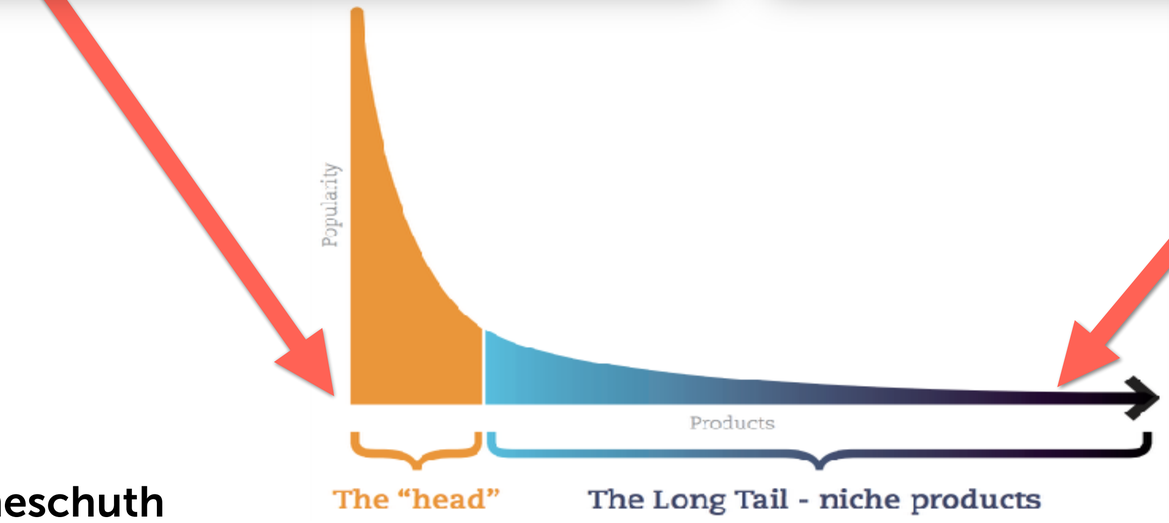
- Should be personalised
  - Lots of content
  - Lots of different tastes



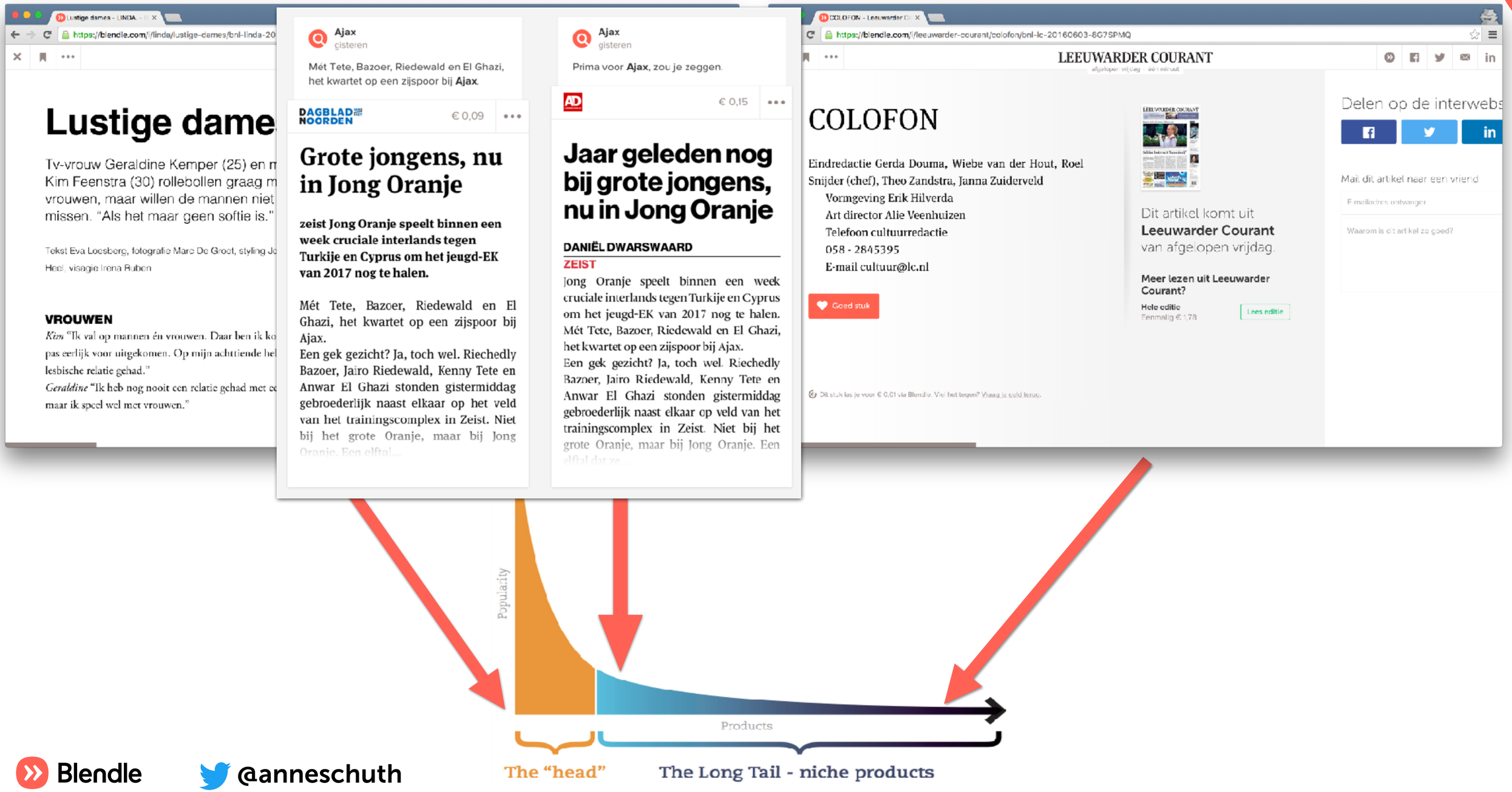




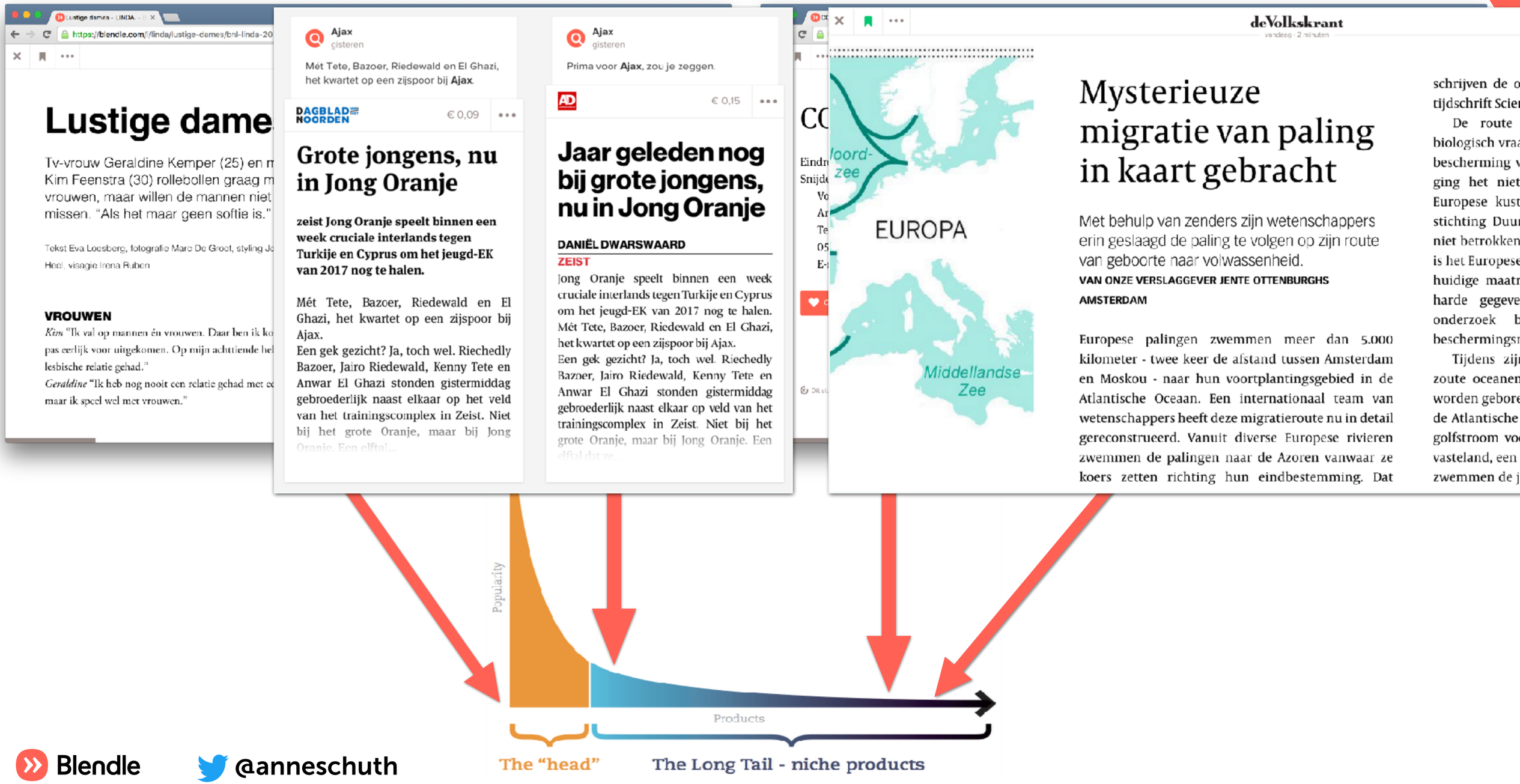




Blendle @anneschuth



Blendle @anneschuth



Blendle @anneschuth

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our users read

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our users read
  - What they didn’t like

Rankings a

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our user
  - What they didn't like

Blendle @anneschuth

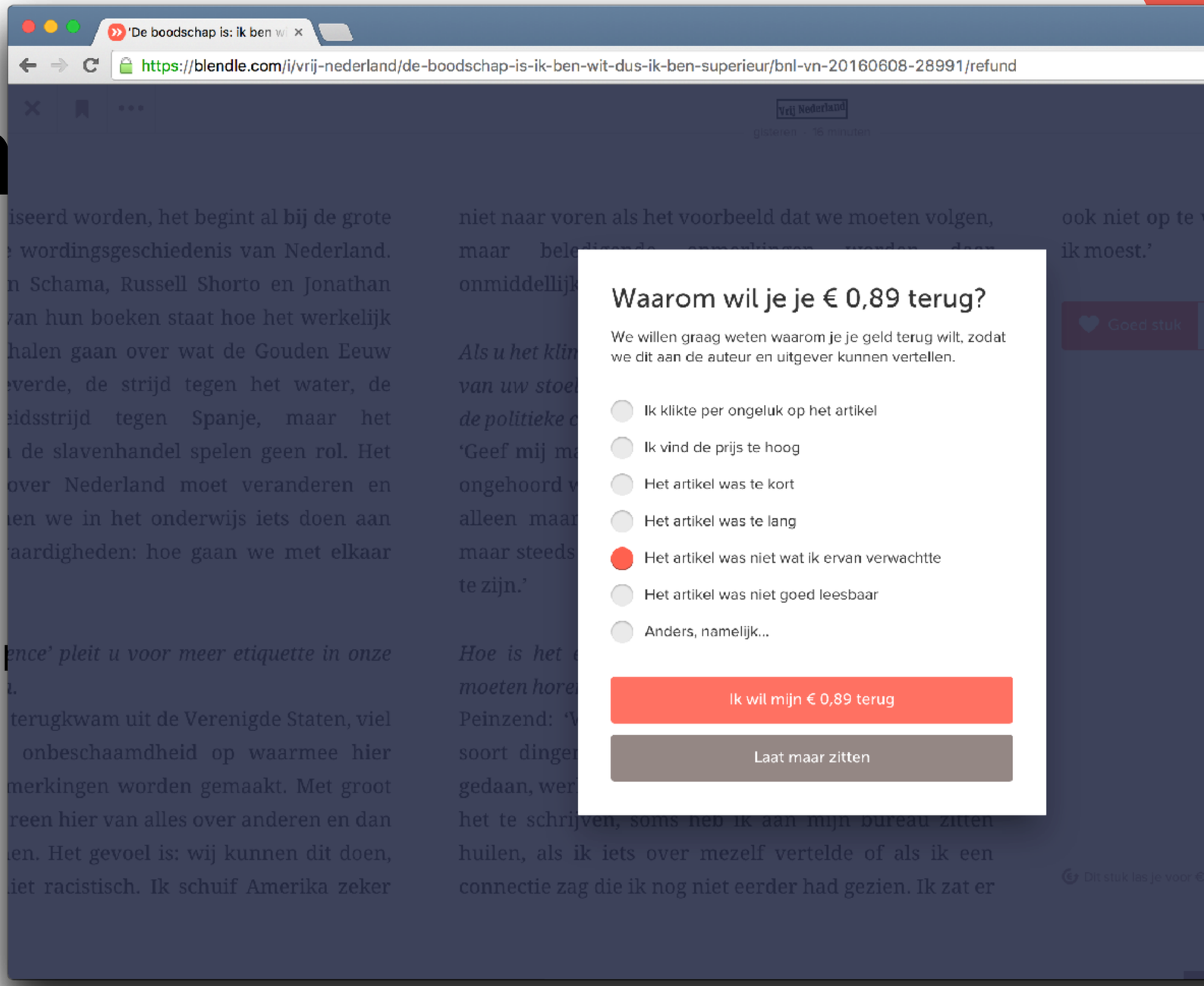
Rankings a

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our user
  - What they didn't like

Blendle @anneschuth

Rankings a

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our user
  - What they didn't like



Blendle @anneschuth

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our users read
  - What they didn’t like

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our users read
  - What they didn’t like
  - When they read

Rankings at Blendle

- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our users read
  - What they didn’t like
  - When they read
  - …

Personalised recommendations

Music

- **Repeated** consumption
- More **niche** than movies/series
- **Tens of millions** of songs and users
- **Long** (often unlimited) shelf life
- Feedback: implicit (listening behaviour)

Personalised recommendations

Music

- **Repeated** consumption
- More **niche** than movies/series
- **Tens of millions** of songs and users
- **Long** (often unlimited) shelf life
- Feedback: implicit (listening behaviour)

Movies and series

- **Single** consumption
- Less niche than music
- 75.000.000 users
- **Long** shelf life
- Feedback: explicit ratings

Personalised recommendations

Music

- **Repeated** consumption
- More **niche** than movies/series
- **Tens of millions** of songs and users
- **Long** (often unlimited) shelf life
- Feedback: implicit (listening behaviour)

Articles

- **Single** consumption

Movies and series

- **Single** consumption
- Less niche than music
- 75.000.000 users
- **Long** shelf life
- Feedback: explicit ratings

Personalised recommendations

Music

- **Repeated** consumption
- More **niche** than movies/series
- **Tens of millions** of songs and users
- **Long** (often unlimited) shelf life
- Feedback: implicit (listening behaviour)

Movies and series

- **Single** consumption
- Less niche than music
- 75.000.000 users
- **Long** shelf life
- Feedback: explicit ratings

Articles

- **Single** consumption
- Very **short** shelf-life
  - **Cold start problem, Collaborative Filtering won't work**

Personalised recommendations

Music

- **Repeated** consumption
- More **niche** than movies/series
- **Tens of millions** of songs and users
- **Long** (often unlimited) shelf life
- Feedback: implicit (listening behaviour)

Movies and series

- **Single** consumption
- Less niche than music
- 75.000.000 users
- **Long** shelf life
- Feedback: explicit ratings

Articles

- **Single** consumption
- Very **short** shelf-life
  - **Cold start problem, Collaborative Filtering won't work**
- **Diversity** of suggestions very important
  - **Content Based Recommendation won't work**

Personalised recommendations

Music

- **Repeated** consumption
- More **niche** than movies/series
- **Tens of millions** of songs and users
- **Long** (often unlimited) shelf life
- Feedback: implicit (listening behaviour)

Movies and series

- **Single** consumption
- Less niche than music
- 75.000.000 users
- **Long** shelf life
- Feedback: explicit ratings

Articles

- **Single** consumption
- Very **short** shelf-life
  - **Cold start problem, Collaborative Filtering won't work**
- **Diversity** of suggestions very important
  - **Content Based Recommendation won't work**
- A lot of information **overlap**

Personalised recommendations

Music

- **Repeated** consumption
- More **niche** than movies/series
- **Tens of millions** of songs and users
- **Long** (often unlimited) shelf life
- Feedback: implicit (listening behaviour)

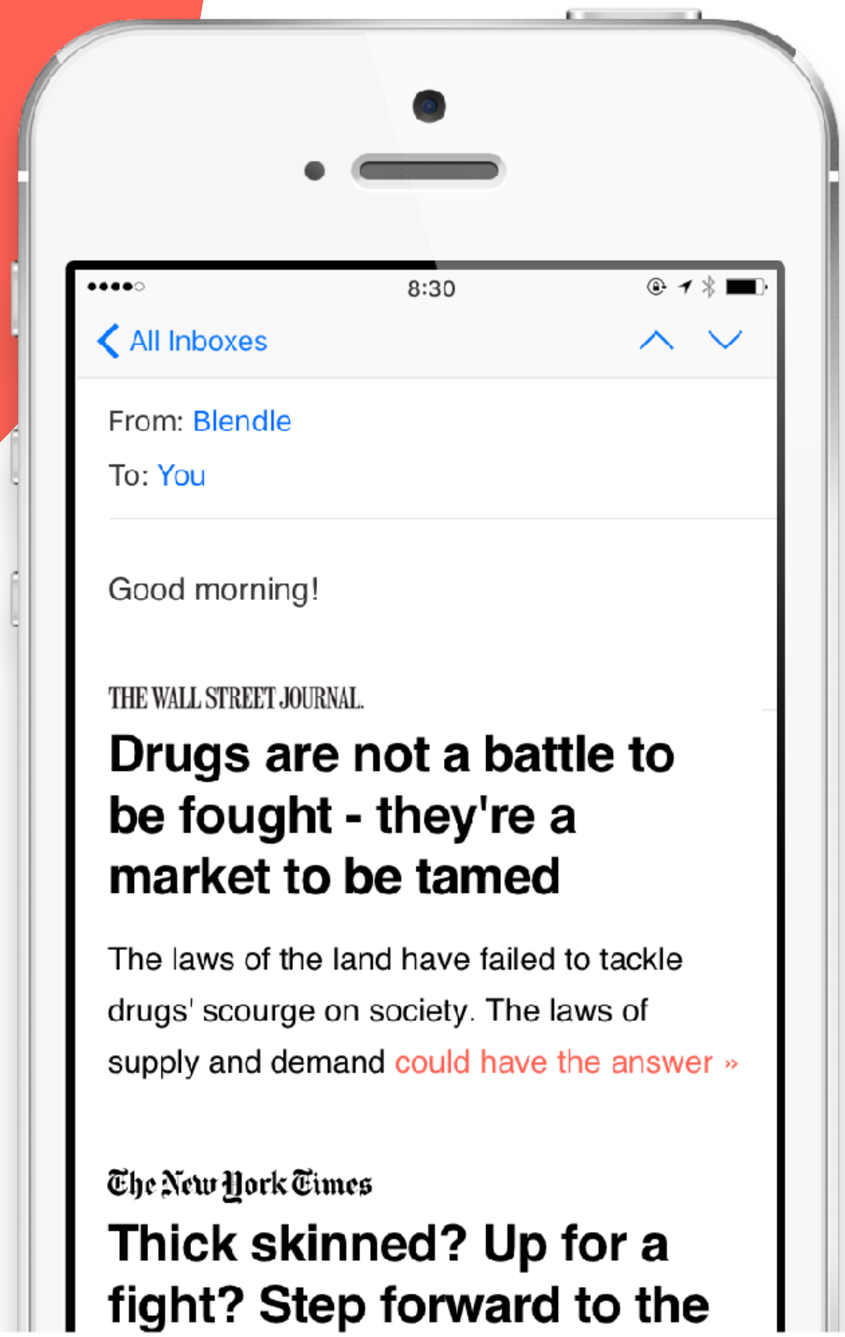
Movies and series

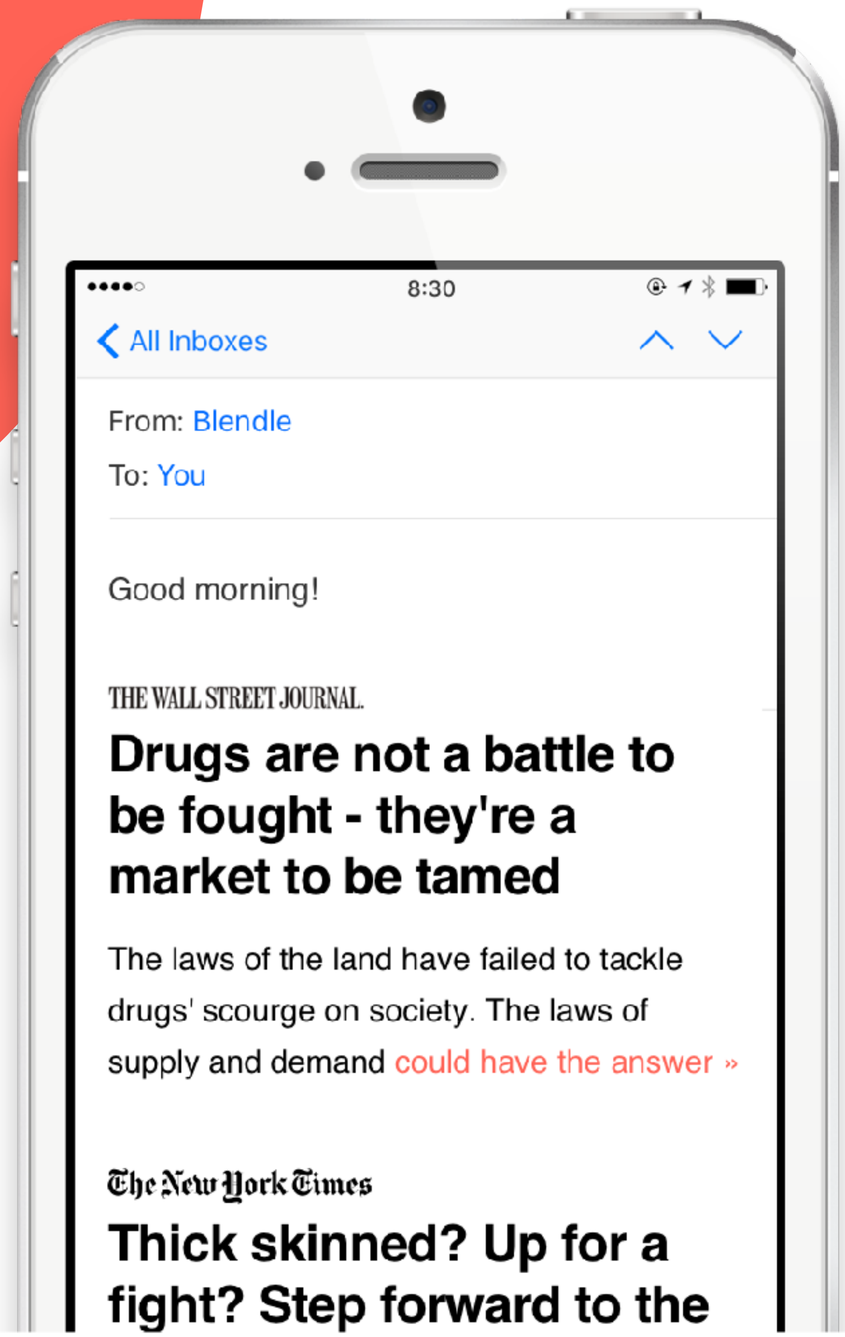
- **Single** consumption
- Less niche than music
- 75.000.000 users
- **Long** shelf life
- Feedback: explicit ratings

Articles

- **Single** consumption
- Very **short** shelf-life
  - **Cold start problem, Collaborative Filtering won't work**
- **Diversity** of suggestions very important
  - **Content Based Recommendation won't work**
- A lot of information **overlap**
- Feedback: implicit through reading behaviour, refunds, also 'like' option

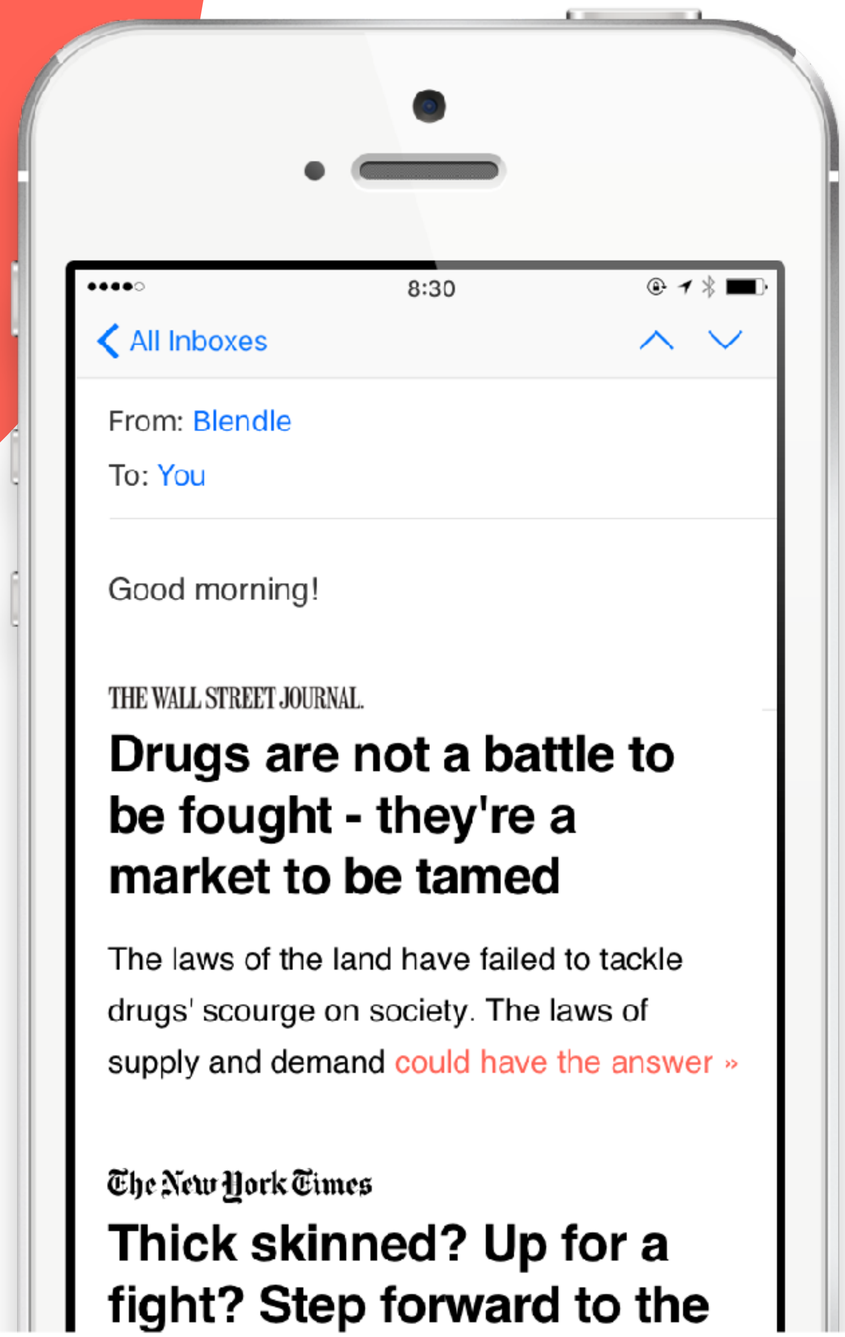
Cold start problem





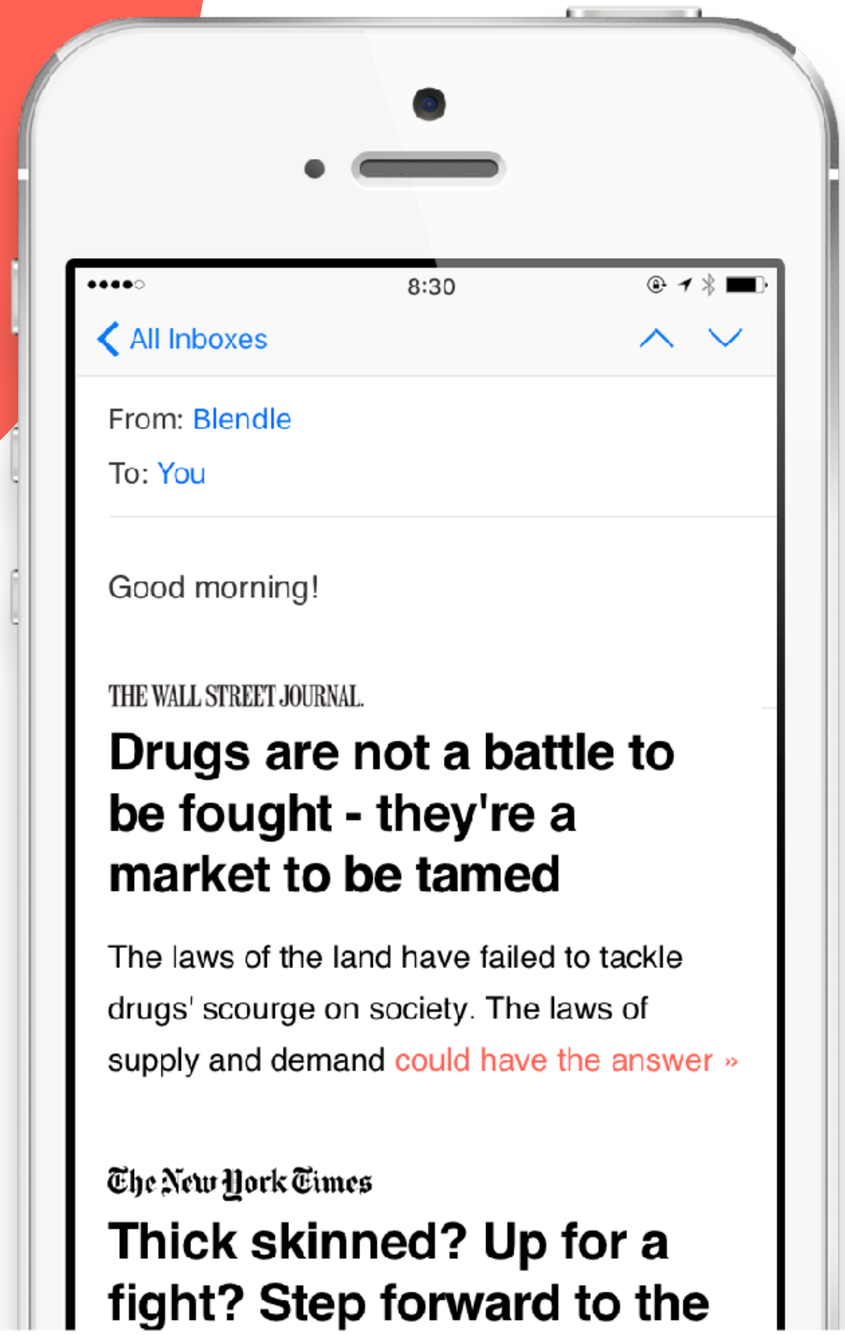
Cold start problem

- >6K new articles every night



Cold start problem

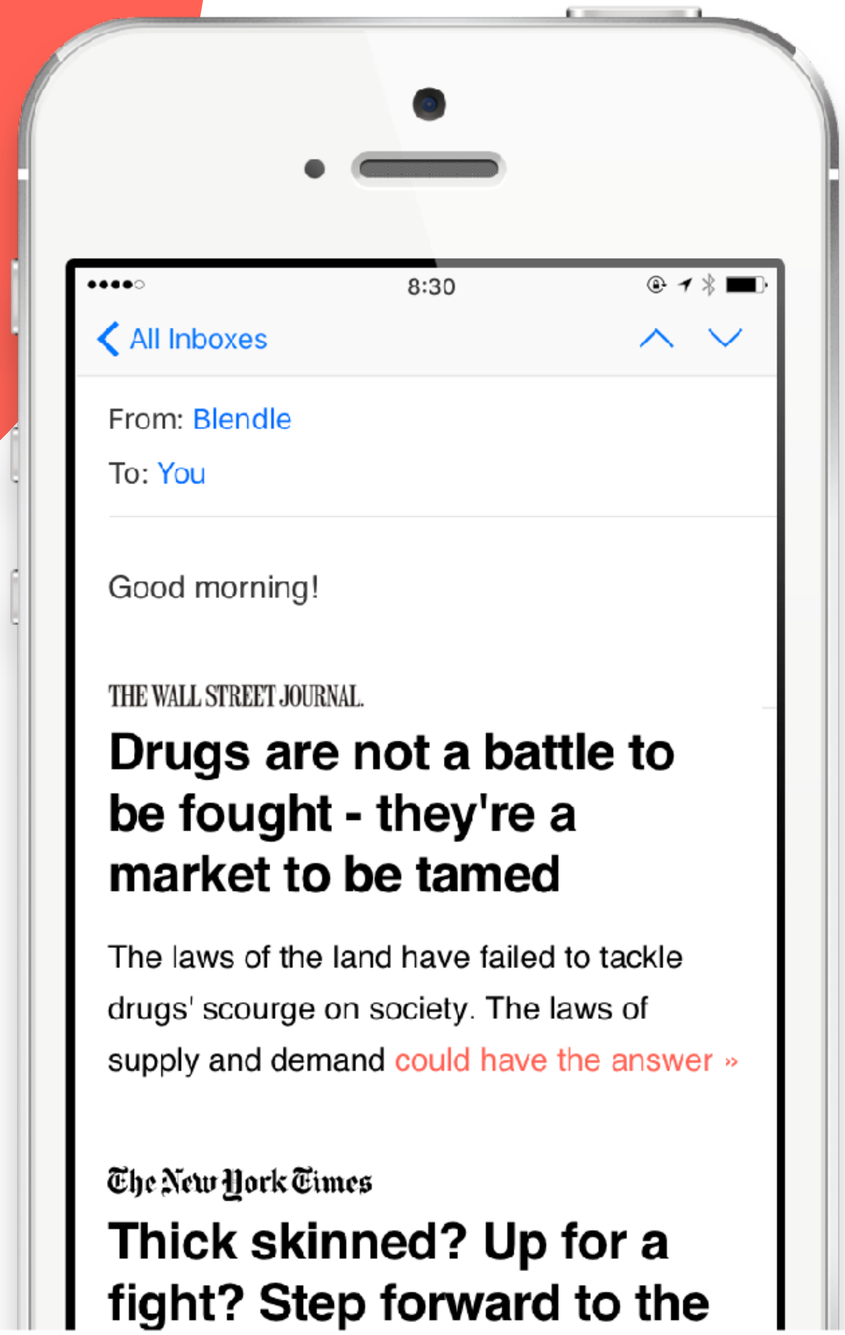
- >6K new articles every night
- Our newsletter is an important traffic driver

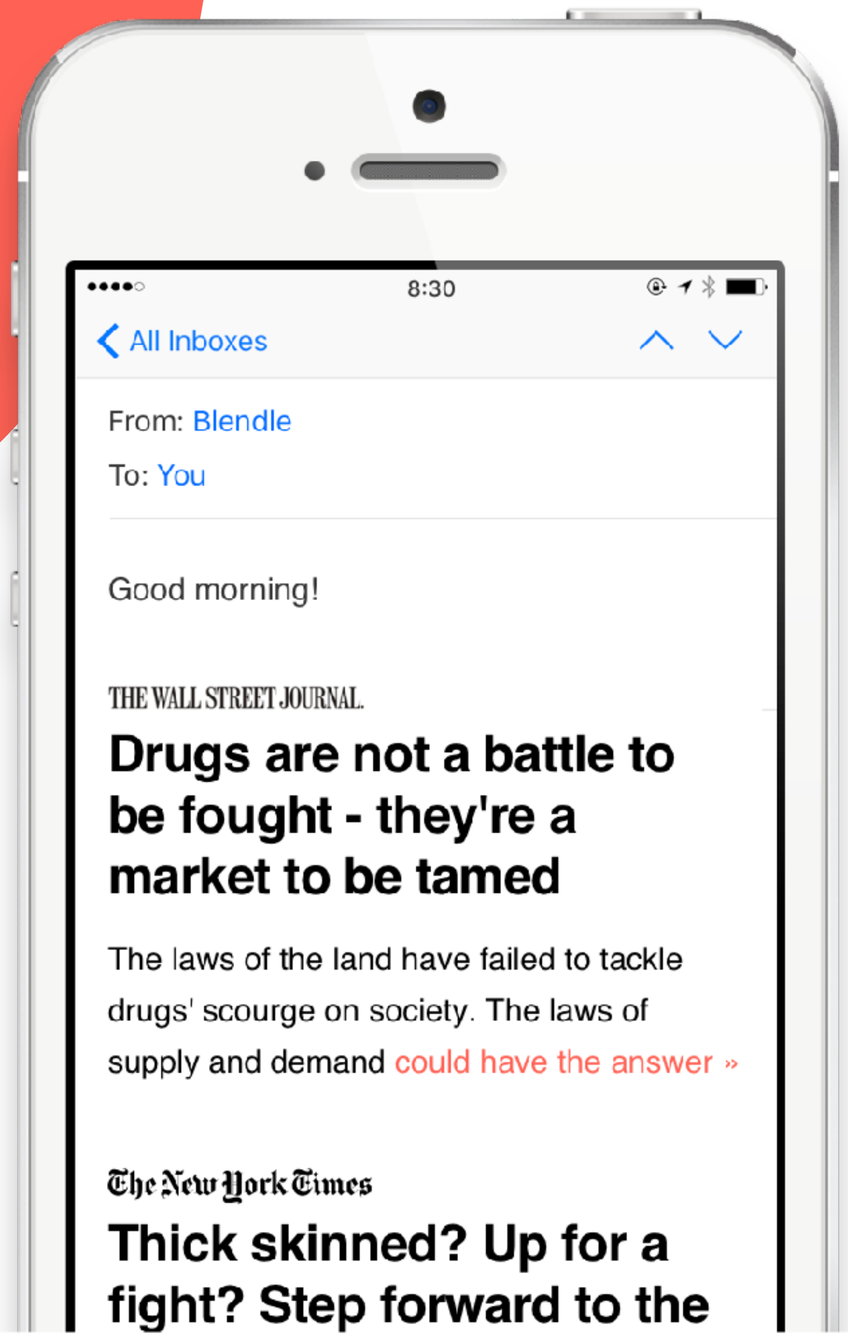


Cold start problem

- >6K new articles every night
- Our newsletter is an important traffic driver
- No usage info to *rank* the newsletter before we *send* the newsletter

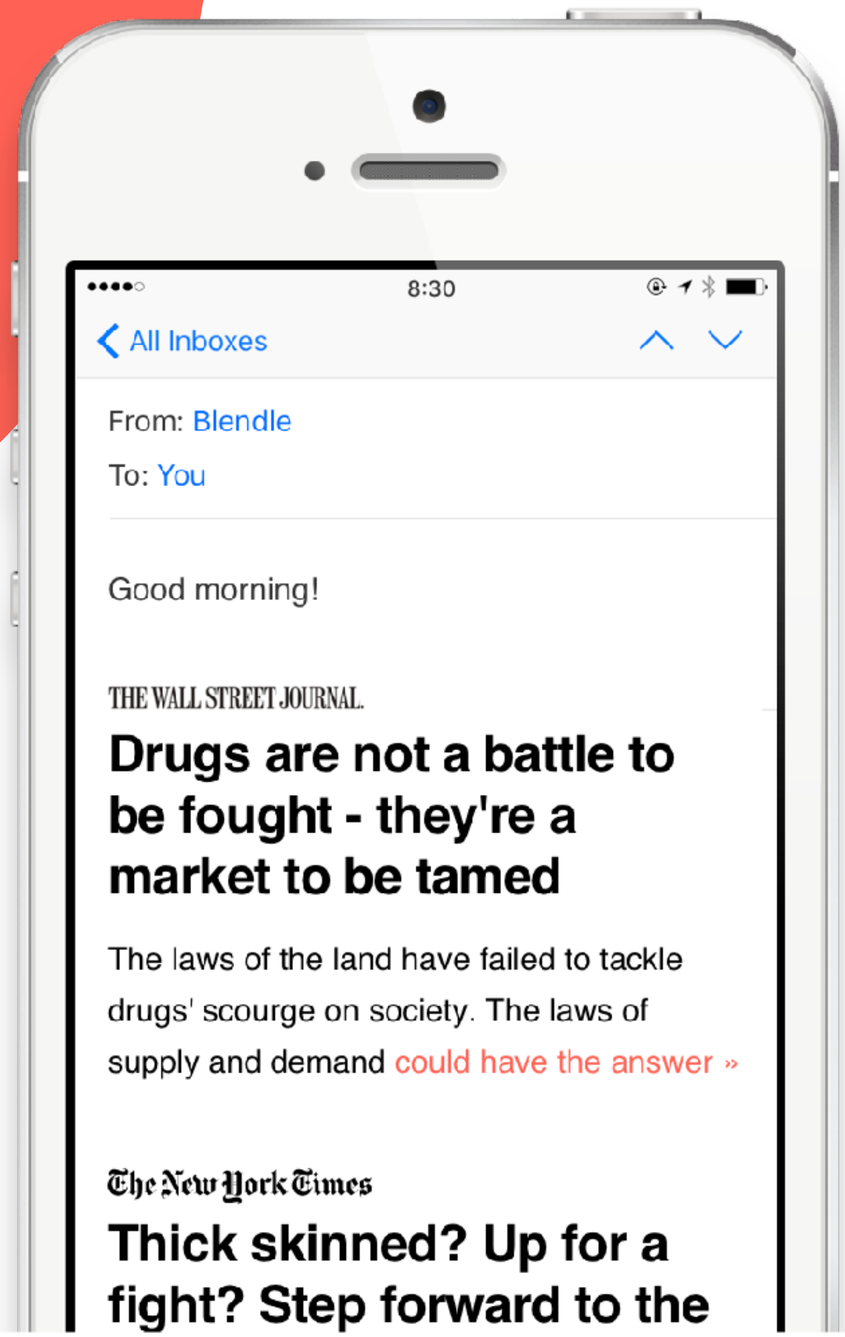
Timing problem





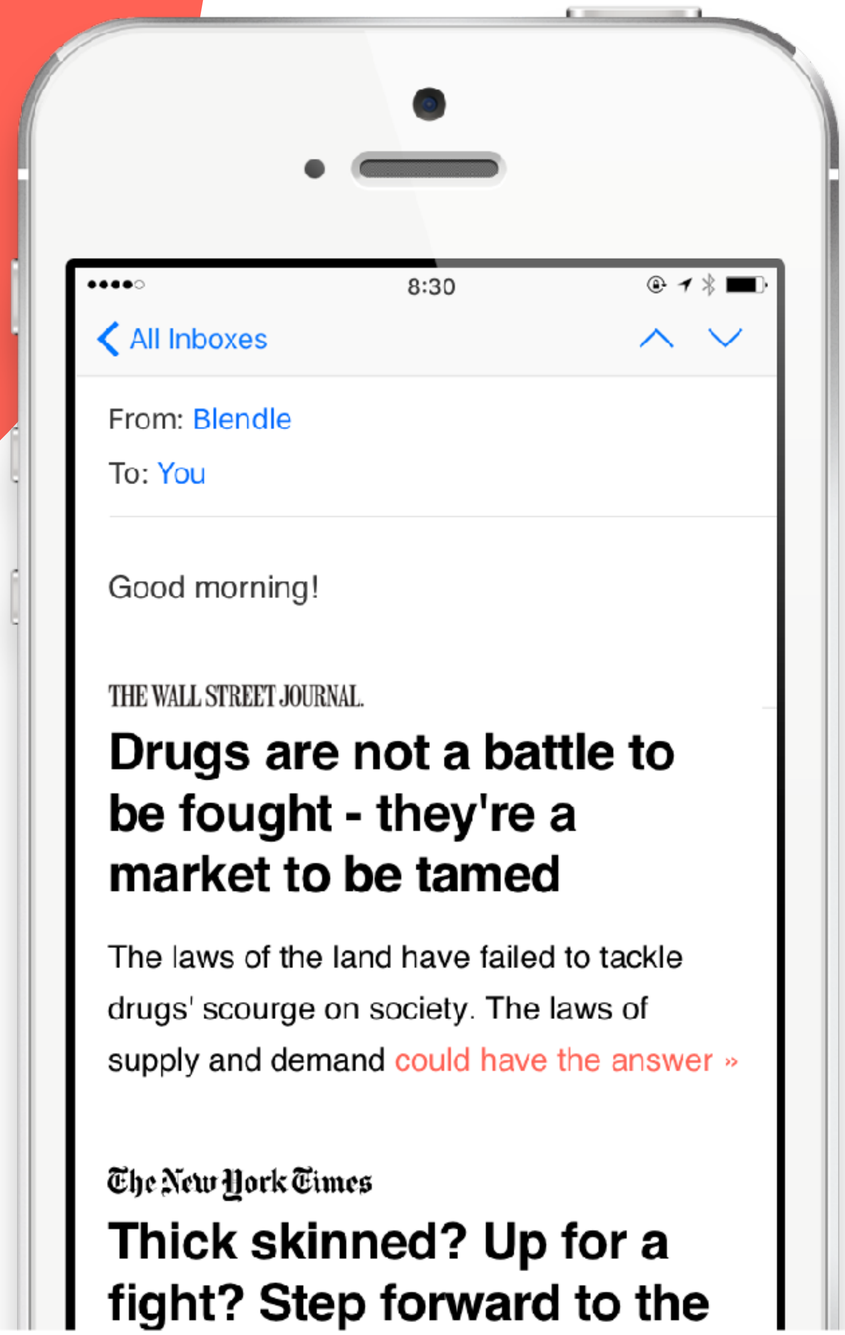
Timing problem

- Our editors wake up at 5am and are done reading at 8am



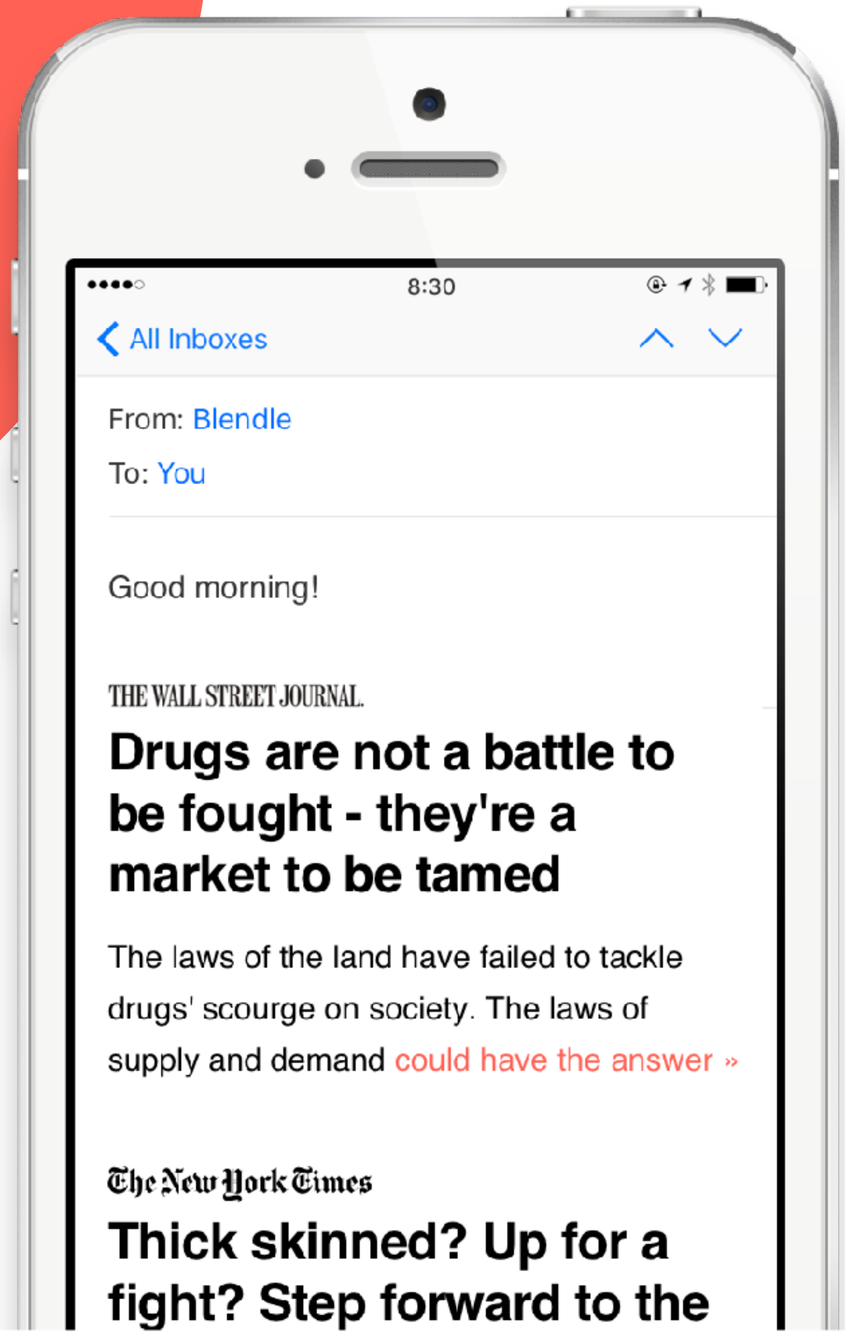
Timing problem

- Our editors wake up at 5am and are done reading at 8am
- Which is also when we want to send our newsletter



Timing problem

- Our editors wake up at 5am and are done reading at 8am
- Which is also when we want to send our newsletter
- We simply can’t wait for a batch process

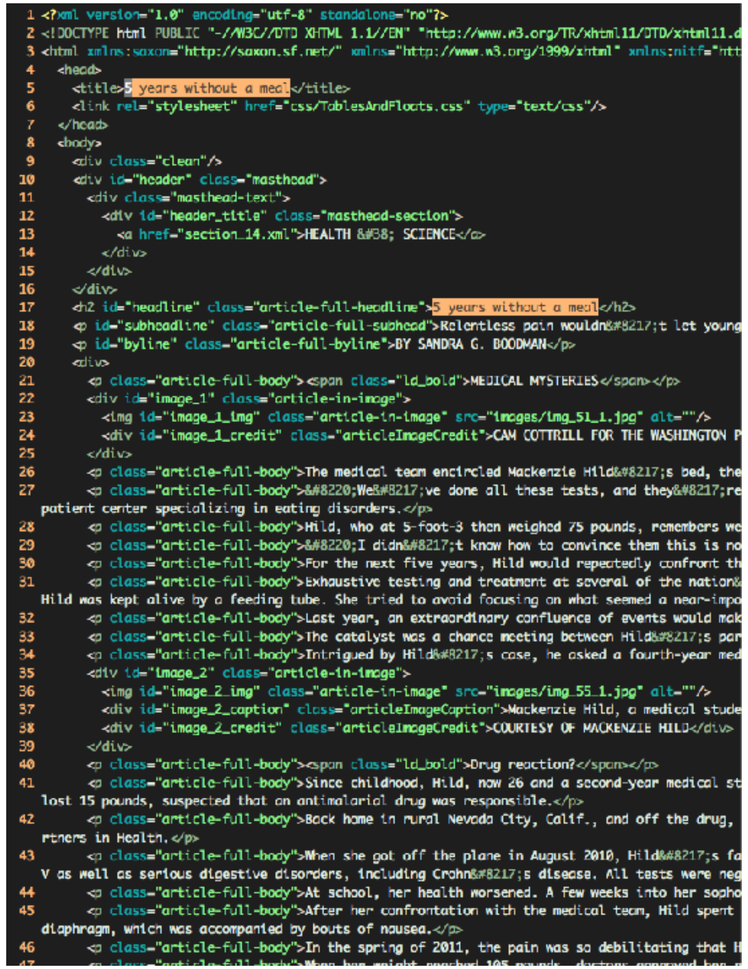


Cold start problem

- >6K new articles every night
- Our newsletter is an important traffic driver
- No usage info to *rank* the newsletter before we *send* the newsletter

Birth of a Blendle article

Delivered as PDF



```
<?xml version="1.0" encoding="utf-8" standalone="no"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN" "http://www.w3.org/TR/xhtml11/DTD/xhtml11.d
<html xmlns:saxon="http://saxon.sf.net/" xmlns="http://www.w3.org/1999/xhtml" xmlns:nitf="htt
  <head>
    <title>5 years without a meal</title>
    <link rel="stylesheet" href="css/TablesAndFloats.css" type="text/css"/>
  </head>
  <body>
    <div class="clean"/>
    <div id="header" class="masthead">
      <div class="masthead-text">
        <div id="header_title" class="masthead-section">
          <a href="section_14.xml">HEALTH &#38; SCIENCE</a>
        </div>
      </div>
    </div>
    <h2 id="headline" class="article-full-headline">5 years without a meal</h2>
    <p id="subheadline" class="article-full-subhead">Relentless pain wouldn&#8217;t let young
    <p id="byline" class="article-full-byline">BY SANDRA G. BOODMAN</p>
    <div>
      <p class="article-full-body"><span class="ld_bold">MEDICAL MYSTERIES</span></p>
      <div id="image_1" class="article-in-image">
        <img id="image_1_img" class="article-in-image" src="images/img_51_1.jpg" alt=""/>
        <div id="image_1_credit" class="articleImageCredit">CAM COTTRILL FOR THE WASHINGTON P
      </div>
      <p class="article-full-body">The medical team encircled Mackenzie Hild&#8217;s bed, the
      <p class="article-full-body">&#8220;We&#8217;ve done all these tests, and they&#8217;re
patient center specializing in eating disorders.</p>
      <p class="article-full-body">Hild, who at 5-foot-3 then weighed 75 pounds, remembers we
      <p class="article-full-body">&#8220;I didn&#8217;t know how to convince them this is no
      <p class="article-full-body">For the next five years, Hild would repeatedly confront th
      <p class="article-full-body">Exhaustive testing and treatment at several of the nation&
Hild was kept alive by a feeding tube. She tried to avoid focusing on what seemed a near-impo
      <p class="article-full-body">Last year, an extraordinary confluence of events would mak
      <p class="article-full-body">The catalyst was a chance meeting between Hild&#8217;s par
      <p class="article-full-body">Intrigued by Hild&#8217;s case, he asked a fourth-year med
      <div id="image_2" class="article-in-image">
        <img id="image_2_img" class="article-in-image" src="images/img_55_1.jpg" alt=""/>
        <div id="image_2_caption" class="articleImageCaption">Mackenzie Hild, a medical stude
        <div id="image_2_credit" class="articleImageCredit">COURTESY OF MACKENZIE HILD</div>
      </div>
      <p class="article-full-body"><span class="ld_bold">Drug reaction?</span></p>
      <p class="article-full-body">Since childhood, Hild, now 26 and a second-year medical st
lost 15 pounds, suspected that an antimalarial drug was responsible.</p>
      <p class="article-full-body">Back home in rural Nevada City, Calif., and off the drug,
rtners in Health.</p>
      <p class="article-full-body">When she got off the plane in August 2010, Hild&#8217;s fa
V as well as serious digestive disorders, including Crohn&#8217;s disease. All tests were neg
      <p class="article-full-body">At school, her health worsened. A few weeks into her sopho
      <p class="article-full-body">After her confrontation with the medical team, Hild spent
diaphragm, which was accompanied by bouts of nausea.</p>
      <p class="article-full-body">In the spring of 2011, the pain was so debilitating that H
```

Processed as epub

```
"format version": 5,
"id": "bnl-washingtonpost-20160927-51_3",
"date": "2016-09-27T00:00:00+00:00",
"provider": {
    "id": 'washingtonpost'
},
"body": [
    {
        "type": 'hl1",
        "content': "5 years without a meal"
    },
    {
        "type": 'hl2",
        "content': "Relentless pain wouldnâ€™t let young woman eat; only a
        feeding tube kept her alive"
    },
    {
        "type": 'byline",
        "content': "BY SANDRA G. BOODMAN"
    },
    {
        "type": 'p",
        "content': "<b>MEDICAL MYSTERIES</b>'
    },
    {
        "type": 'p",
        "content': "The medical team encircled Mackenzie Hildâ€™s bed, their
        somber expressions reflecting the gravity of the news they were
        about to impart to the Harvard sophomore and her mother, newly
        arrived from California."
    },
    {
        "type": 'p",
        "content': "â€œWeâ€™ve done all these tests, and theyâ€™re all
        normal,â€ Hild recalls one doctor at the renowned Boston hospital
        telling them. To treat Hildâ€™s life-threatening weight loss, which
        the 19-year-old claimed was the result of searing abdominal pain
        triggered by eating, doctorsâ€¦"
    }
],
"images": [
    {
        "_links": {
            "small": {
                "href": "https://static.blendle.nl/publication/washingtonpost
                /2016/09/27/item/51_3/version/1/image/small/90d02c1b211b8c2ab
                3cd58d1addeb6894c918ac3.jpg",
                "width": 231,
                "height": 300
            },
            "medium": {
                "href": "https://static.blendle.nl/publication/washingtonpost
                /2016/09/27/item/51_3/version/1/image/medium/90d02c1b211b8c2a
```

To the Blendle format

Blendle @anneschuth

Blendle @anneschuth

Article enrichment pipeline

Language detection → Tokenization

Article enrichment pipeline

Language detection

Tokenization

Stylometry

Length, word variation, vocabulary richness, …

Article enrichment pipeline

Language detection

Tokenization

Stylometry

Length, word variation, vocabulary richness, …

Sentiment analysis

Polarity scores (negative, positive)

Article enrichment pipeline

Het Nieuwsblad

Waarom eindigde Quentin als kanonnenvoer voor IS?

Language detection

Tokenization

Stylometry

Sentiment analysis

PoS-tagging

Length, word variation, vocabulary richness, …

Polarity scores (negative, positive)

Article enrichment pipeline

Language detection

Tokenization

Stylometry

Sentiment analysis

PoS-tagging

Named Entity Recognition

Length, word variation, vocabulary richness, …

Polarity scores (negative, positive)

Locations, persons, organizations

Article enrichment pipeline

Waarom eindigde Quentin als kanonnenvoer voor IS?

Language detection

Tokenization

Stylometry

Sentiment analysis

PoS-tagging

Author extraction

Named Entity Recognition

Length, word variation, vocabulary richness, …

Polarity scores (negative, positive)

Locations, persons, organizations

Article enrichment pipeline

Waarom eindigde Quentin als kanonnenvoer voor IS?

Language detection

Tokenization

Stylometry

Length, word variation, vocabulary richness, …

Sentiment analysis

Polarity scores (negative, positive)

PoS-tagging

Named Entity Recognition

Locations, persons, organizations

Author extraction

Semantic linking

Wikipedia concepts

Article enrichment pipeline

Language detection

Tokenization

Stylometry

Length, word variation, vocabulary richness, …

Sentiment analysis

Polarity scores (negative, positive)

PoS-tagging

Named Entity Recognition

Locations, persons, organizations

Author extraction

Semantic linking

Wikipedia concepts

Topic modeling

Spark LDA EM

Article enrichment pipeline

Blendle

```
Feature: Enrich item pipeline

  Scenario: Enrich an item
    Given I have the following item data
      """json
      {"_links":{"self":{"href":"s3://static.blendle.nl/publication/par/2015/01/27/item/3895280/version/1/private/content.json"}},
      "body":[{"content":"Smulpapen","type":"kicker"},
      {"content":"Tom Kellerhuis stoomt uit allerhande keukens bedrieglijk simpele gerechten klaar.","type":"intro"},
      {"content":"Dronken vegetariër","type":"hl1"},{"content":"TOM KELLERHUIS","type":"byline"},
      {"content":"Ik ken een flink aantal restaurateurs dat gruwt bij het woord vegetariër. Ik ken er zelfs een paar die
      hem bewust uit hun zaak weren. \"U eet geen vlees of vis? Wat doet u hier dan?\" hoorde ik ooit als pissig antwoord
      op de vraag of er soms ook gerechten zónder vis of vlees op de kaart stonden.","type":"p"},{"content":"Een goeie
      kok wil nu eenmaal laten zien wat hij in huis heeft. Dat zit hem vooral in de vlees- en visbereidingen, of in de
      sauzen, waarin vaak ook vlees- en visproducten zijn verwerkt. Vanuit dat oogpunt snap ik de kok wel. Maar steeds
      meer mensen kiezen ervoor om minder vlees of vis te consumeren, om hen moverende redenen. De één denkt aan het
      milieu en de consumptiemaatschappij, de ander aan het dierenwelzijn, weer een ander louter aan zijn eigen
      gezondheid.","type":"p"},
      {"content":"Hoewel er aan het woord vegetariër nog steeds een sektarisch randje kleeft, mogen we toch wel stellen
      dat hij tegenwoordig lang zo suf niet meer is als het aloude stereotype beeld: het op sandalen - waarin platvoeten
      in geitenwollensokken - sjokkende en van elke humor gespeende tiepje, met linnen draagtas op zoek naar gedroogde
      sojabonen en tempé voor de avondmaaltijd.","type":"p"},{"content":"Ik kreeg laatst het receptenboek De dronken
      vegetariër van Floor Overgoor, Hanny Roskamp en Yannick Slagter toegezonden, dat gehakt maakt van dat oude cliché.
      Allemaal vegagerechten met een tic (en soms flink wat, zoals in bijgaand recept. Roskamp schrijft in haar voorwoord:
      'We zijn gaan beseffen dat die dronken vegetariër een betere ambassadeur is voor een diervriendelijke wereld dan een
      vegetariër die zichzelf opsluit achter een muur van morele superioriteit.'","type":"p"},{"content":"Welaan, het
      beeld is dus aan het kantelen. Ik hoop snel eens zo'n dronken vegetariër tegen te komen!","type":"p"},
      {"content":"Reageren? t.kellerhuis@parool.nl","type":"p"},{"content":"Wijnfondue","type":"ph"},
      {"content":"(voor 2 personen)","type":"p"},{"content":"1/2 liter vega-bouillon, 1/2 fles pinot noir, 1/2 aubergine
      in plakken, 200 g oesterzwammen, 200 g bospaddenstoelen, 1 courgette in plakken, 200 g tofoe in repen.","type":"p"},
      {"content":"Simpeler en romantischer dan dit wordt het niet. Maak de bouillon, breng aan de kook, giet de wijn erbij
      en breng opnieuw aan de kook (maar laat niet doorkoken: zonde van de wijn!).","type":"p"},{"content":"Zet dit
      heerlijke vocht op een fonduestel, warmhoudplaatje of theelicht. Gooi een paar flinke handen groentes en tofoe in
      het hete vocht en haal de stukken er na een minuut of 5 weer uit; liefst met zo'n Chinees fonduemandje, maar een
      vork werkt ook prima.","type":"p"},{"content":"Laat de volgende portie groente en tofoe garen in het vocht, terwijl
      je de eerste portie opeet.","type":"p"},{"content":"Eet de stukjes groente met brood en knoflookboter. Dippen in
      knoflooksaus kan uiteraard ook.","type":"p"}],
      "date":"2015-01-27T00:00:00+00:00",
      "format_version":4,
      "id":"bnl-par-20150127-3895280",
      "issue":{"id":"bnl-par-20150127"},
      "item_index":70,
      "provider":{"id":"par"}}
      """
```

```
Feature: Enrich item pipeline

  Scenario: Enrich an item
    Given I have the following item data
      """json
      {"_links":{"self":{"href":"s3://static.blendle.nl/publication/par/2015/01/27/item/3895280/version/1/private/content.json"}}
      "body":[{"cont
      {"content":"To
      {"content":"Dr
      {"content":"Ik
      hem bewust uit
      op de vraag of
      kok wil nu een
      sauzen, waarin
      meer mensen ki
      milieu en de o
      gezondheid.","
      {"content":"Ho
      dat hij tegenw
      in geitenwolle
      sojabonen en t
      vegetariër var
      Allemaal vegag
      'We zijn gaan
      vegetariër die
      beeld is dus a
      {"content":"Re
      {"content":"(v
      in plakken, 20
      {"content":"Si
      en breng opnie
      heerlijke voch
      het hete vocht
      vork werkt ook
      je de eerste p
      knoflooksaus k
      "date":"2015-0
      "format_versio
      "id":"bnl-par-
      "issue":{"id":
      "item_index":7
      "provider":{"i
      """
```

Het Parool

dinsdag 27 januari 2015 · 2 minuten

Smulpapen

Dronken vegetariër

Tom Kellerhuis stoomt uit allerhande keukens bedrieglijk simpele gerechten klaar.

TOM KELLERHUIS

Ik ken een flink aantal restaurateurs dat gruwt bij het woord vegetariër. Ik ken er zelfs een paar die hem bewust uit hun zaak weren. "U eet geen vlees of vis? Wat doet u hier dan?" hoorde ik ooit als pissig antwoord op de vraag of er soms ook gerechten zónder vis of vlees op de kaart stonden.

Een goeie kok wil nu eenmaal laten zien wat hij in huis heeft. Dat zit hem vooral in de vlees- en visbereidingen, of in de sauzen, waarin vaak ook vlees- en visproducten zijn verwerkt. Vanuit dat oogpunt snap ik de kok wel. Maar steeds meer mensen kiezen ervoor om minder vlees of vis te consumeren, om hen moverende redenen. De één denkt aan het milieu en de consumptiemaatschappij, de ander aan het dierenwelzijn, weer een ander louter aan zijn eigen gezondheid.

Hoewel er aan het woord vegetariër nog steeds een sektarisch randje kleeft, mogen we toch wel stellen dat hij tegenwoordig lang zo suf niet meer is als het aloude stereotype beeld: het op sandalen - waarin platvoeten in geitenwollensokken - sjokkende en van elke humor gespeende tiepje, met linnen draagtas op zoek naar gedroogde sojabonen en tempé voor de avondmaaltijd.

Ik kreeg laatst het receptenboek De dronken vegetariër van Floor Overgoor, Hanny Roskamp en Yannick Slagter toegezonden, dat gehakt maakt van dat oude cliché. Allemaal vegagerechten met een tic (en soms flink wat, zoals in bijgaand recept. Roskamp schrijft in haar voorwoord: 'We zijn gaan beseffen dat die dronken vegetariër een betere ambassadeur is voor een diervriendelijke wereld da zichzelf opsluit achter een superioriteit.'

Welaan, het beeld is dus aan snel eens zo'n dronken vegetariër

Reageren? t.kellerhuis@paroc

Wijnfondue

(voor 2 personen)

1/2 liter vega-bouillon, 1/2 aubergine in plakken, 200 g oe bospaddenstoelen, 1 courgette in in repen.

Simpeler en romantischer da Maak de bouillon, breng aan de k en breng opnieuw aan de k doorkoken: zonde van de wijn!).

Zet dit heerlijke vocht warmhoudplaatje of theelicht. handen groentes en tofoe in het stukken er na een minuut of 5 w

```
Feature: Enrich item pipeline

  Scenario: Enrich an item
    Given I have the following item data
      """json
      {"_links":{"self":{"href":"s3://static.blendle.nl/publication/par/2015/01/27/item/3895280/version/1/private/content.json"}},
      "body":[{"content":"Smulpapen","type":"kicker"},
      {"content":"Tom Kellerhuis stoomt uit allerhande keukens bedrieglijk simpele gerechten klaar.","type":"intro"},
      {"content":"Dronken vegetariër","type":"hl1"},{"content":"TOM KELLERHUIS","type":"byline"},
      {"content":"Ik ken een flink aantal restaurateurs dat gruwt bij het woord vegetariër. Ik ken er zelfs een paar die
      hem bewust uit hun zaak weren. \"U eet geen vlees of vis? Wat doet u hier dan?\" hoorde ik ooit als pissig antwoord
      op de vraag of er soms ook gerechten zónder vis of vlees op de kaart stonden.","type":"p"},{"content":"Een goeie
      kok wil nu eenmaal laten zien wat hij in huis heeft. Dat zit hem vooral in de vlees- en visbereidingen, of in de
      sauzen, waarin vaak ook vlees- en visproducten zijn verwerkt. Vanuit dat oogpunt snap ik de kok wel. Maar steeds
      meer mensen kiezen ervoor om minder vlees of vis te consumeren, om hen moverende redenen. De één denkt aan het
      milieu en de consumptiemaatschappij, de ander aan het dierenwelzijn, weer een ander louter aan zijn eigen
      gezondheid.","type":"p"},
      {"content":"Hoewel er aan het woord vegetariër nog steeds een sektarisch randje kleeft, mogen we toch wel stellen
      dat hij tegenwoordig lang zo suf niet meer is als het aloude stereotype beeld: het op sandalen - waarin platvoeten
      in geitenwollensokken - sjokkende en van elke humor gespeende tiepje, met linnen draagtas op zoek naar gedroogde
      sojabonen en tempé voor de avondmaaltijd.","type":"p"},{"content":"Ik kreeg laatst het receptenboek De dronken
      vegetariër van Floor Overgoor, Hanny Roskamp en Yannick Slagter toegezonden, dat gehakt maakt van dat oude cliché.
      Allemaal vegagerechten met een tic (en soms flink wat, zoals in bijgaand recept. Roskamp schrijft in haar voorwoord:
      'We zijn gaan beseffen dat die dronken vegetariër een betere ambassadeur is voor een diervriendelijke wereld dan een
      vegetariër die zichzelf opsluit achter een muur van morele superioriteit.'","type":"p"},{"content":"Welaan, het
      beeld is dus aan het kantelen. Ik hoop snel eens zo'n dronken vegetariër tegen te komen!","type":"p"},
      {"content":"Reageren? t.kellerhuis@parool.nl","type":"p"},{"content":"Wijnfondue","type":"ph"},
      {"content":"(voor 2 personen)","type":"p"},{"content":"1/2 liter vega-bouillon, 1/2 fles pinot noir, 1/2 aubergine
      in plakken, 200 g oesterzwammen, 200 g bospaddenstoelen, 1 courgette in plakken, 200 g tofoe in repen.","type":"p"},
      {"content":"Simpeler en romantischer dan dit wordt het niet. Maak de bouillon, breng aan de kook, giet de wijn erbij
      en breng opnieuw aan de kook (maar laat niet doorkoken: zonde van de wijn!).","type":"p"},{"content":"Zet dit
      heerlijke vocht op een fonduestel, warmhoudplaatje of theelicht. Gooi een paar flinke handen groentes en tofoe in
      het hete vocht en haal de stukken er na een minuut of 5 weer uit; liefst met zo'n Chinees fonduemandje, maar een
      vork werkt ook prima.","type":"p"},{"content":"Laat de volgende portie groente en tofoe garen in het vocht, terwijl
      je de eerste portie opeet.","type":"p"},{"content":"Eet de stukjes groente met brood en knoflookboter. Dippen in
      knoflooksaus kan uiteraard ook.","type":"p"}],
      "date":"2015-01-27T00:00:00+00:00",
      "format_version":4,
      "id":"bnl-par-20150127-3895280",
      "issue":{"id":"bnl-par-20150127"},
      "item_index":70,
      "provider":{"id":"par"}}
      """
```

When I send it through the enrichment pipeline language|tokenizer|postagger|ner|stylometry|author|semanticizer|sentiment
Then I should have the following enrichments in the result

enrichment	value
language	nl
author	TOM KELLERHUIS

And I should have the following stylometry enrichments in the result

enrichment	value
number_of_paragraphs	12
number_of_words	432
article_length	2542
vocab_richness	40.5376820773
word_variation	0.585648148148
avg_word_length	4.98611111111
word_size_short	0.553240740741
word_size_medium	0.391203703704
word_size_long	0.0555555555556
hapax_legomena	0.4375
hapax_dislegomena	0.0763888888889
number_of_images	1

And I should have the following semantic links in the result

url	title	text	senseProbability	count	norm
http://nl.wikipedia.org/wiki/Avondmaaltijd	Avondmaaltijd	avondmaaltijd	0.36666666666666664	1	http://dbpedia.org/resource/Supper
http://nl.wikipedia.org/wiki/Tic	Tic	tic	1.0416666666666667	1	http://dbpedia.org/resource/Tic
http://nl.wikipedia.org/wiki/Platvoeten	Platvoeten	platvoeten	0.47058823529411764	1	http://dbpedia.org/resource/Flat_feet
http://nl.wikipedia.org/wiki/Groente	Groente	groente	0.41304347826086957	2	http://dbpedia.org/resource/Vegetable
http://nl.wikipedia.org/wiki/Sojaboon	Sojaboon	sojabonen	0.6206896551724138	1	http://dbpedia.org/resource/Soybean
http://nl.wikipedia.org/wiki/Vegetarisme	Vegetarisme	vegetariër	2.6956521739130435	6	http://dbpedia.org/resource/Vegetarianism
http://nl.wikipedia.org/wiki/Pinot%20noir	Pinot noir	pinot noir	0.8827586206896552	1	http://dbpedia.org/resource/Pinot_noir
http://nl.wikipedia.org/wiki/Knoflooksaus	Knoflooksaus	knoflooksaus	0.35294117647058826	1	
http://nl.wikipedia.org/wiki/Consumptiemaatschappij	Consumptiemaatschappij	consumptiemaatschappij	0.5121951219512195	1	
http://nl.wikipedia.org/wiki/Courgette	Courgette	courgette	0.8421052631578947	1	http://dbpedia.org/resource/Zucchini
http://nl.wikipedia.org/wiki/Tofoe	Tofoe	tofoe	0.8275862068965517	3	http://dbpedia.org/resource/Tofu
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	- bouillon	0.5752212389380531	1	http://dbpedia.org/resource/Broth
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	bouillon	0.5752212389380531	2	http://dbpedia.org/resource/Broth
http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29	Aubergine (plant)	aubergine	0.8666666666666667	1	http://dbpedia.org/resource/Eggplant
http://nl.wikipedia.org/wiki/Kensoort	Kensoort	ken	0.8787878787878788	2	

And I should have the following named entities in the result

entity	tag	count
TOM KELLERHUIS	PER	1
FLOOR OVERGOOR	PER	1
HANNY ROSKAMP	PER	2
YANNICK SLAGTER	PER	1
WIJNFONDUE	LOC	1
SIMPELER	PER	1

When I send it through the enrichment pipeline language|tokenizer|postagger|ner|stylometry|author|semanticizer|sentiment
Then I should have the following enrichments in the result

enrichment	value
language	nl
author	TOM KELLERHUIS

avg_word_length	4.98611111111
word_size_short	0.553240740741
word_size_medium	0.391203703704
word_size_long	0.0555555555556
hapax_legomena	0.4375
hapax_dislegomena	0.0763888888889
number_of_images	1

And I should have the following semantic links in the result

url	title	text	senseProbability	count	norm
http://nl.wikipedia.org/wiki/Avondmaaltijd	Avondmaaltijd	avondmaaltijd	0.36666666666666664	1	http://dbpedia.org/resource/Supper
http://nl.wikipedia.org/wiki/Tic	Tic	tic	1.0416666666666667	1	http://dbpedia.org/resource/Tic
http://nl.wikipedia.org/wiki/Platvoeten	Platvoeten	platvoeten	0.47058823529411764	1	http://dbpedia.org/resource/Flat_feet
http://nl.wikipedia.org/wiki/Groente	Groente	groente	0.41304347826086957	2	http://dbpedia.org/resource/Vegetable
http://nl.wikipedia.org/wiki/Sojaboon	Sojaboon	sojabonen	0.6206896551724138	1	http://dbpedia.org/resource/Soybean
http://nl.wikipedia.org/wiki/Vegetarisme	Vegetarisme	vegetariër	2.6956521739130435	6	http://dbpedia.org/resource/Vegetarianism
http://nl.wikipedia.org/wiki/Pinot%20noir	Pinot noir	pinot noir	0.8827586206896552	1	http://dbpedia.org/resource/Pinot_noir
http://nl.wikipedia.org/wiki/Knoflooksaus	Knoflooksaus	knoflooksaus	0.35294117647058826	1	
http://nl.wikipedia.org/wiki/Consumptiemaatschappij	Consumptiemaatschappij	consumptiemaatschappij	0.5121951219512195	1	
http://nl.wikipedia.org/wiki/Courgette	Courgette	courgette	0.8421052631578947	1	http://dbpedia.org/resource/Zucchini
http://nl.wikipedia.org/wiki/Tofoe	Tofoe	tofoe	0.8275862068965517	3	http://dbpedia.org/resource/Tofu
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	- bouillon	0.5752212389380531	1	http://dbpedia.org/resource/Broth
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	bouillon	0.5752212389380531	2	http://dbpedia.org/resource/Broth
http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29	Aubergine (plant)	aubergine	0.8666666666666667	1	http://dbpedia.org/resource/Eggplant
http://nl.wikipedia.org/wiki/Kensoort	Kensoort	ken	0.8787878787878788	2	

And I should have the following named entities in the result

entity	tag	count
TOM KELLERHUIS	PER	1
FLOOR OVERGOOR	PER	1
HANNY ROSKAMP	PER	2
YANNICK SLAGTER	PER	1
WIJNFONDUE	LOC	1
SIMPELER	PER	1

When I send it through the enrichment pipeline language|tokenizer|postagger|ner|stylometry|author|semanticizer|sentiment

enrichment	value
number_of_paragraphs	12
number_of_words	432
article_length	2542
vocab_richness	40.5376820773
word_variation	0.585648148148
avg_word_length	4.98611111111
word_size_short	0.553240740741
word_size_medium	0.391203703704
word_size_long	0.0555555555556
hapax_legomena	0.4375
hapax_dislegomena	0.0763888888889
number_of_images	1

entity	tag	count
TOM KELLERHUIS	PER	1
FLOOR OVERGOOR	PER	1
HANNY ROSKAMP	PER	2
YANNICK SLAGTER	PER	1
WIJNFONDUE	LOC	1
SIMPELER	PER	1

	count	norm
564	1	http://dbpedia.org/resource/Supper
67	1	http://dbpedia.org/resource/Tic
764	1	http://dbpedia.org/resource/Flat_feet
957	2	http://dbpedia.org/resource/Vegetable
38	1	http://dbpedia.org/resource/Soybean
35	6	http://dbpedia.org/resource/Vegetarianism
52	1	http://dbpedia.org/resource/Pinot_noir
326	1	
95	1	
47	1	http://dbpedia.org/resource/Zucchini
17	3	http://dbpedia.org/resource/Tofu
31	1	http://dbpedia.org/resource/Broth
31	2	http://dbpedia.org/resource/Broth
67	1	http://dbpedia.org/resource/Eggplant
38	2	

url
http://nl.wikipedia.org/wiki/Avondmaaltijd
http://nl.wikipedia.org/wiki/Tic
http://nl.wikipedia.org/wiki/Platvoeten
http://nl.wikipedia.org/wiki/Groente
http://nl.wikipedia.org/wiki/Sojaboon
http://nl.wikipedia.org/wiki/Vegetarisme
http://nl.wikipedia.org/wiki/Pinot%20noir
http://nl.wikipedia.org/wiki/Knoflooksaus
http://nl.wikipedia.org/wiki/Consumptiemaatschappij
http://nl.wikipedia.org/wiki/Courgette
http://nl.wikipedia.org/wiki/Tofoe
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29
http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29
http://nl.wikipedia.org/wiki/Kensoort

e/Supper
e/Tic
e/Flat_feet
e/Vegetable
e/Soybean
e/Vegetarianism
e/Pinot_noir

e/Zucchini
e/Tofu
e/Broth
e/Broth
e/Eggplant

```
url
http://nl.wikipedia.org/wiki/Avondmaaltijd
```

Hoofdgerecht - Wikipedia

https://nl.wikipedia.org/wiki/Hoofdgerecht

WIKIPEDIA
De vrije encyclopedie

Hoofdpagina
Vind een artikel
Vandaag
Etalage
Categorieën
Recente wijzigingen
Nieuwe artikelen
Willekeurige pagina

Informatie
Gebruikersportaal
Snelcursus
Hulp en contact
Donaties

Hulpmiddelen
Links naar deze pagina

Een **hoofdgerecht** wordt in de klassieke keuken als het belangrijkste deel van een diner gezien. Het kan als gang op een menukaart te vinden zijn.

Kenmerk van een hoofdgerecht is dat er nadruk op wordt gelegd door deze uitgebreider dan de andere gangen te serveren. In het Frans wordt het wel uitgelegd als "pièce de resistance".[1] Het zwaarste deel van de maaltijd, meestal rondom vlees, gebraad of gevogelte. Tezamen met het vlees kunnen groenten, sauzen en bijgerechten worden geserveerd.

In de zin van "belangrijkste onderdeel van de maaltijd" kan een hoofdgerecht in de moderne keuken evengoed rondom bijvoorbeeld vis of een groenteschotel zijn gecreëerd.

Referenties [bewerken]

Hoofdgerecht van een kerstdiner in schalen op tafel.

https://nl.wikipedia.org/wiki/Avondmaaltijd

url
http://nl.wikipedia.org/wiki/Avondmaaltijd
http://nl.wikipedia.org/wiki/Tic
http://nl.wikipedia.org/wiki/Platvoeten
http://nl.wikipedia.org/wiki/Groente
http://nl.wikipedia.org/wiki/Sojaboon
http://nl.wikipedia.org/wiki/Vegetarisme
http://nl.wikipedia.org/wiki/Pinot%20noir
http://nl.wikipedia.org/wiki/Knoflooksaus
http://nl.wikipedia.org/wiki/Consumptiemaatschappij
http://nl.wikipedia.org/wiki/Courgette
http://nl.wikipedia.org/wiki/Tofoe
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29
http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29
http://nl.wikipedia.org/wiki/Kensoort

e/Supper
e/Tic
e/Flat_feet
e/Vegetable
e/Soybean
e/Vegetarianism
e/Pinot_noir

e/Zucchini
e/Tofu
e/Broth
e/Broth
e/Eggplant

url
http://nl.wikipedia.org/wiki/Avondmaaltijd

Avondmaaltijd - Wikipedia

https://nl.wikipedia.org/wiki/Avondmaaltijd

Чăвашла
Dansk
Deutsch
English
Esperanto
Español
Euskara
Suomi
Français
Galego
עברית
Bahasa Indonesia
Íslenska
Italiano
日本語
한국어

https://en.wikipedia.org/wiki/Supper

steeds gebruikt om de middagmaaltijd aan te duiden.[1] In Parijs ontstond de gewoonte om de belangrijkste maaltijd van de dag 's avonds te eten.[1]

Avondmaal, door Lucas Cranach

Vandaag betekent diner in België en Nederland een vaste maaltijd die met meerdere personen wordt genuttigd. In praktijk gebeurd dit vaak 's avonds. Al begint een diner soms al in de middaguren. Er wordt doorgaans meer tijd besteed om het maal te nuttigen. De locatie kan ook buitenshuis plaatsvinden. Wordt er tijdens het diner de gelegenheid geboden om te dansen, dan spreekt men van een *diner dansant*. Een officieel diner dat door een staatshoofd wordt gegeven, noemt men een *staatsdiner* of *staatsbanket*; alleen genodigden mogen hieraan deelnemen. Vaak wordt bij een staatsbezoek een staatsdiner aangeboden.

Een ander woord voor het avondmaal is het 'souper'. De oorsprong hiervan ligt bij het eten van de inhoud van een ketel (stoverij of soep), die een tijdlang had staan garen.[1]

Kerstdiner [bewerken]

http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29
http://nl.wikipedia.org/wiki/Kensoort

url
http://nl.wikipedia.org/wiki/Avondmaaltijd
http://nl.wikipedia.org/wiki/Tic
http://nl.wikipedia.org/wiki/Platvoeten
http://nl.wikipedia.org/wiki/Groente
http://nl.wikipedia.org/wiki/Sojaboon
http://nl.wikipedia.org/wiki/Vegetarisme
http://nl.wikipedia.org/wiki/Pinot%20noir
http://nl.wikipedia.org/wiki/Knoflooksaus
http://nl.wikipedia.org/wiki/Consumptiemaatschappij
http://nl.wikipedia.org/wiki/Courgette
http://nl.wikipedia.org/wiki/Tofoe
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29
http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29
http://nl.wikipedia.org/wiki/Kensoort

e/Supper
e/Tic
e/Flat_feet
e/Vegetable
e/Soybean
e/Vegetarianism
e/Pinot_noir

e/Zucchini
e/Tofu
e/Broth
e/Broth
e/Eggplant

norm
http://dbpedia.org/resource/Supper
http://dbpedia.org/resource/Tic
http://dbpedia.org/resource/Flat_feet
http://dbpedia.org/resource/Vegetable
http://dbpedia.org/resource/Soybean
http://dbpedia.org/resource/Vegetarianism
http://dbpedia.org/resource/Pinot_noir
http://dbpedia.org/resource/Zucchini
http://dbpedia.org/resource/Tofu
http://dbpedia.org/resource/Broth
http://dbpedia.org/resource/Broth
http://dbpedia.org/resource/Eggplant

nt | norm |
| --- |
| http://dbpedia.org/resource/Supper |
| http://dbpedia.org/resource/Tic |
| http://dbpedia.org/resource/Flat_feet |
| http://dbpedia.org/resource/Vegetable |
| http://dbpedia.org/resource/Soybean |
| http://dbpedia.org/resource/Vegetarianism |
| http://dbpedia.org/resource/Pinot_noir |
| |
| |
| http://dbpedia.org/resource/Zucchini |
| http://dbpedia.org/resource/Tofu |
| http://dbpedia.org/resource/Broth |
| http://dbpedia.org/resource/Broth |
| http://dbpedia.org/resource/Eggplant |
| |

```
When I send it through the enrichment pipeline language|tokenizer|postagger|ner|stylometry|author|semanticizer|sentiment
Then I should have the following  enrichments in the result
  | enrichment           | value           |
  | language             | nl              |
  | author               | TOM KELLERHUIS  |

And I should have the following stylometry enrichments in the result
  | enrichment           | value           |
  | number_of_paragraphs | 12              |
  | number_of_words      | 432             |
  | article_length       | 2542            |
  | vocab_richness       | 40.5376820773   |
  | word_variation       | 0.585648148148  |
  | avg_word_length      | 4.98611111111   |
  | word_size_short      | 0.553240740741  |
  | word_size_medium     | 0.391203703704  |
  | word_size_long       | 0.0555555555556 |
  | hapax_legomena       | 0.4375          |
  | hapax_dislegomena    | 0.0763888888889 |
  | number_of_images     | 1               |

And I should have the following semantic links in the result
  | url                                                | title                | text                | senseProbability     | count | norm                                  |
  | http://nl.wikipedia.org/wiki/Avondmaaltijd         | Avondmaaltijd        | avondmaaltijd       | 0.36666666666666664  | 1     | http://dbpedia.org/resource/Supper    |
  | http://nl.wikipedia.org/wiki/Tic                   | Tic                  | tic                 | 1.0416666666666667   | 1     | http://dbpedia.org/resource/Tic       |
  | http://nl.wikipedia.org/wiki/Platvoeten            | Platvoeten           | platvoeten          | 0.47058823529411764  | 1     | http://dbpedia.org/resource/Flat_feet |
  | http://nl.wikipedia.org/wiki/Groente               | Groente              | groente             | 0.41304347826086957  | 2     | http://dbpedia.org/resource/Vegetable |
  | http://nl.wikipedia.org/wiki/Sojaboon              | Sojaboon             | sojabonen           | 0.6206896551724138   | 1     | http://dbpedia.org/resource/Soybean   |
  | http://nl.wikipedia.org/wiki/Vegetarisme           | Vegetarisme          | vegetariër          | 2.6956521739130435   | 6     | http://dbpedia.org/resource/Vegetarianism |
  | http://nl.wikipedia.org/wiki/Pinot%20noir          | Pinot noir           | pinot noir          | 0.8827586206896552   | 1     | http://dbpedia.org/resource/Pinot_noir |
  | http://nl.wikipedia.org/wiki/Knoflooksaus          | Knoflooksaus         | knoflooksaus        | 0.35294117647058826  | 1     |                                       |
  | http://nl.wikipedia.org/wiki/Consumptiemaatschappij | Consumptiemaatschappij | consumptiemaatschappij | 0.5121951219512195   | 1     |                                       |
                                                                                                 | 0.8421052631578947   | 1     | http://dbpedia.org/resource/Zucchini  |
                                                                                                 | 0.8275862068965517   | 3     | http://dbpedia.org/resource/Tofu      |
                                                                                                 | 0.5752212389380531   | 1     | http://dbpedia.org/resource/Broth     |
                                                                                                 | 0.5752212389380531   | 2     | http://dbpedia.org/resource/Broth     |
                                                                                                 | 0.8666666666666667   | 1     | http://dbpedia.org/resource/Eggplant  |
                                                                                                 | 0.8787878787878788   | 2     |                                       |
```

entity	tag	count
TOM KELLERHUIS	PER	1
FLOOR OVERGOOR	PER	1
HANNY ROSKAMP	PER	2
YANNICK SLAGTER	PER	1

When I send it through the enrichment pipeline language|tokenizer|postagger|ner|stylometry|author|semanticizer|sentiment
Then I should have the following enrichments in the result

enrichment	value
language	nl
author	TOM KELLERHUIS

And I should have the following stylometry enrichments in the result

enrichment	value
number_of_paragraphs	12
number_of_words	432
article_length	2542
vocab_richness	40.5376820773
word_variation	0.585648148148
avg_word_length	4.98611111111
word_size_short	0.553240740741
word_size_medium	0.391203703704
word_size_long	0.0555555555556
hapax_legomena	0.4375
hapax_dislegomena	0.0763888888889
number_of_images	1

And I should have the following semantic links in the result

url	title	text	senseProbability	count	norm
http://nl.wikipedia.org/wiki/Avondmaaltijd	Avondmaaltijd	avondmaaltijd	0.36666666666666664	1	http://dbpedia.org/resource/Supper
http://nl.wikipedia.org/wiki/Tic	Tic	tic	1.0416666666666667	1	http://dbpedia.org/resource/Tic
http://nl.wikipedia.org/wiki/Platvoeten	Platvoeten	platvoeten	0.47058823529411764	1	http://dbpedia.org/resource/Flat_feet
http://nl.wikipedia.org/wiki/Groente	Groente	groente	0.41304347826086957	2	http://dbpedia.org/resource/Vegetable
http://nl.wikipedia.org/wiki/Sojaboon	Sojaboon	sojabonen	0.6206896551724138	1	http://dbpedia.org/resource/Soybean
http://nl.wikipedia.org/wiki/Vegetarisme	Vegetarisme	vegetariër	2.6956521739130435	6	http://dbpedia.org/resource/Vegetarianism
http://nl.wikipedia.org/wiki/Pinot%20noir	Pinot noir	pinot noir	0.8827586206896552	1	http://dbpedia.org/resource/Pinot_noir
http://nl.wikipedia.org/wiki/Knoflooksaus	Knoflooksaus	knoflooksaus	0.35294117647058826	1	
http://nl.wikipedia.org/wiki/Consumptiemaatschappij	Consumptiemaatschappij	consumptiemaatschappij	0.5121951219512195	1	
http://nl.wikipedia.org/wiki/Courgette	Courgette	courgette	0.8421052631578947	1	http://dbpedia.org/resource/Zucchini
http://nl.wikipedia.org/wiki/Tofoe	Tofoe	tofoe	0.8275862068965517	3	http://dbpedia.org/resource/Tofu
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	- bouillon	0.5752212389380531	1	http://dbpedia.org/resource/Broth
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	bouillon	0.5752212389380531	2	http://dbpedia.org/resource/Broth
http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29	Aubergine (plant)	aubergine	0.8666666666666667	1	http://dbpedia.org/resource/Eggplant
http://nl.wikipedia.org/wiki/Kensoort	Kensoort	ken	0.8787878787878788	2	

And I should have the following named entities in the result

entity	tag	count
TOM KELLERHUIS	PER	1
FLOOR OVERGOOR	PER	1
HANNY ROSKAMP	PER	2
YANNICK SLAGTER	PER	1
WIJNFONDUE	LOC	1
SIMPELER	PER	1

User enrichment pipeline

```
Feature: Combine enriched items into a user profile
  Scenario combine two enriched items
    Given "item1" with these enrichments
      | enrichment           | value           |
      | language             | nl              |
      | author               | TOM KELLERHUIS  |
    And these stylometry enrichments
      | enrichment           | value           |
      | number_of_paragraphs | 12              |
      | number_of_words      | 432             |
      | article_length       | 2542            |
      | vocab_richness       | 40.5376820773   |
      | word_variation       | 0.585648148148  |
      | avg_word_length      | 4.98611111111   |
    And these named entities
      | entity           | tag | count |
      | YANNICK SLAGTER  | PER | 1     |
      | WIJNFONDUE       | LOC | 1     |
      | SIMPELER         | PER | 1     |

    Given "item2" with these enrichments
      | enrichment           | value           |
      | language             | de              |
      | author               | PETER LANDERS   |
    And these stylometry enrichments
      | enrichment           | value               |
      | number_of_paragraphs | 7                   |
      | number_of_words      | 600                 |
      | article_length       | 4560                |
      | vocab_richness       | 61.62260711030082   |
      | word_variation       | 0.6265060240963856  |
      | avg_word_length      | 5.578313253012048   |
    And these named entities
      | entity                          | tag | count |
      | ANDRÉ SCHÜRRLE                  | PER | 3     |
      | VFL WOLFSBURG                   | ORG | 4     |
      | KAA                             | ORG | 5     |
```

```
Feature: Combine enriched items into a user profile
  Scenario combine two enriched items
    Given "item1" with these enrichments
      | enrichment           | value           |
      | language             | nl              |
      | author               | TOM KELLERHUIS  |
    And these stylometry enrichments
      | enrichment           | value           |
      | number_of_paragraphs | 12              |
      | number_of_words      | 432             |
      | article_length       | 2542            |
      | vocab_richness       | 40.5376820773   |
      | word_variation       | 0.585648148148  |
      | avg_word_length      | 4.98611111111   |
    And these named entities
      | entity           | tag | count |
      | YANNICK SLAGTER  | PER | 1     |
      | WIJNFONDUE       | LOC | 1     |
      | SIMPELER         | PER | 1     |

    Given "item2" with these enrichments
      | enrichment           | value           |
      | language             | de              |
      | author               | PETER LANDERS   |
    And these stylometry enrichments
      | enrichment           | value               |
      | number_of_paragraphs | 7                   |
      | number_of_words      | 600                 |
      | article_length       | 4560                |
      | vocab_richness       | 61.62260711030082   |
      | word_variation       | 0.6265060240963856  |
      | avg_word_length      | 5.578313253012048   |
    And these named entities
      | entity                          | tag | count |
      | ANDRÉ SCHÜRRLE                  | PER | 3     |
      | VFL WOLFSBURG                   | ORG | 4     |
      | KAA                             | ORG | 5     |
```

```
    When I run the UserEnrichmenTask task and it's dependencies
    Then the user "anne" should have the following profile:
      """json
      {
        "count": 2,
        "enrichments": {
          "language": [
            {
              "count": 1,
              "name": "nl"
            },
            {
              "count": 1,
              "name": "en"
            }
          ],
          "author": [
            {
              "count": 1,
              "name": "TOM KELLERHUIS"
            },
            {
              "count": 1,
              "name": "PETER LANDERS"
            }
          ]
          "stylometry": [
            {
              "count": 2,
              "mean": 2177.5,
              "stdev": 515.4808434849931,
              "name": "article_length"
            },
            {
              "count": 2,
              "mean": 5.165410504361017,
              "stdev": 0.2535676336592832,
              "name": "avg_word_length"
            }
          ]
        }
        "last_100_purchased_items": ["item1", "item2"],
        "tracking_uid": "anne"
      }
      """
```

```
Feature: Combine enriched items into a user profile
  Scenario combine two enriched items
    Given "item1" with these enrichments
      | enrichment           | value           |
      | language             | nl              |
      | author               | TOM KELLERHUIS  |
    And these stylometry enrichments
      | enrichment           | value           |
      | number_of_paragraphs | 12              |
      | number_of_words      | 432             |
      | article_length       | 2542            |
      | vocab_richness       | 40.5376820773   |
      | word_variation       | 0.585648148148  |
      | avg_word_length      | 4.98611111111   |
    And these named entities
      | entity           | tag | count |
      | YANNICK SLAGTER  | PER | 1     |
      | WIJNFONDUE       | LOC | 1     |
      | SIMPELER         | PER | 1     |

    Given "item2" with these enrichments
      | enrichment           | value           |
      | language             | de              |
      | author               | PETER LANDERS   |
    And these stylometry enrichments
      | enrichment           | value               |
      | number_of_paragraphs | 7                   |
      | number_of_words      | 600                 |
      | article_length       | 4560                |
      | vocab_richness       | 61.62260711030082   |
      | word_variation       | 0.6265060240963856  |
      | avg_word_length      | 5.578313253012048   |
    And these named entities
      | entity                          | tag | count |
      | ANDRÉ SCHÜRRLE                  | PER | 3     |
      | VFL WOLFSBURG                   | ORG | 4     |
      | KAA                             | ORG | 5     |
```

```
  When I run the UserEnrichmenTask task and it's dependencies
  Then the user "anne" should have the following profile:
"""
{
  "language": [
    {
      "count": 1,
      "name": "nl"
    },
    {
      "count": 1,
      "name": "en"
    }
  ],
  "stylometry": [
    {
      "count": 2,
      "mean": 2177.5,
      "stdev": 515.4808434849931,
      "name": "article_length"
    },
    {
      "count": 2,
      "mean": 5.165410504361017,
      "stdev": 0.2535676336592832,
      "name": "avg_word_length"
    }
  ]
}
  "last_100_purchased_items": ["item1", "item2"],
  "tracking_uid": "anne"
}
"""
```

```
Feature: Combine enriched items into a user profile
  Scenario combine two enriched items
    Given "item1" with these enrichments
      | enrichment           | value           |
      | language             | nl              |
      | author               | TOM KELLERHUIS  |
    And these stylometry enrichments
      | enrichment           | value           |
      | number_of_paragraphs | 12              |
      | number_of_words      | 432             |
      | article_length       | 2542            |
      | vocab_richness       | 40.5376820773   |
      | word_variation       | 0.585648148148  |
      | avg_word_length      | 4.98611111111   |
    And these named entities
      | entity           | tag | count |
      | YANNICK SLAGTER  | PER | 1     |
      | WIJNFONDUE       | LOC | 1     |
      | SIMPELER         | PER | 1     |

    Given "item2" with these enrichments
      | enrichment           | value           |
      | language             | de              |
      | author               | PETER LANDERS   |
    And these stylometry enrichments
      | enrichment           | value               |
      | number_of_paragraphs | 7                   |
      | number_of_words      | 600                 |
      | article_length       | 4560                |
      | vocab_richness       | 61.62260711030082   |
      | word_variation       | 0.6265060240963856  |
      | avg_word_length      | 5.578313253012048   |
    And these named entities
      | entity                          | tag | count |
      | ANDRÉ SCHÜRRLE                  | PER | 3     |
      | VFL WOLFSBURG                   | ORG | 4     |
      | KAA                             | ORG | 5     |
```

```
  When I run the UserEnrichmenTask task and it's dependencies
  Then the user "anne" should have the following profile:
"""
{
  "language": [
    {
      "count": 1,
      "name": "nl"
    },
    {
      "count": 1,
      "name": "de"
    }
  ],
  "author": [
    {
      "count": 1,
      "name": "TOM KELLERHUIS"
    },
    {
      "count": 1,
      "name": "PETER LANDERS"
    }
  ],
  "named_entities": [
    {
      "count": 3,
      "name": "ANDRÉ SCHÜRRLE",
      "tag": "PER"
    },
    {
      "count": 4,
      "name": "VFL WOLFSBURG",
      "tag": "ORG"
    },
    {
      "count": 5,
      "name": "KAA",
      "tag": "ORG"
    }
  ],
  "stylometry": [
    {
      "count": 2,
      "mean": 2177.5,
      "stdev": 515.4808434849931,
      "name": "article_length"
    },
    {
      "count": 2,
      "mean": 5.165410504361017,
      "stdev": 0.2535676336592832,
      "name": "avg_word_length"
    }
  ],
  "last_100_purchased_items": ["item1", "item2"],
  "tracking_uid": "anne"
}
"""
```

```
Feature: Combine enriched items into a user profile
  Scenario combine two enriched items
    Given "item1" with these enrichments
      | enrichment           | value           |
      | language             | nl              |
      | author               | TOM KELLERHUIS  |
    And these stylometry enrichments
      | enrichment           | value           |
      | number_of_paragraphs | 12              |
      | number_of_words      | 432             |
      | article_length       | 2542            |
      | vocab_richness       | 40.5376820773   |
      | word_variation       | 0.585648148148  |
      | avg_word_length      | 4.98611111111   |
    And these named entities
      | entity           | tag | count |
      | YANNICK SLAGTER  | PER | 1     |
      | WIJNFONDUE       | LOC | 1     |
      | SIMPELER         | PER | 1     |

    Given "item2" with these enrichments
      | enrichment           | value           |
      | language             | de              |
      | author               | PETER LANDERS   |
    And these stylometry enrichments
      | enrichment           | value               |
      | number_of_paragraphs | 7                   |
      | number_of_words      | 600                 |
      | article_length       | 4560                |
      | vocab_richness       | 61.62260711030082   |
      | word_variation       | 0.6265060240963856  |
      | avg_word_length      | 5.578313253012048   |
    And these named entities
      | entity                          | tag | count |
      | ANDRÉ SCHÜRRLE                  | PER | 3     |
      | VFL WOLFSBURG                   | ORG | 4     |
      | KAA                             | ORG | 5     |
```

```
When I run the UserEnrichmenTask task and it's dependencies
Then the user "anne" should have the following profile:
"""
{
  "language": [
    {
      "author": [
        {
          "stylometry": [
            {
              "count": 2,
              "mean": 2177.5,
              "stdev": 515.4808434849931,
              "name": "article_length"
            },
            {
              "count": 2,
              "mean": 5.165410504361017,
              "stdev": 0.2535676336592832,
              "name": "avg_word_length"
            }
          ]
        }
      ]
    }
  ],
  "last_100_purchased_items": ["item1", "item2"],
  "tracking_uid": "anne"
}
"""
```

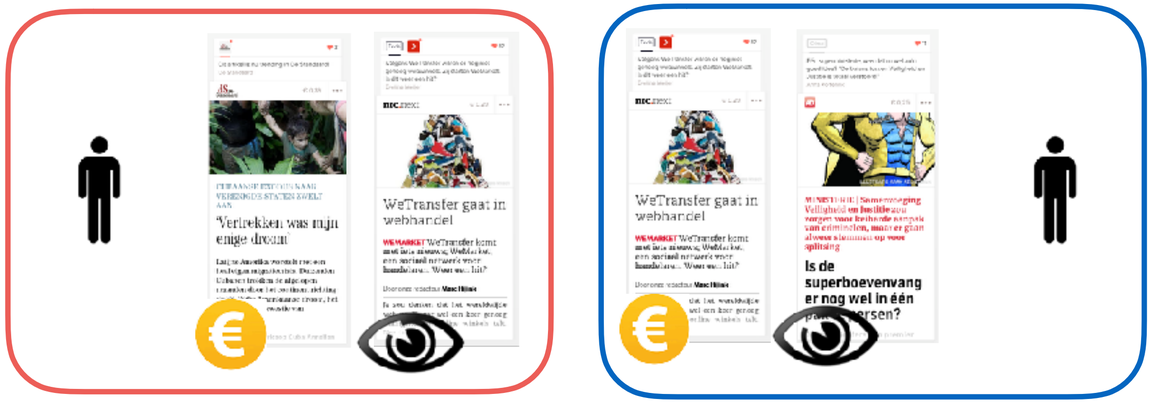
Learning to rank: preference learning

Blendle

Learning to rank: preference learning

Blendle

Learning to rank: preference learning



Blendle

Learning to rank: preference learning

Enrich

Enrich

Enrich

Enrich

Blendle

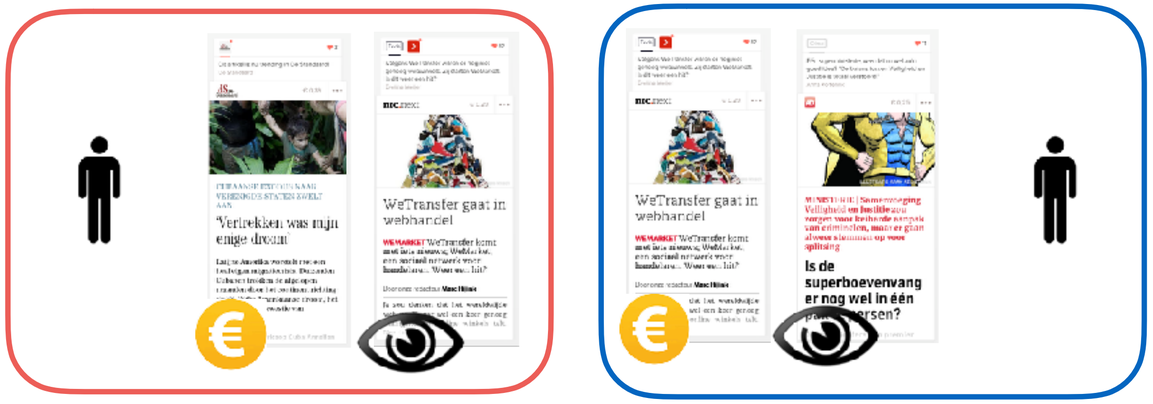
Learning to rank: preference learning

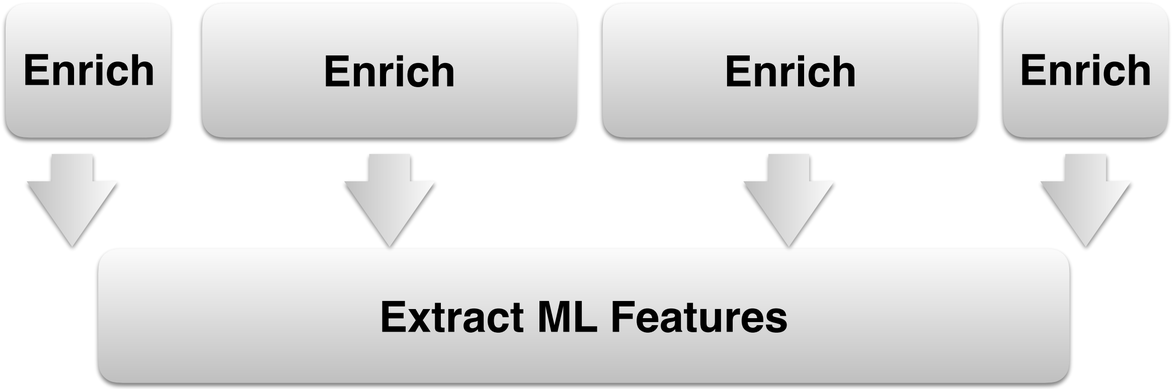
Enrich **Enrich** **Enrich** **Enrich**

Extract ML Features

Blendle

Learning to rank: preference learning

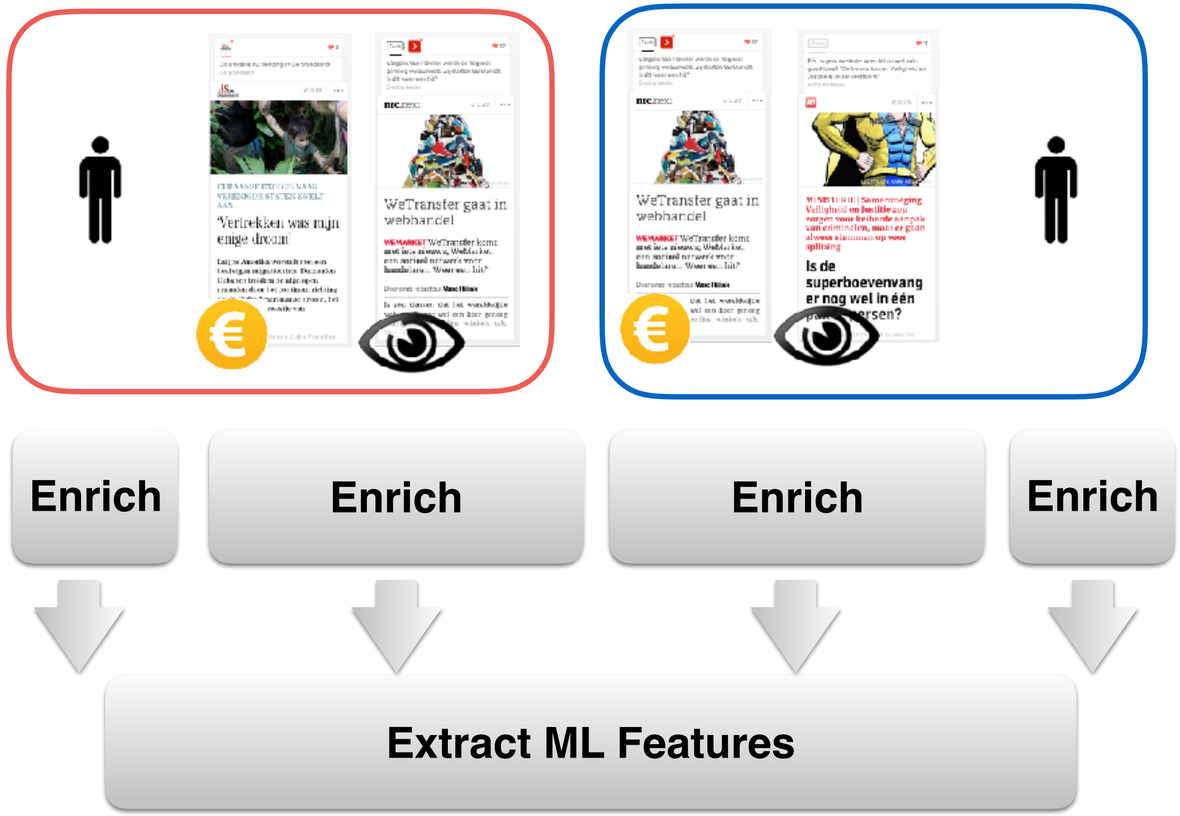





```
When I compute features
Then I should have the following features and values:
  | feature                              | value              |
  | user_nr_purchased_items              | 5                  |
  | user_nr_channels                     | 1                  |
  | user_nr_subscriptions                | 2                  |
  | user_nr_providers                    | 1                  |
  | user_vmoney_almost_spent             | 1                  |
  | user_alerts                          | 1                  |
  | user_new_edition                     | 1                  |
  | user_digest                          | 1                  |
  | user_magazine_digest                 | 1                  |
  | user_image_mean                      | 0.5                |
  | user_likes_it_long                   | 0                  |
  | purchased_channel_intersect          | 1                  |
  | purchased_channel_intersect_size     | 1                  |
  | newsletter_purchased_channel_score   | 0.05               |
  | newsletter_purchased_provider_score  | 0.025              |
  | item_provider_followed_by_user       | 0                  |
  | item_channel_followed_by_user        | 0                  |
  | item_rating_score                    | 0.9                |
  | item_number_of_images                | 1                  |
  | item_length_class                    | 0                  |
  | user_item_language_score             | 0.5                |
  | user_item_author_in_top3             | 1                  |
  | user_item_semlink_in_top10           | 1                  |
  | user_item_semantic_overlap           | 0.09777777777777   |
  | user_item_named_entity_overlap       | 0.04081632653061   |
  | user_item_length_pdf                 | 0.00316108841344   |
```

Blendle

Learning to rank: preference learning




```
When I compute features
Then I should have the following features and values:
  | feature                                | value                |
  | user_nr_purchased_items                | 5                    |
  | user_nr_channels                       | 1                    |
  | user_nr_subscriptions                  | 2                    |
  | user_nr_providers                      | 1                    |
  | user_vmoney_almost_spent               | 1                    |
  | user_alerts                            | 1                    |
  | user_new_edition                       | 1                    |
  | user_digest                            | 1                    |
  | user_magazine_digest                   | 1                    |
  | user_image_mean                        | 0.5                  |
  | user_likes_it_long                     | 0                    |
  | purchased_channel_intersect            | 1                    |
  | purchased_channel_intersect_size       | 1                    |
  | newsletter_purchased_channel_score     | 0.05                 |
  | newsletter_purchased_provider_score    | 0.025                |
  | item_provider_followed_by_user         | 0                    |
  | item_channel_followed_by_user          | 0                    |
  | item_rating_score                      | 0.9                  |
  | item_number_of_images                  | 1                    |
  | item_length_class                      | 0                    |
  | user_item_semlink_in_top10             | 1                    |
  | user_item_semantic_overlap             | 0.09777777777777     |
  | user_item_named_entity_overlap         | 0.04081632653061     |
  | user_item_length_pdf                   | 0.00316108841344     |
```

```
| user_item_author_in_top3                | 1
```

Blendle

Learning to rank: preference learning

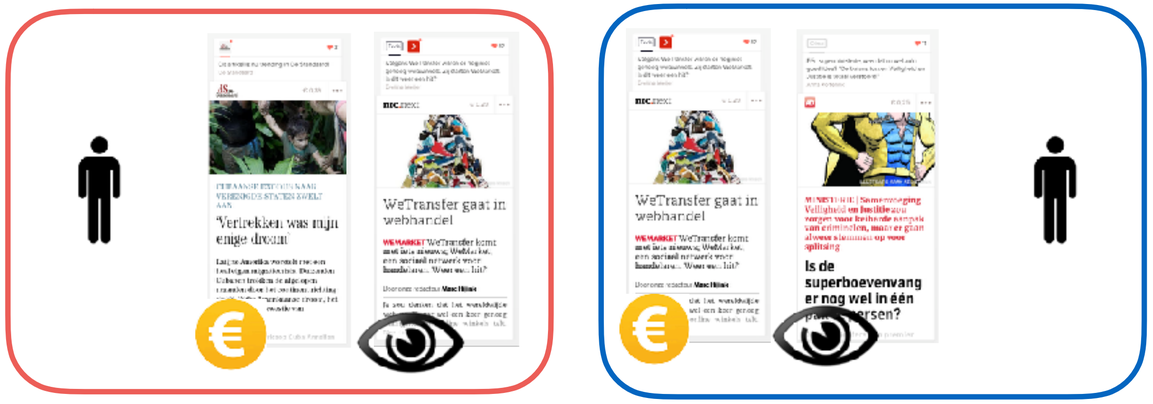
Enrich | Enrich | Enrich | Enrich

Extract ML Features

```
When I compute features
Then I should have the following features and values:
  | feature                              | value              |
  | user_nr_purchased_items              | 5                  |
  | user_nr_channels                     | 1                  |
  | user_nr_subscriptions                | 2                  |
  | user_nr_providers                    | 1                  |
  | user_vmoney_almost_spent             | 1                  |
  | user_alerts                          | 1                  |
  | user_new_edition                     | 1                  |
  | user_digest                          | 1                  |
  | user_magazine_digest                 | 1                  |
  | user_image_mean                      | 0.5                |
  | user_likes_it_long                   | 0                  |
  | purchased_channel_intersect          | 1                  |
  | purchased_channel_intersect_size     | 1                  |
  | newsletter_purchased_channel_score   | 0.05               |
  | newsletter_purchased_provider_score  | 0.025              |
  | item_provider_followed_by_user       | 0                  |
  | item_channel_followed_by_user        | 0                  |
  | user_item_language_score             | 0.5                |
  | item_length_class                    | 0                  |
  | user_item_language_score             | 0.5                |
  | user_item_author_in_top3             | 1                  |
  | user_item_semlink_in_top10           | 1                  |
  | user_item_semantic_overlap           | 0.09777777777777   |
  | user_item_named_entity_overlap       | 0.04081632653061   |
  | user_item_length_pdf                 | 0.00316108841344   |
```

Blendle

Learning to rank: preference learning



Enrich | Enrich | Enrich

Extract ML Features

```
When I compute features
Then I should have the following features and values:
  | feature                              | value              |
  | user_nr_purchased_items              | 5                  |
  | user_nr_channels                     | 1                  |
  | user_nr_subscriptions                | 2                  |
  | user_nr_providers                    | 1                  |
  | user_vmoney_almost_spent             | 1                  |
  | user_alerts                          | 1                  |
  | user_new_edition                     | 1                  |
  | user_digest                          | 1                  |
  | user_magazine_digest                 | 1                  |
  | user_likes_it_long                   | 0                  |
  | purchased_channel_intersect          | 1                  |
  | purchased_channel_intersect_size     | 1                  |
  | newsletter_purchased_channel_score   | 0.05               |
  | newsletter_purchased_provider_score  | 0.025              |
  | item_provider_followed_by_user       | 0                  |
  | item_channel_followed_by_user        | 0                  |
  | item_rating_score                    | 0.9                |
  | item_number_of_images                | 1                  |
  | item_length_class                    | 0                  |
  | user_item_language_score             | 0.5                |
  | user_item_author_in_top3             | 1                  |
  | user_item_semlink_in_top10           | 1                  |
  | user_item_semantic_overlap           | 0.09777777777777   |
  | user_item_named_entity_overlap       | 0.04081632653061   |
  | user_item_length_pdf                 | 0.00316108841344   |
```

Blendle

Learning to rank: preference learning

Enrich **Enrich** **Enrich** **Enrich**

Extract ML Features

Blendle

Learning to rank: preference learning

Enrich **Enrich** **Enrich** **Enrich**

Extract ML Features

Pairwise transform

Blendle

Learning to rank: preference learning

Enrich **Enrich** **Enrich** **Enrich**

Extract ML Features

Pairwise transform

Learner to predict

SVM (Spark MLib) with linear kernel

Blendle

Learning to rank: preference learning

Enrich **Enrich** **Enrich** **Enrich**

Extract ML Features

Pairwise transform

Learner to predict

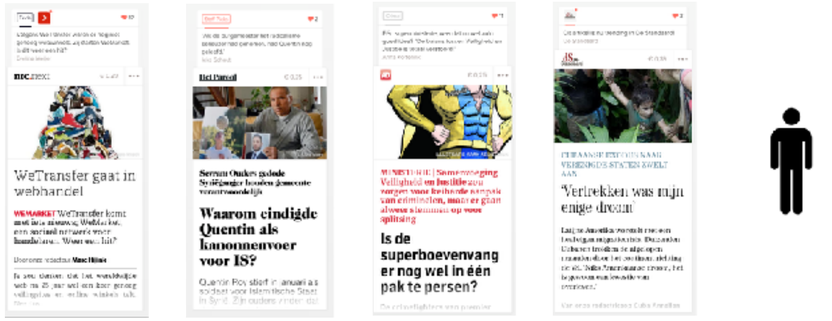
SVM (Spark MLib) with linear kernel

Model

Feature weights

Blendle

Learning to rank: preference learning



Enrich **Enrich** **Enrich** **Enrich**

Extract ML Features

Pairwise transform

Learner to predict €

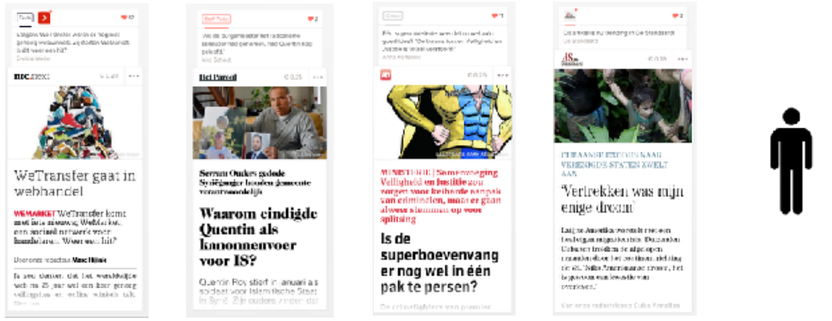
SVM (Spark MLib) with linear kernel

Model

Feature weights

Blendle

Learning to rank: preference learning



Enrich | **Enrich** | **Enrich** | **Enrich** | **Enrich** | **Enrich**

Extract ML Features

Pairwise transform

Learner to predict €

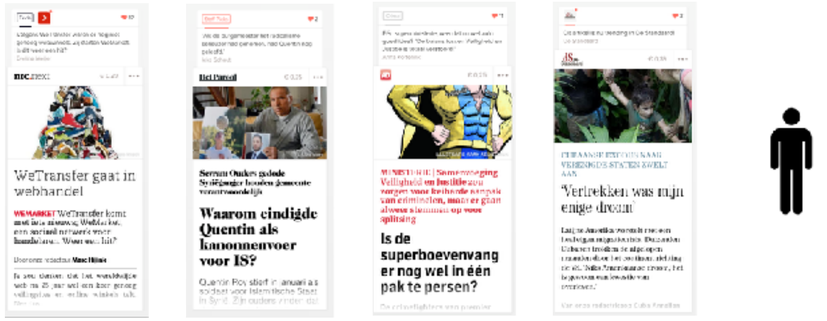
SVM (Spark MLib) with linear kernel

Model

Feature weights

Blendle

Learning to rank: preference learning



Enrich **Enrich** **Enrich** **Enrich**

Extract ML Features

Pairwise transform

Learner to predict €

SVM (Spark MLib) with linear kernel

Model

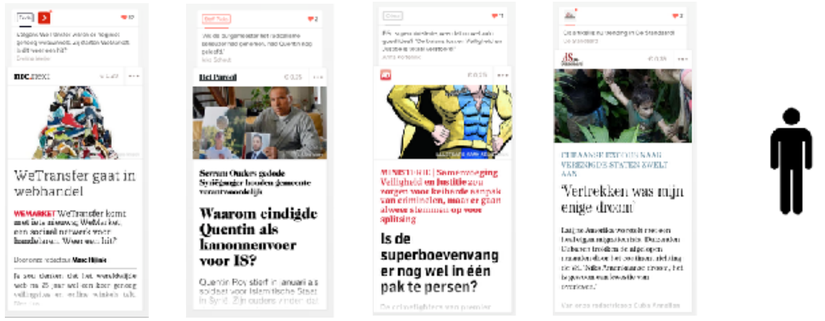
Feature weights

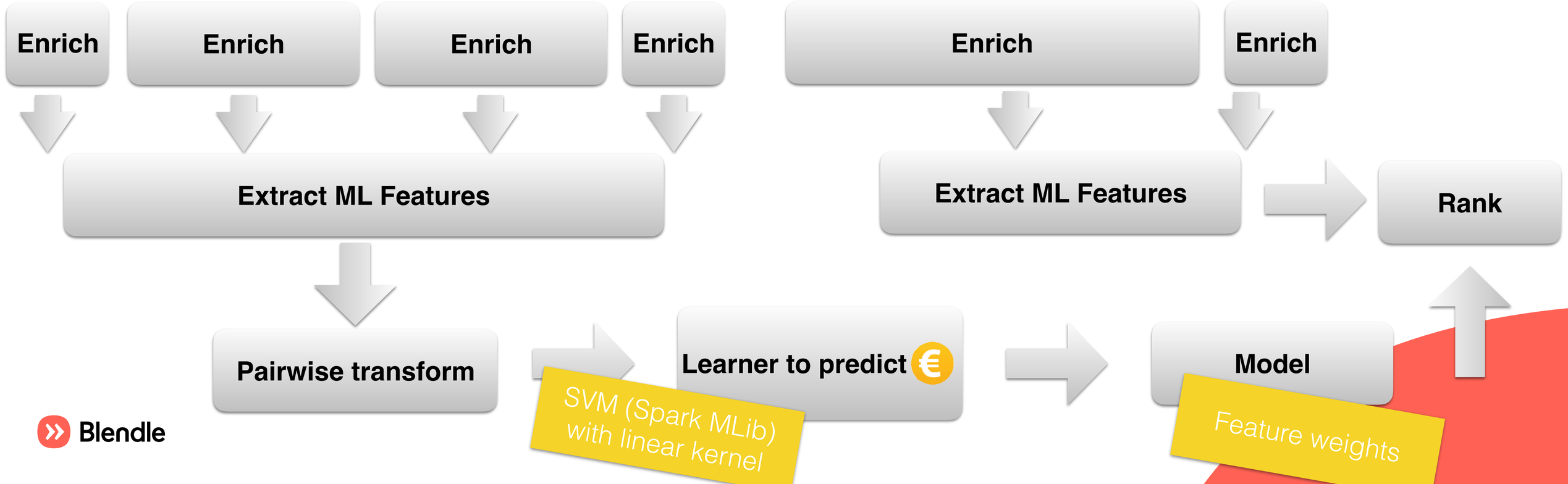
Enrich **Enrich**

Extract ML Features

Blendle

Learning to rank: preference learning





Blendle

Learning to rank: preference learning

Ranking

Enrich

Enrich

Enrich

Enrich

Enrich

Enrich

Extract ML Features

Extract ML Features

Rank

Pairwise transform

Learner to predict

Model

SVM (Spark MLib) with linear kernel

Feature weights

Blendle

Learning to Rank

- What can possibly go wrong?

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
    - users won’t click on things you didn’t show them

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
    - users won’t click on things you didn’t show them
    - user are likely to click on things that appear high

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
    - users won’t click on things you didn’t show them
    - user are likely to click on things that appear high
    - it matters how you present documents

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
    - users won’t click on things you didn’t show them
    - user are likely to click on things that appear high
    - it matters how you present documents
      - snippets, images, colors, font size, grouped with other documents

Learning to R

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

Learning to R

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...

savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

Learning to R

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...

savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30

Feature-specific metric:
- Short-dwell click rate

➔ **improves**! ☺

➔ Good feature!?

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30

Feature-specific metric:
- Short-dwell click rate

➔ **improves**! ☺

➔ Good feature!?

What could have happened?

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

Learning to R

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...

savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30

Feature-specific metric:
- Short-dwell click rate

➔ **improves**! ☺

➔ Good feature!?

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

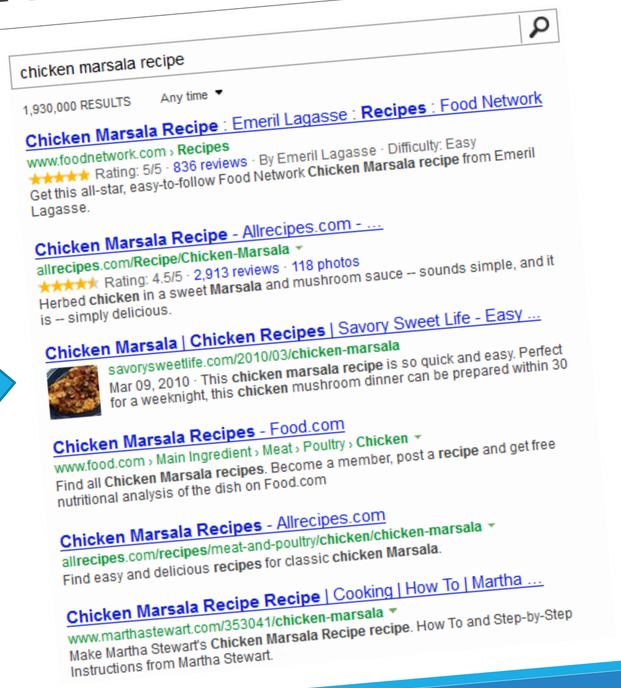
Learning to Rank

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.



Feature-specific metric:
- Short-dwell click rate

➔ **improves**! ☺

➔ Good feature!?

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

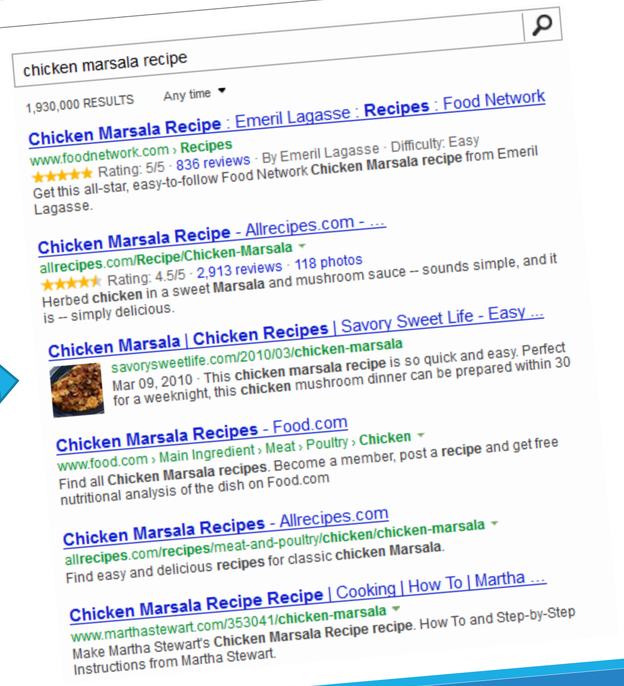
Learning to R

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.



Feature-specific metric:
- Short-dwell click rate

➔ **improves**! ☺

➔ Good feature!?

More generalizable metrics:
- Whole page click-through rate
- #queries per session

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

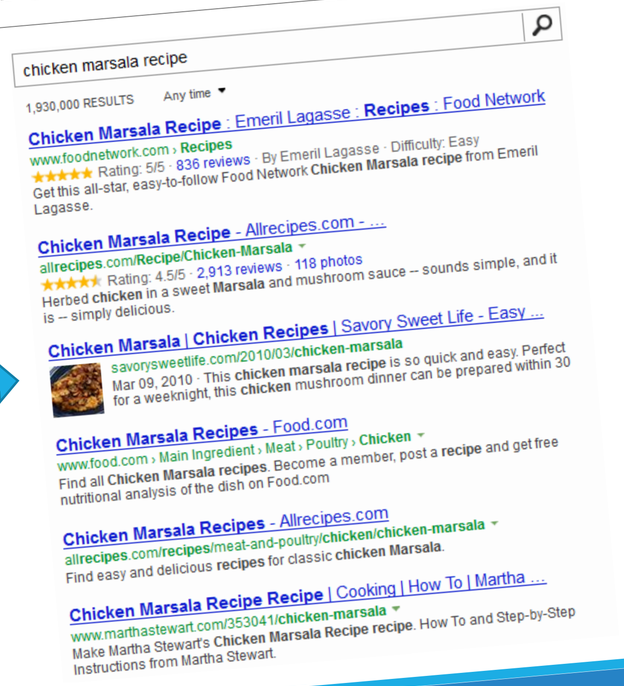
Learning to R

Example: Rich Result Summarization

Idea: Users like to know how dishes look like for recipes.

Let's generate nicer captions by including an image.

Chicken Marsala | Chicken Recipes | Savory Sweet Life - Easy ...
savorysweetlife.com/2010/03/**chicken-marsala**
Mar 09, 2010 · This **chicken marsala recipe** is so quick and easy. Perfect for a weeknight, this **chicken** mushroom dinner can be prepared within 30 minutes.



Challenge: Design feature-specific metrics in a way that they are aligned with overall goodness.

Feature-specific metric:
- Short-dwell click rate

➔ **improves**! ☺

➔ Good feature!?

More generalizable metrics:
- Whole page click-through rate
- #queries per session

Buscher, G. (2013). IR Evaluation: Perspectives From Within a Living Lab.

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
  - Clicks are noisy

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
  - Clicks are noisy
    - They don’t always mean what you hope

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
  - Clicks are noisy
    - They don’t always mean what you hope
    - Even with refunds

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
  - Clicks are noisy
    - They don’t always mean what you hope
    - Even with refunds
  - Absence of clicks is not necessarily negative

Learning to Rank

- What can possibly go wrong?
  - Clicks are biased
  - Clicks are noisy
    - They don’t always mean what you hope
    - Even with refunds
  - Absence of clicks is not necessarily negative
    - Users might learn enough from snippets

Online Learning to Rank

Online Learning to Rank

- Learning with a user in the loop

Online Learning to Rank

- Learning with a user in the loop
- Bias is no longer a problem

Online Learning to Rank

- Learning with a user in the loop
- Bias is no longer a problem
- Noise remains

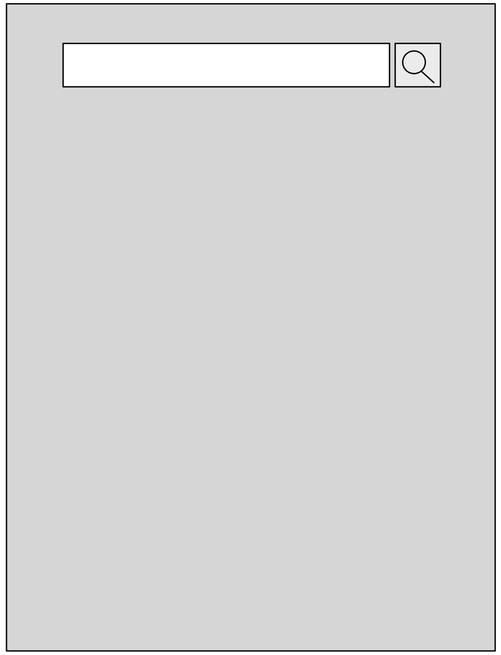
Online Learning to Rank

We are not yet doing this at Blendle

- Learning with a user in the loop
- Bias is no longer a problem
- Noise remains

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

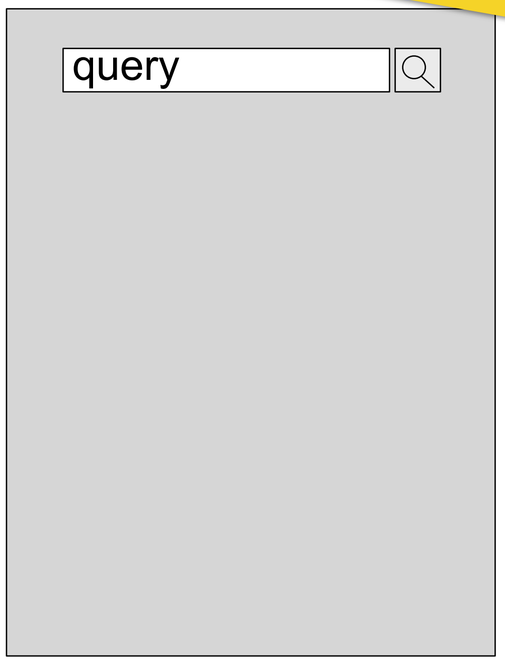


[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

For Blendle the user would be the query

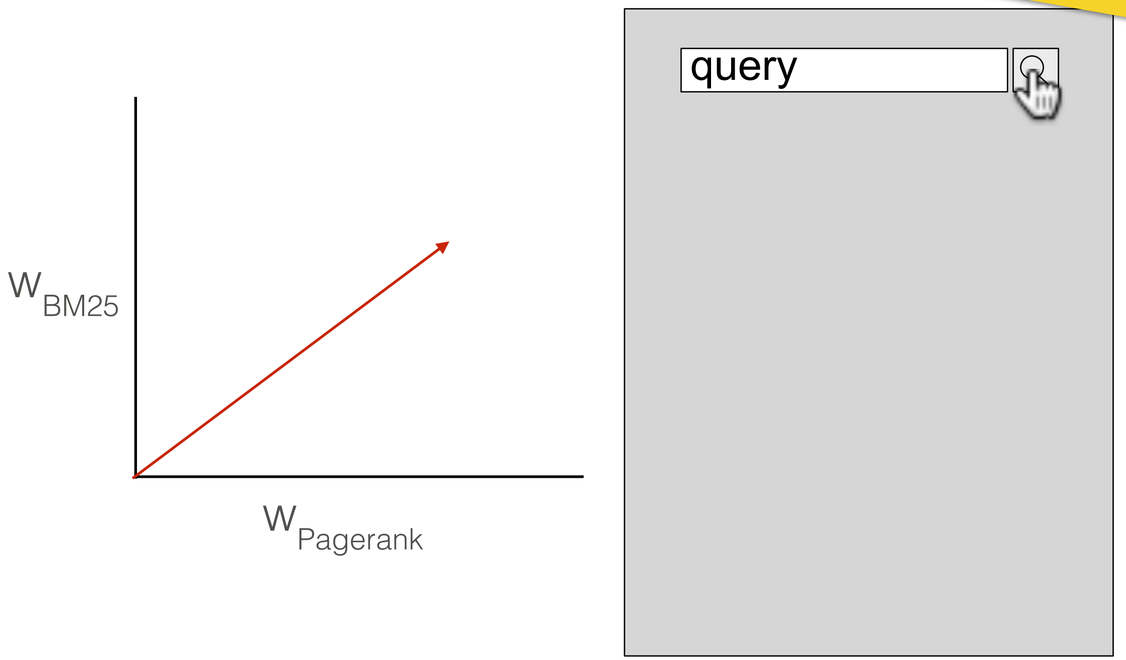


[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

For Blendle the user would be the query

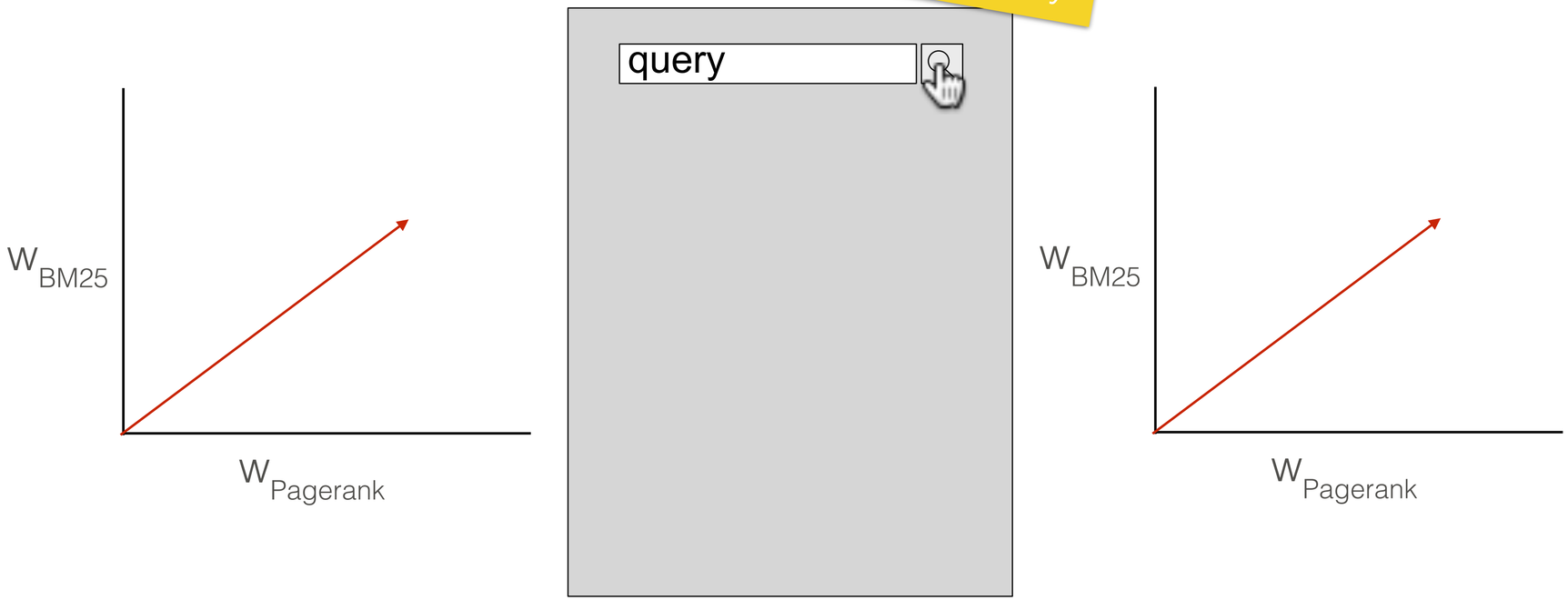


[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

For Blendle the user would be the query

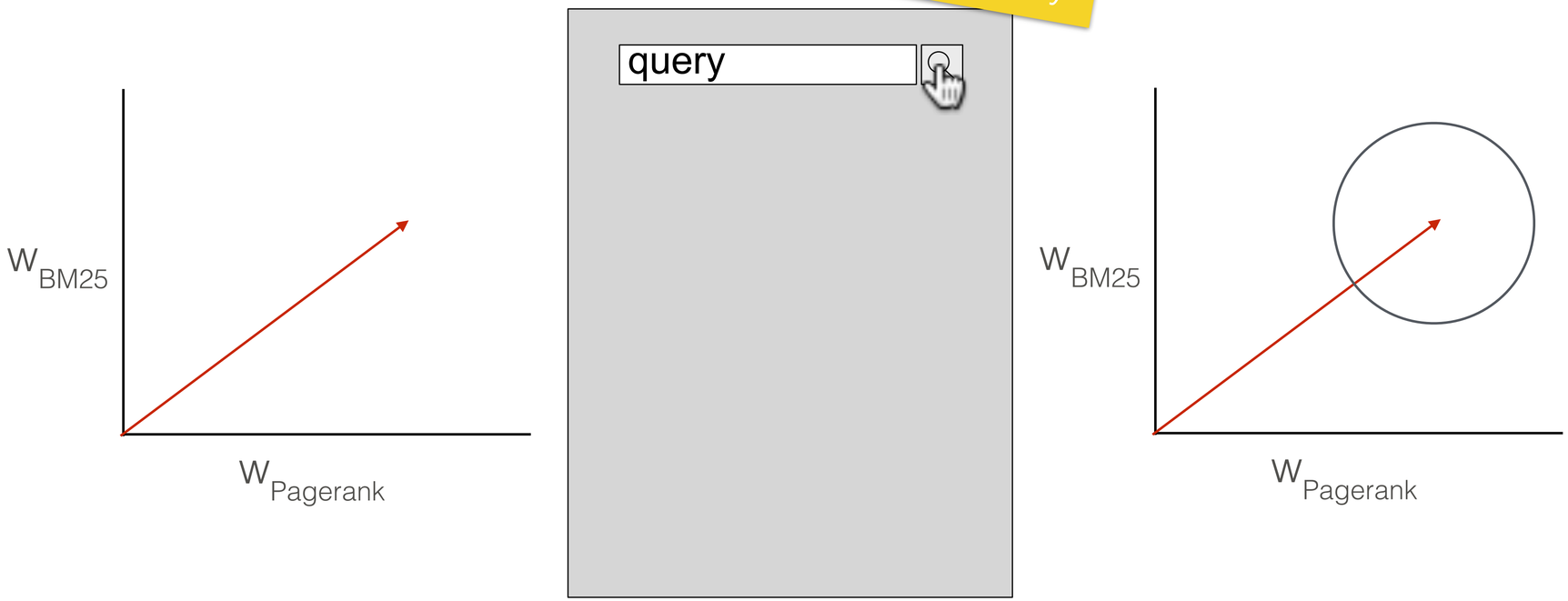


[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

For Blendle the user would be the query

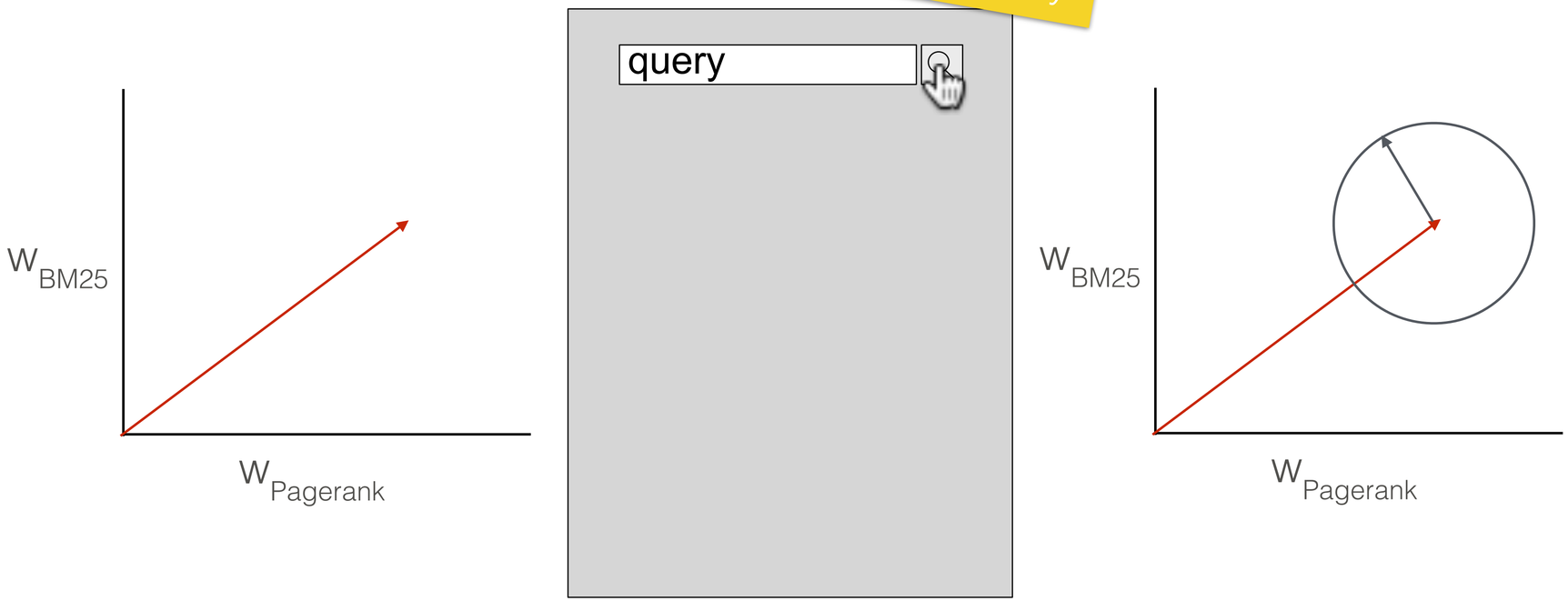


[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

For Blendle the user would be the query

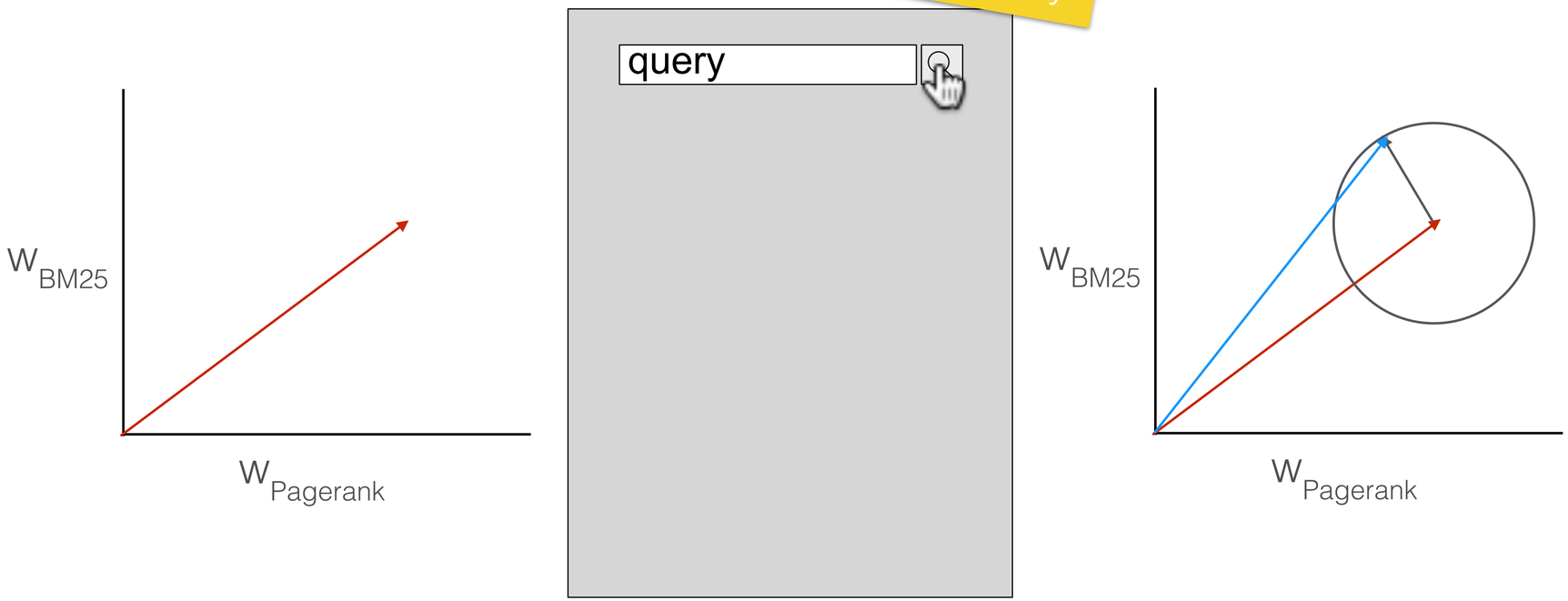


[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

For Blendle the user would be the query



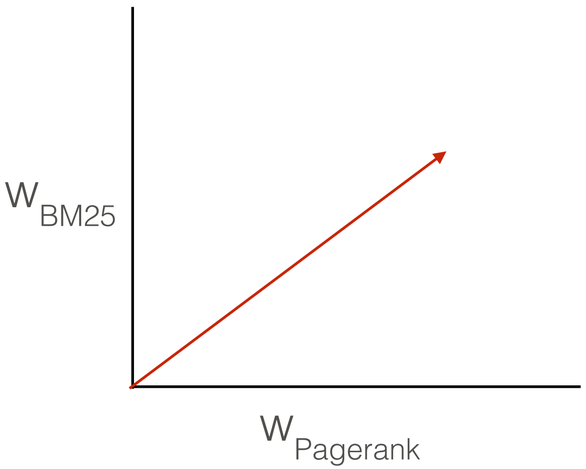
[Yue et al, 2009; Hofmann et al., 2011]

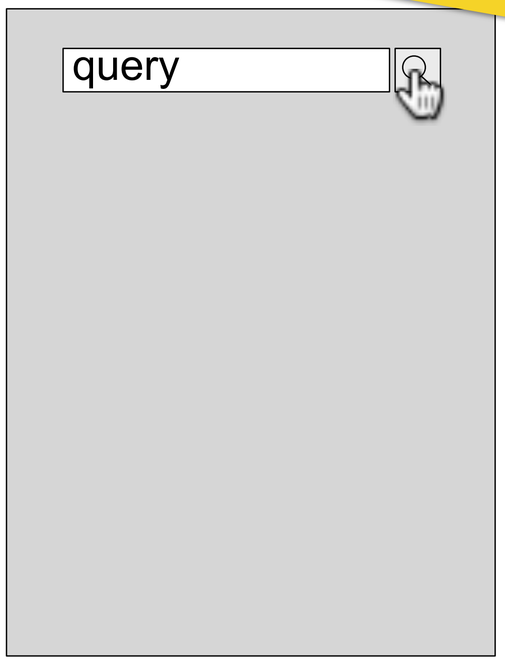
Dueling Bandit Gradient Descent

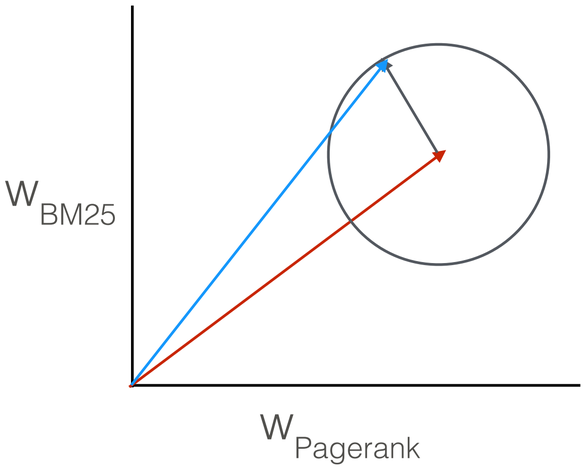
We are not yet doing this at Blendle

For Blendle the user would be the query

Exploitative Ranker







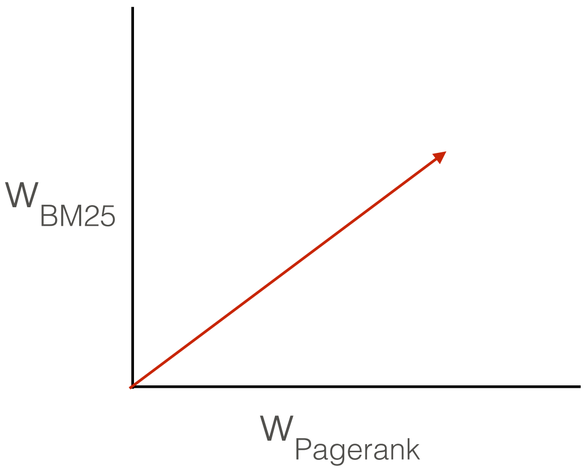
[Yue et al, 2009; Hofmann et al., 2011]

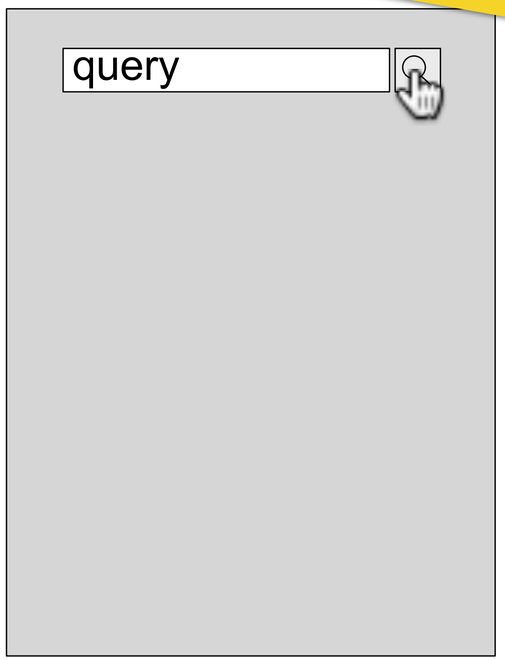
Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

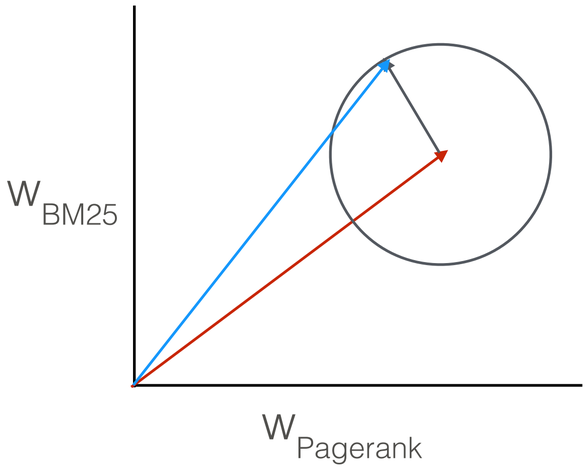
For Blendle the user would be the query

Exploitative Ranker





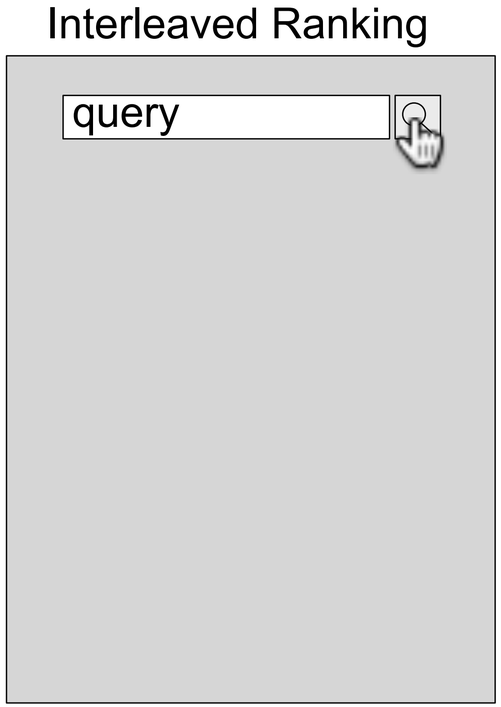
Explorative Ranker



[Yue et al, 2009; Hofmann et al., 2011]

TeamDraft Interleave

We are not yet doing this at Blendle

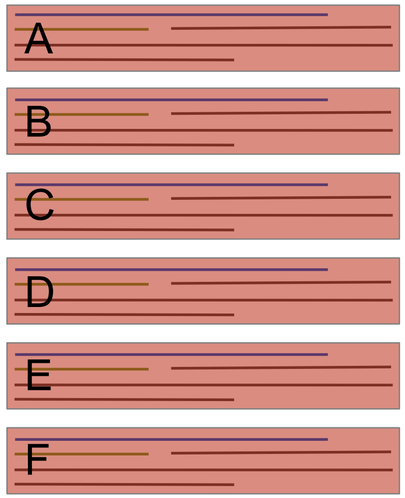


Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

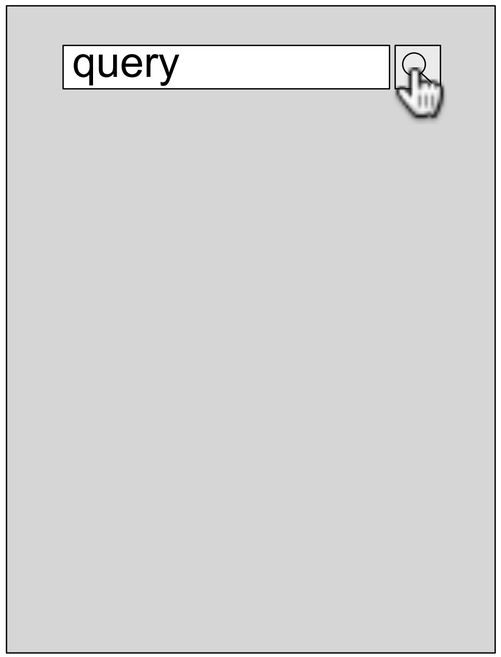
TeamDraft Interleave

We are not yet doing this at Blendle

Exploitative Ranking



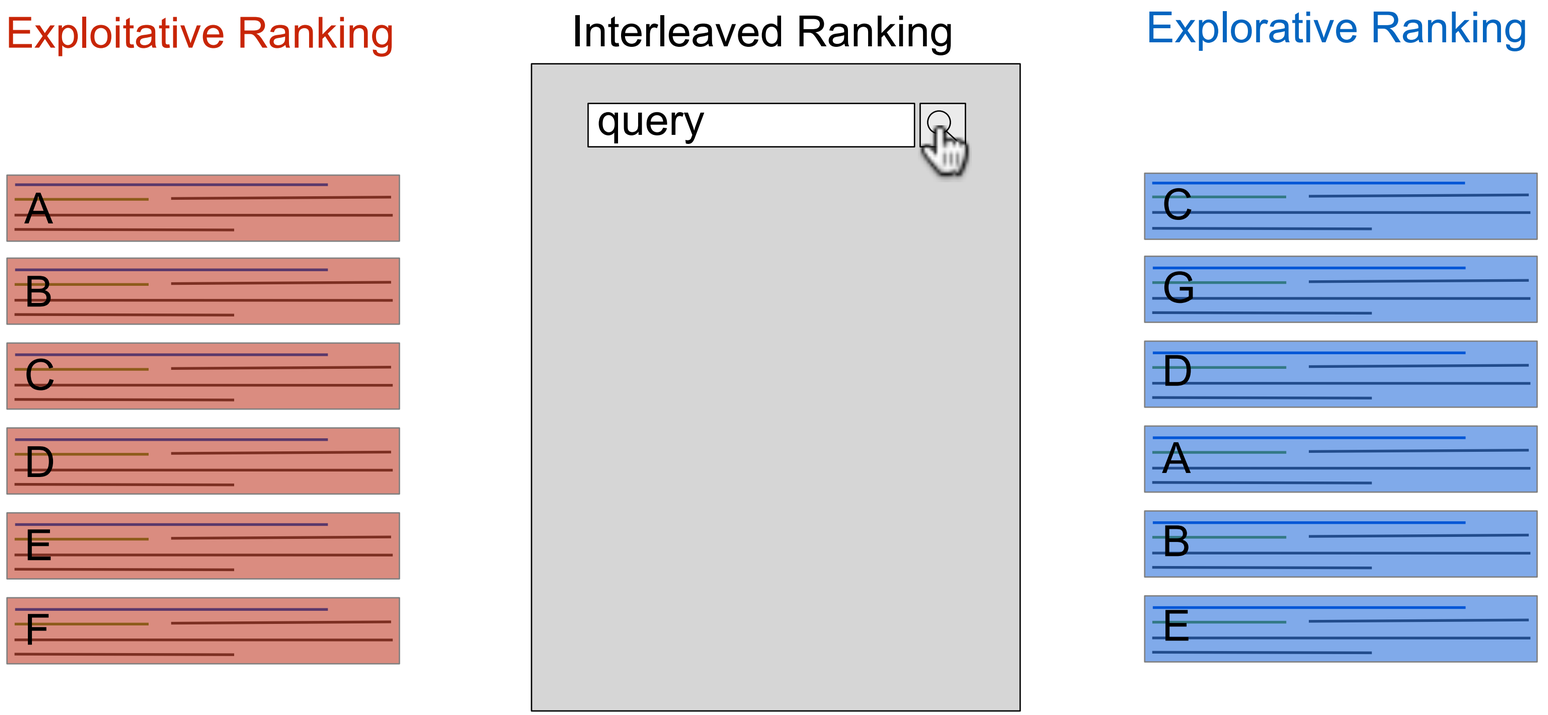
Interleaved Ranking



Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

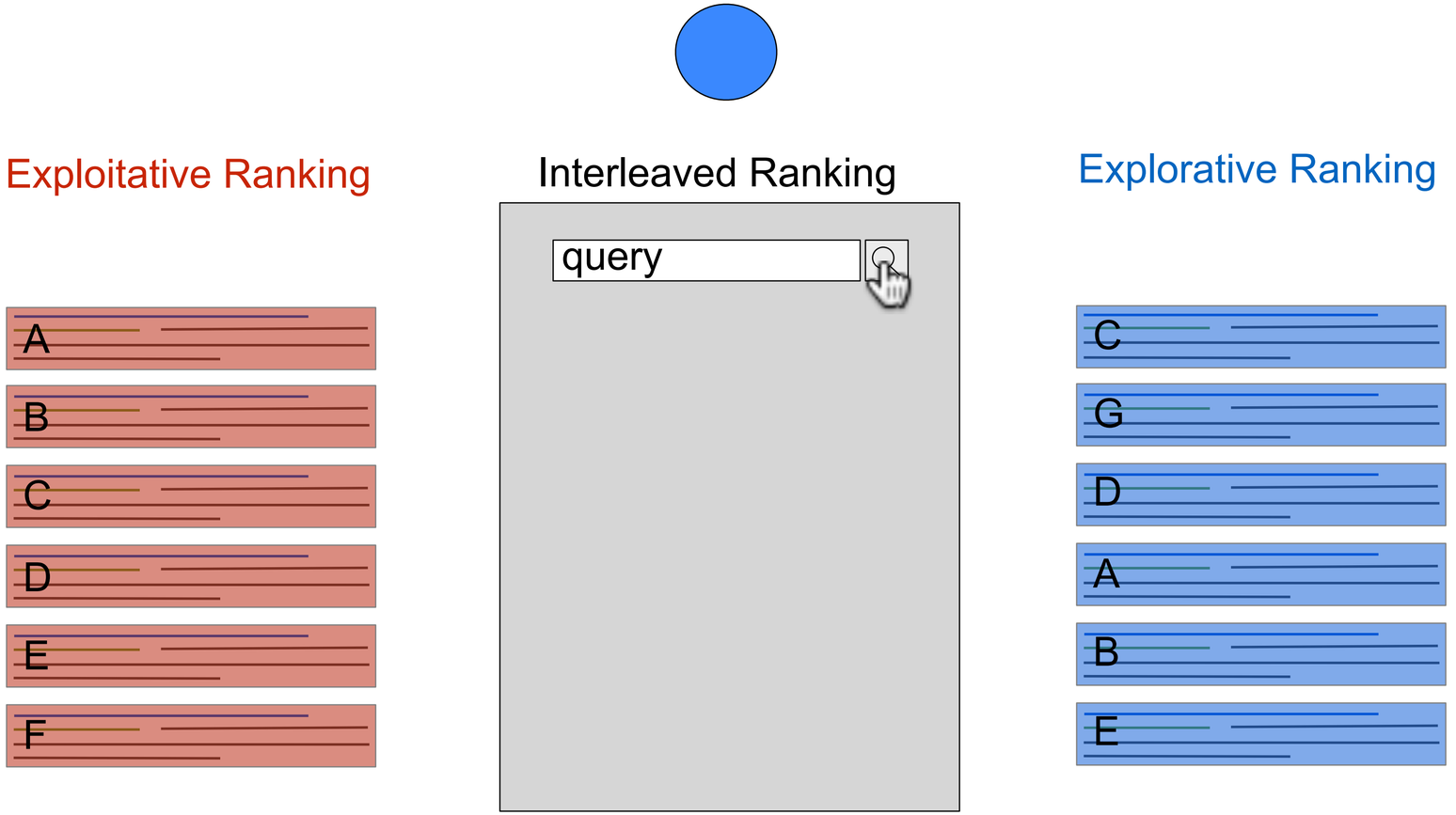
We are not yet doing this at Blendle



Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

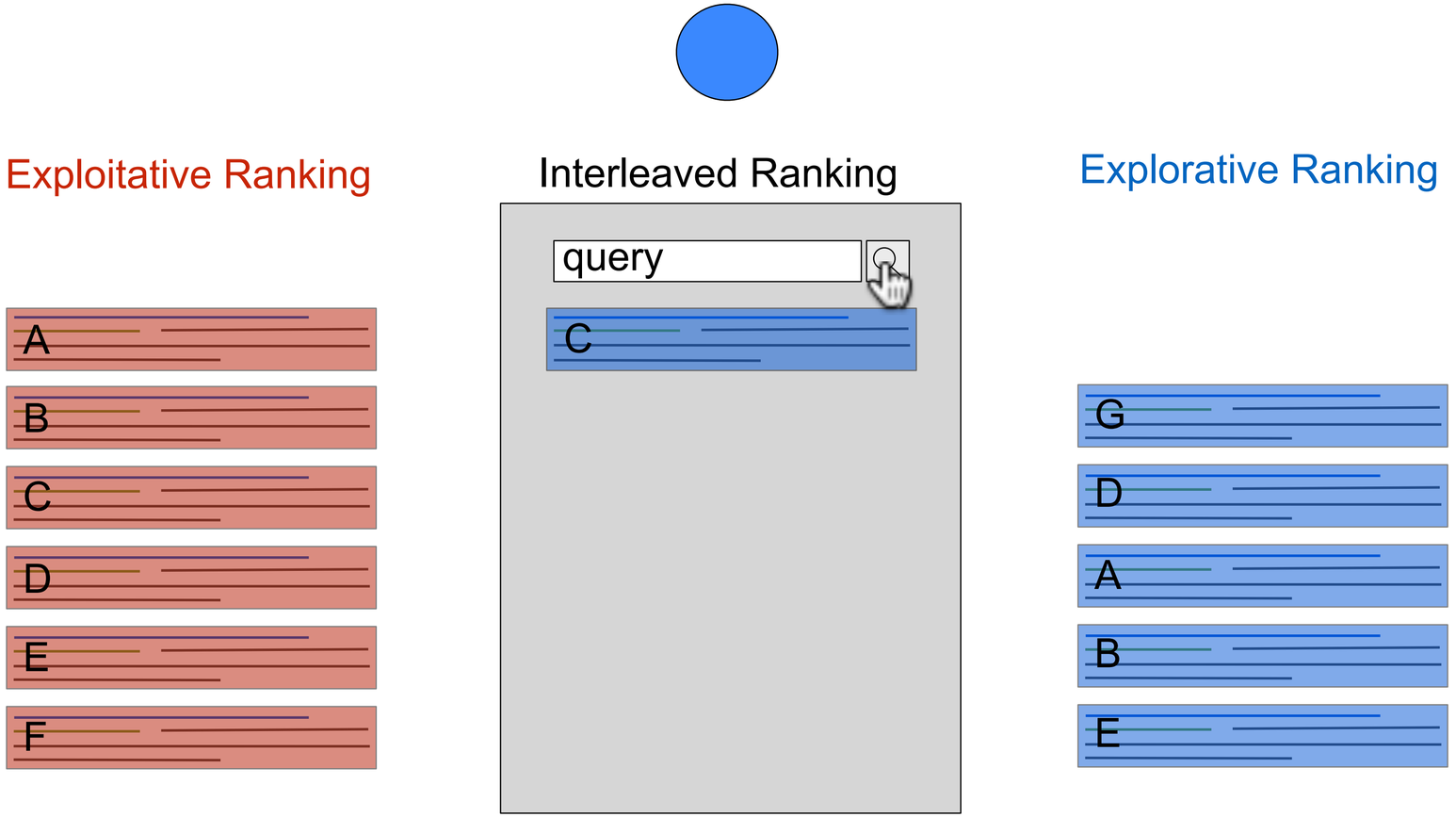
We are not yet doing this at Blendle



Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

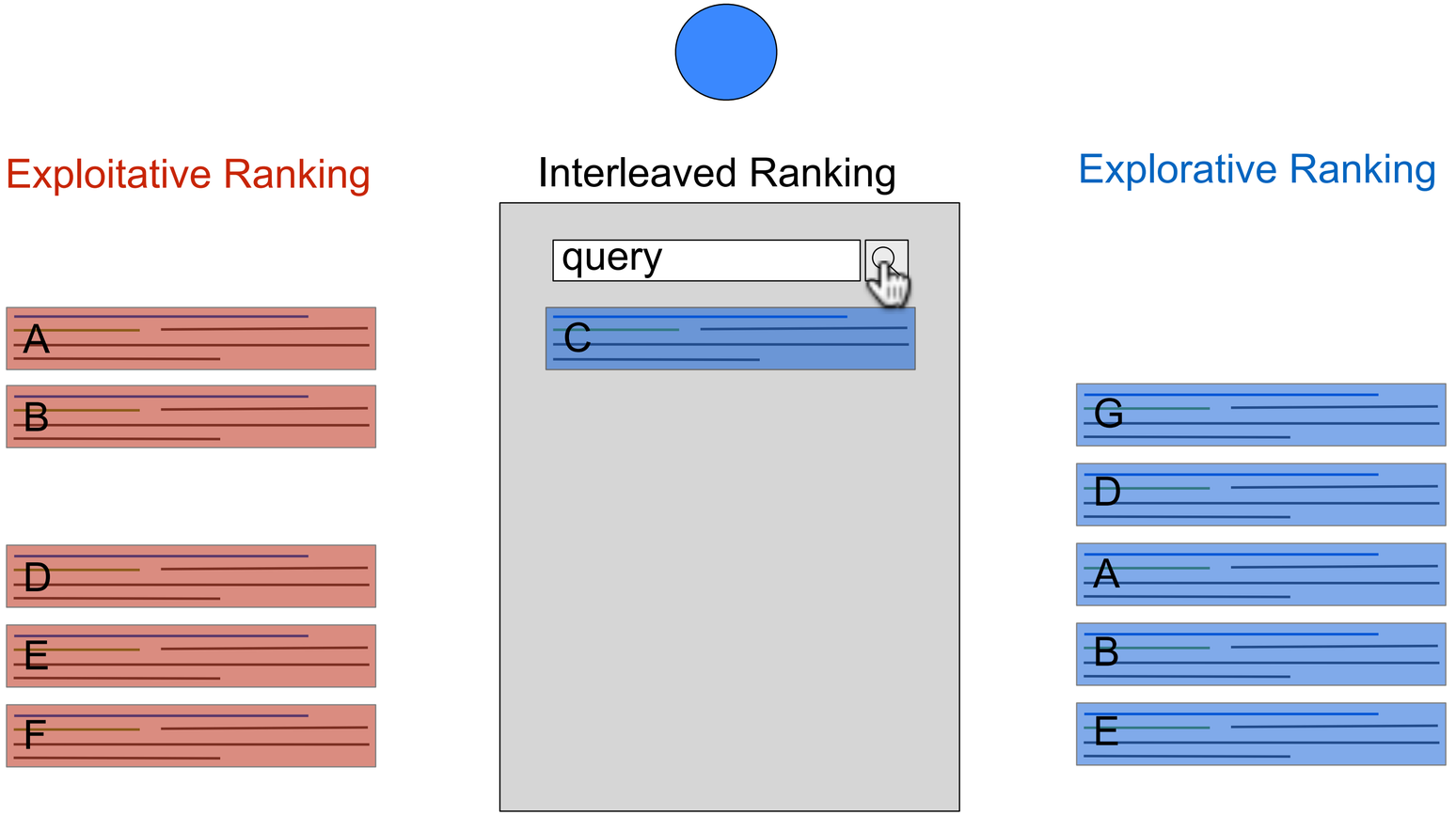
We are not yet doing this at Blendle



Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

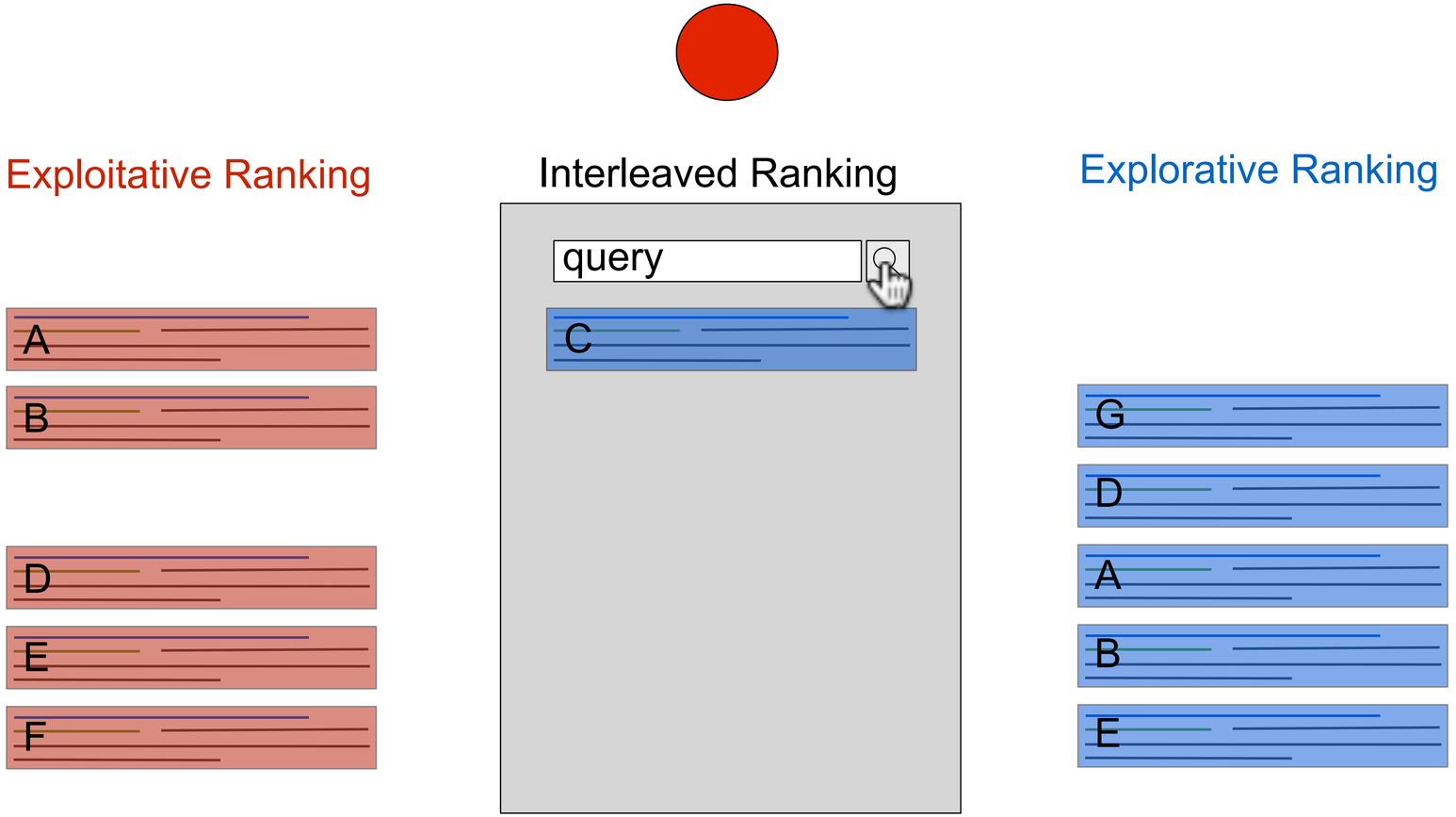
We are not yet doing this at Blendle



Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

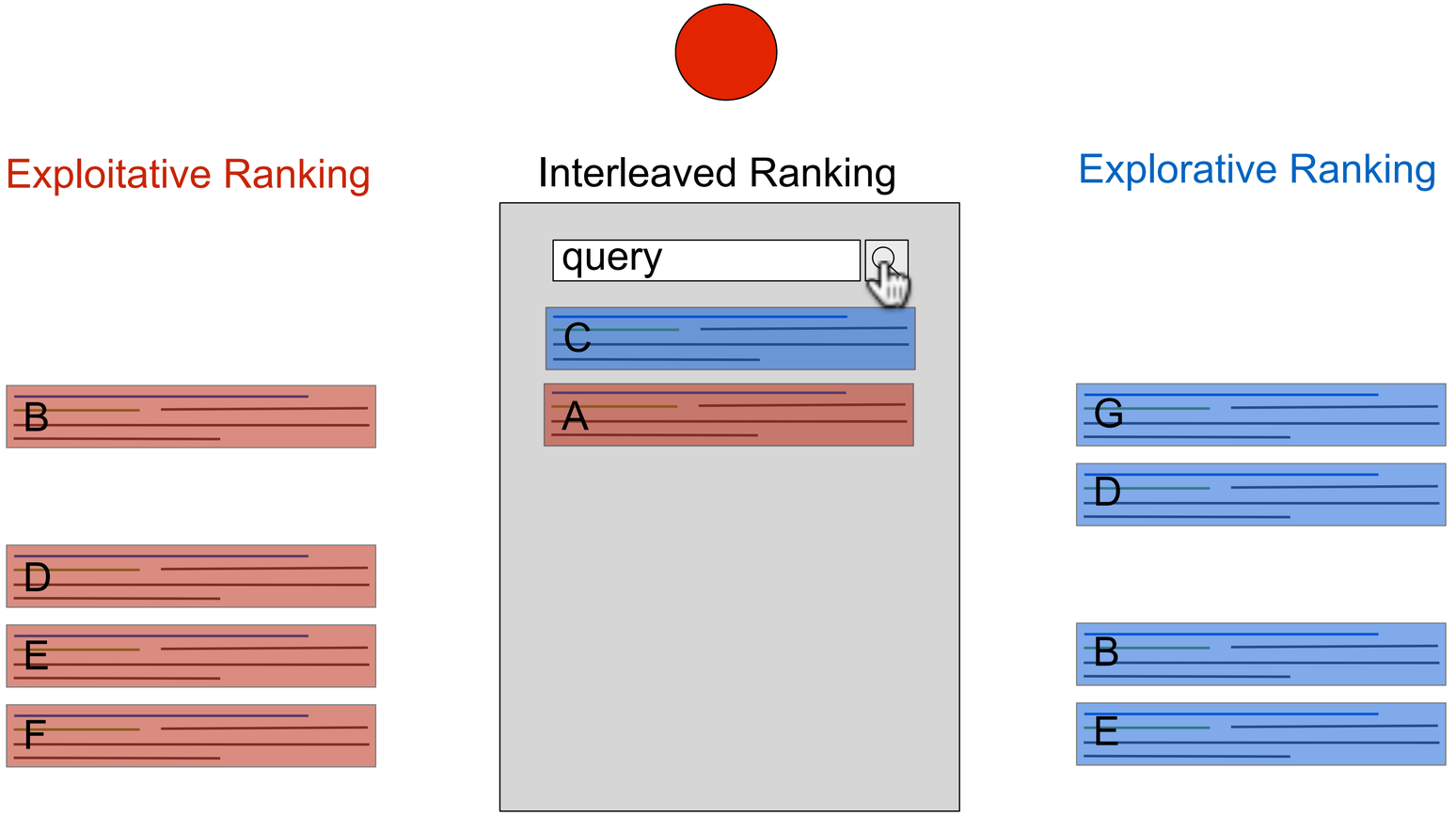
We are not yet doing this at Blendle



Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

We are not yet doing this at Blendle



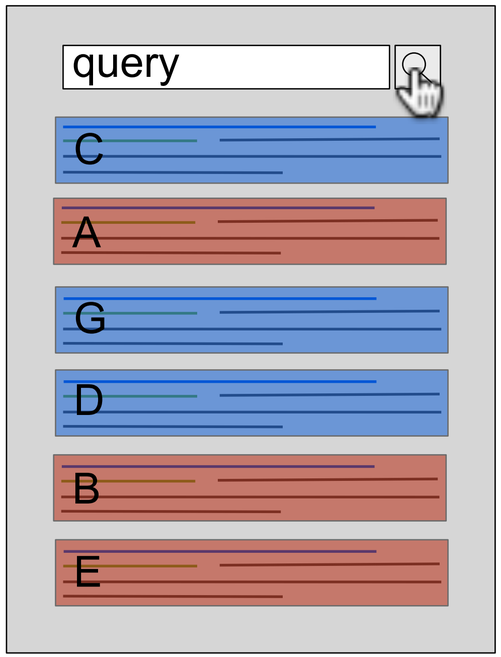
Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

We are not yet doing this at Blendle

Exploitative Ranking

Interleaved Ranking

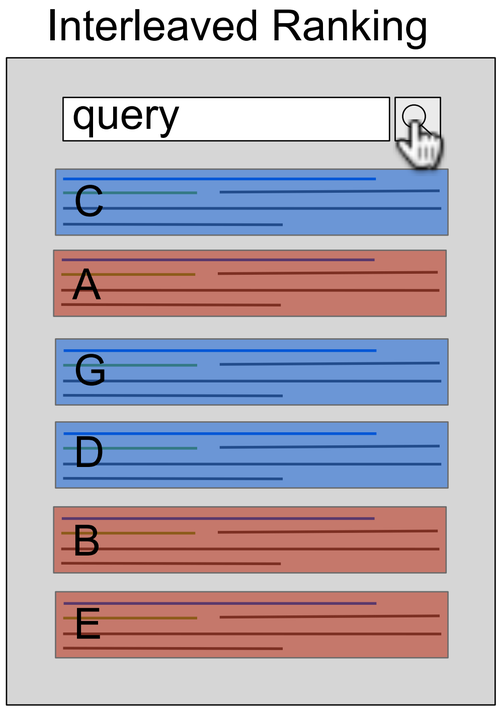


Explorative Ranking

Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

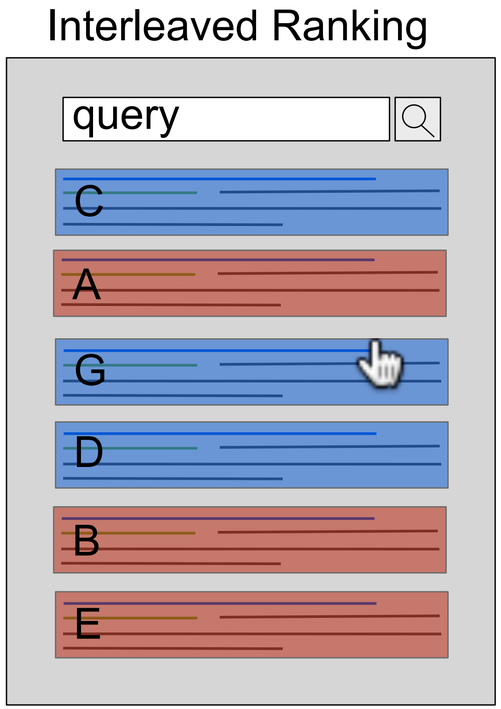
We are not yet doing this at Blendle



Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

TeamDraft Interleave

We are not yet doing this at Blendle

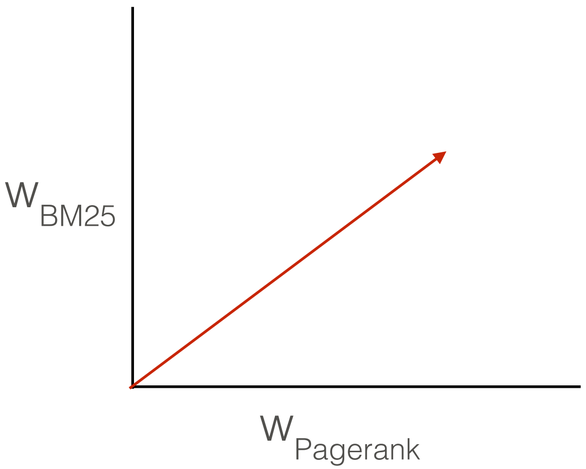


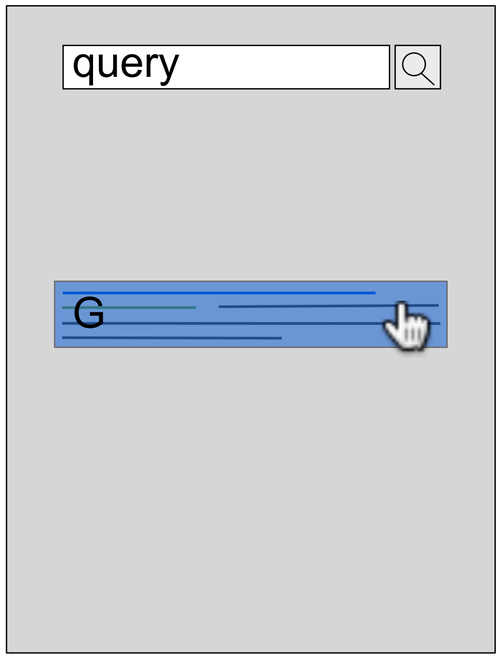
Radlinski, F., Kurup, M., & Joachims, T. (2008). How does clickthrough data reflect retrieval quality? In CIKM '08.

Dueling Bandit Gradient Descent

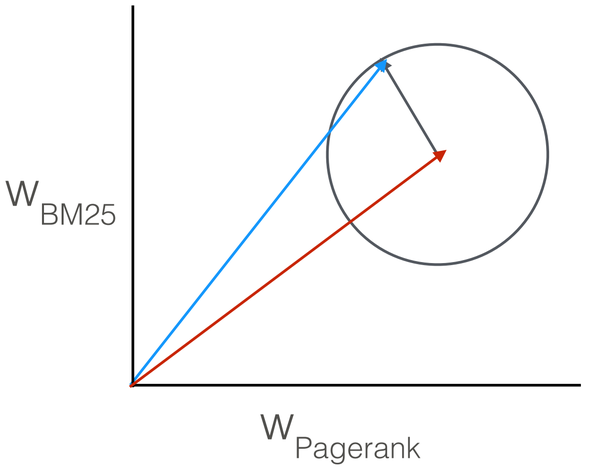
We are not yet doing this at Blendle

Exploitative Ranker





Explorative Ranker

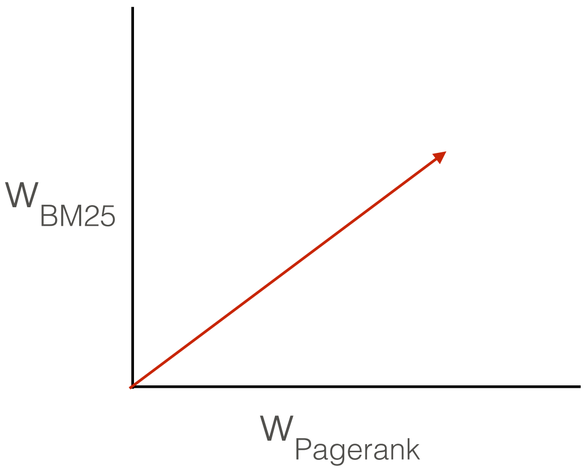


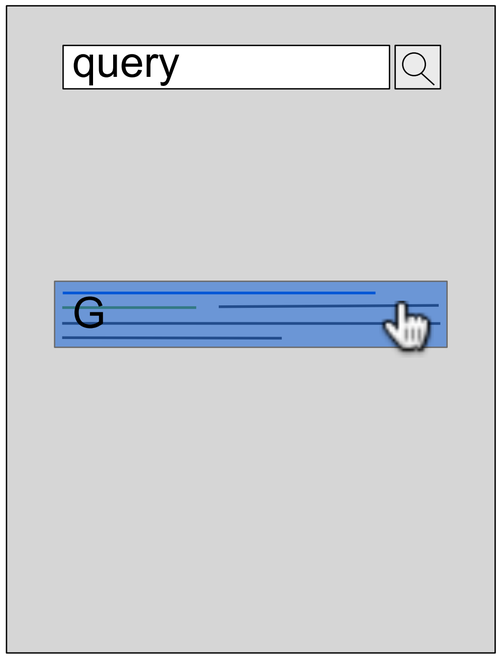
[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

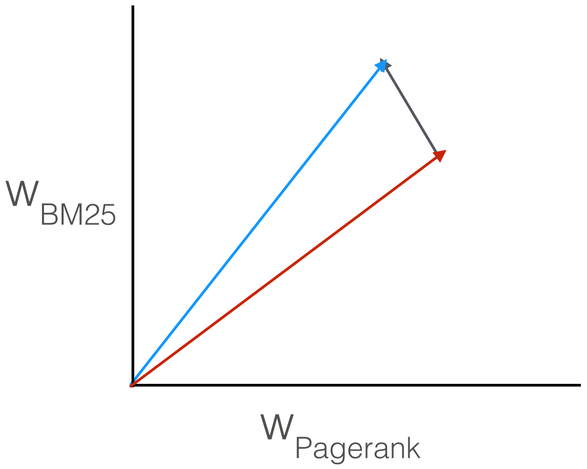
We are not yet doing this at Blendle

Exploitative Ranker





Explorative Ranker

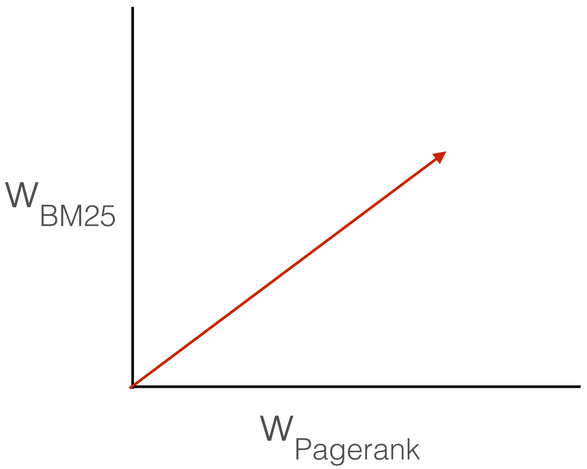


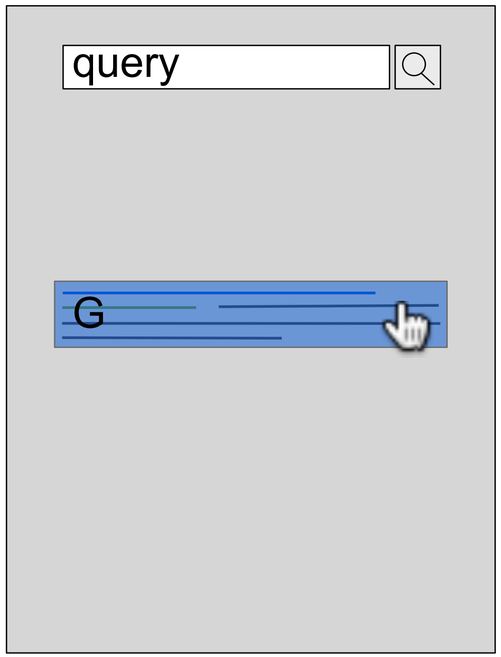
[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

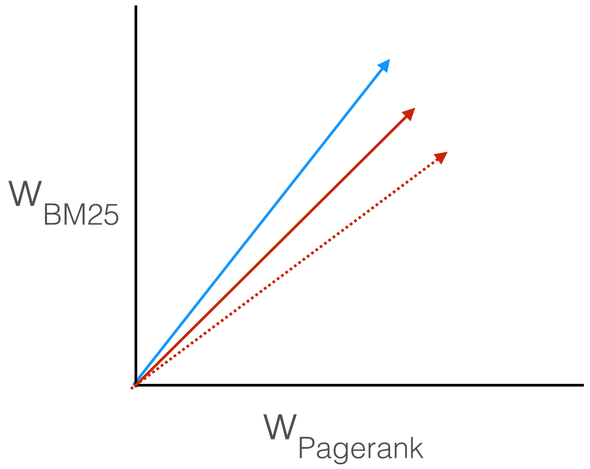
We are not yet doing this at Blendle

Exploitative Ranker





Explorative Ranker

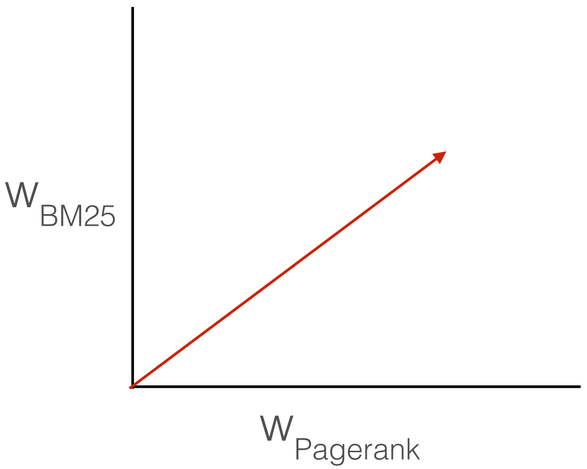


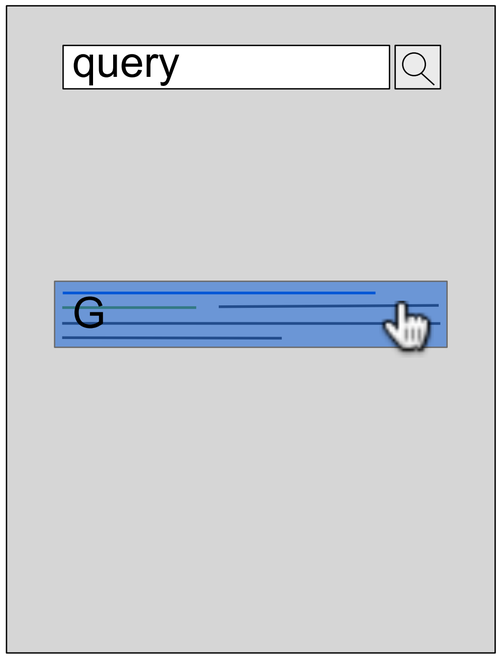
[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

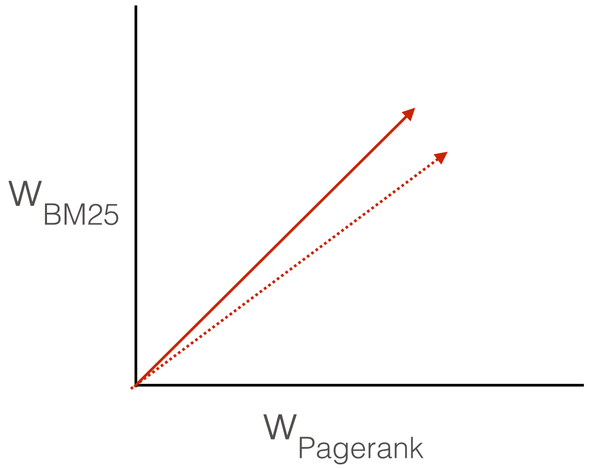
We are not yet doing this at Blendle

Exploitative Ranker





Explorative Ranker

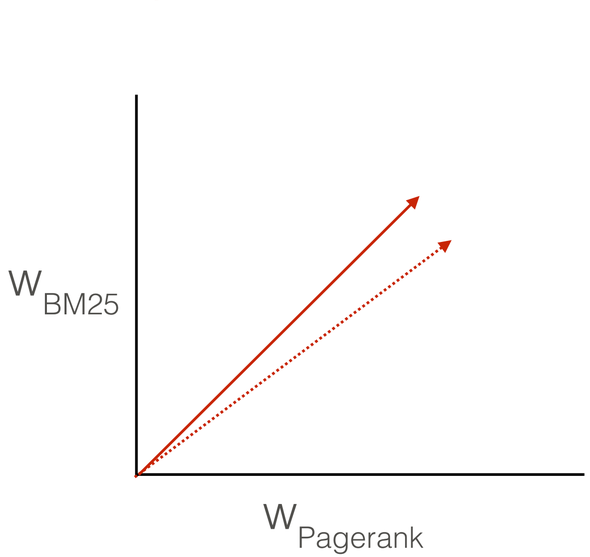


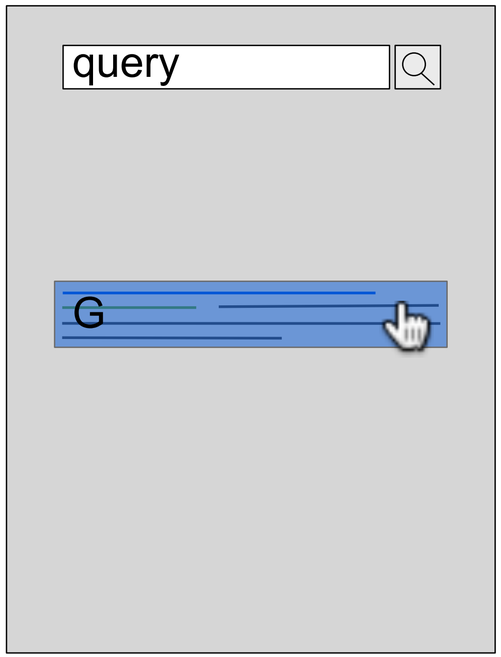
[Yue et al, 2009; Hofmann et al., 2011]

Dueling Bandit Gradient Descent

We are not yet doing this at Blendle

Exploitative Ranker





[Yue et al, 2009; Hofmann et al., 2011]

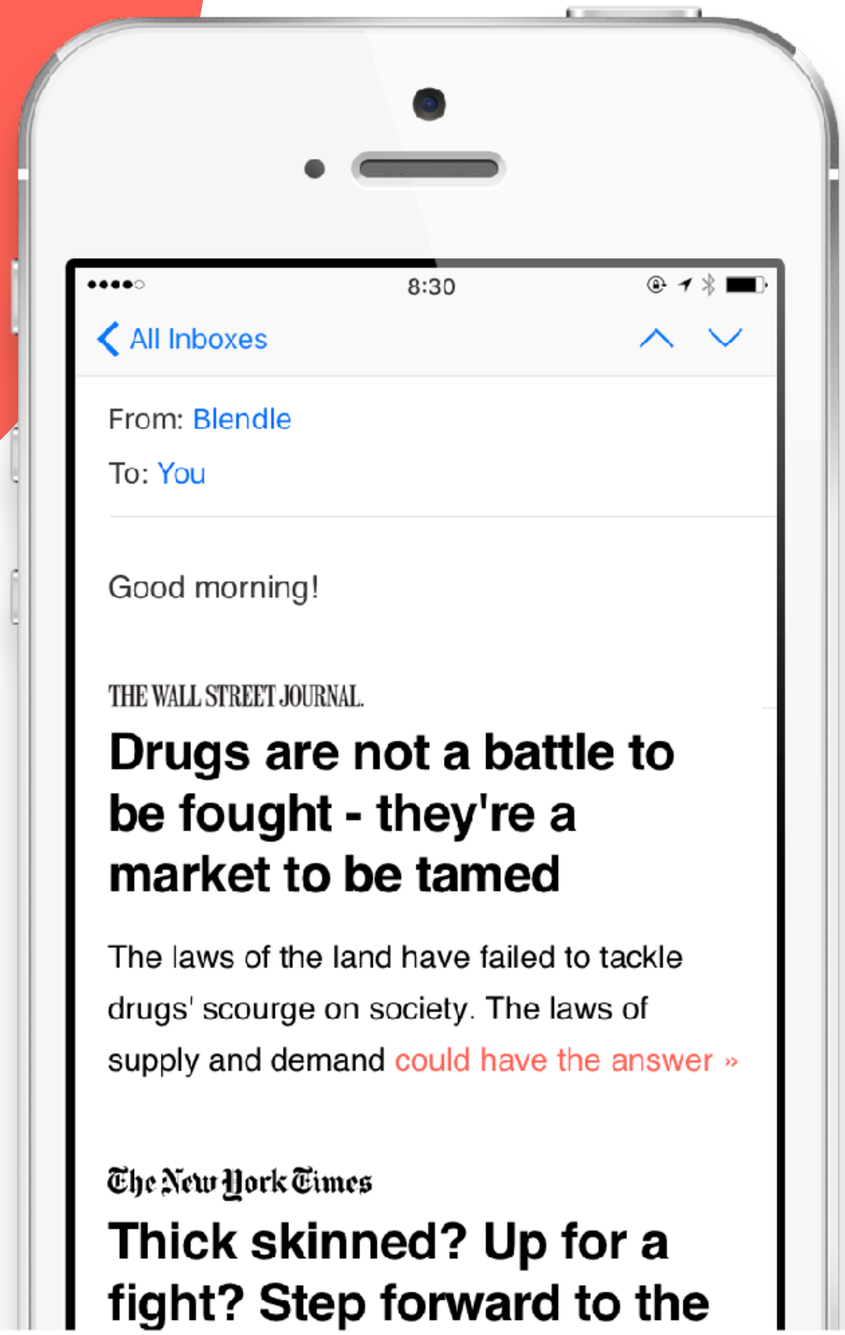
Offline vs Online

Offline Learning

- [x] Better studied, so more ML tools.
- [ ] Not easily adaptable as things change.
- [ ] Often (but not always) the method requires explicit annotations.

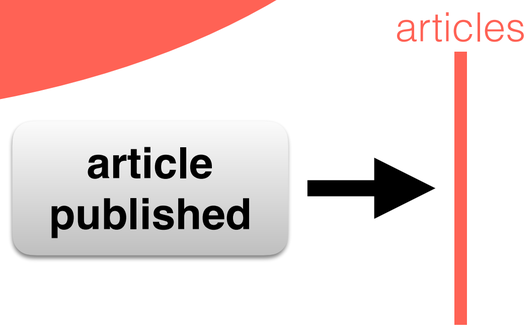
Online Learning

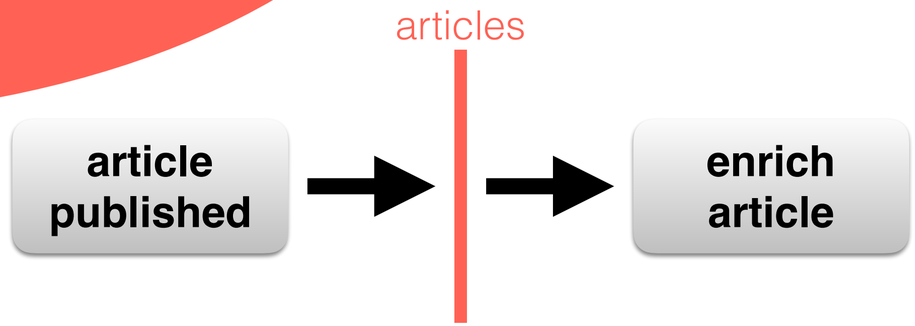
- [x] Better for a dynamic environnement
- [x] More sample efficient since the algorithm plays an active role.
- [ ] Can't (easily) go back and retrain.

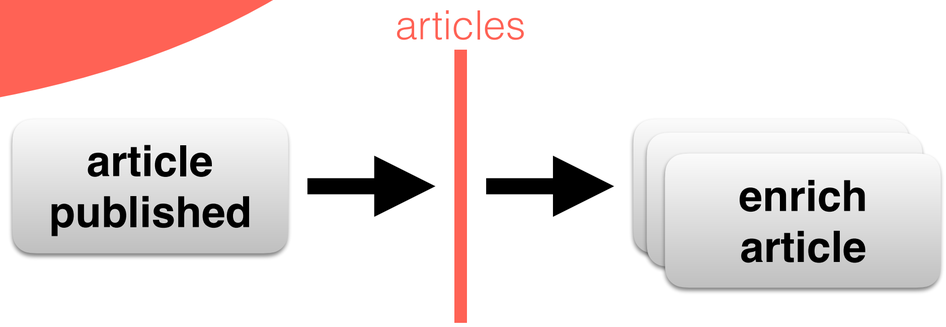


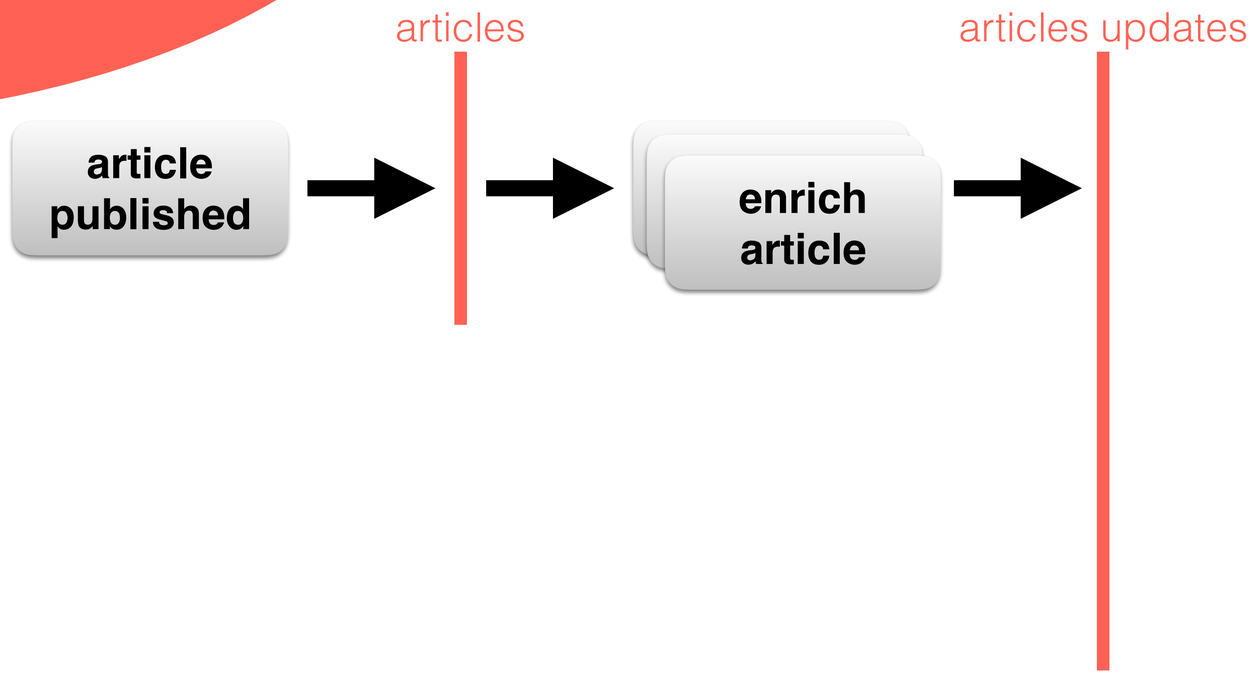
Timing problem

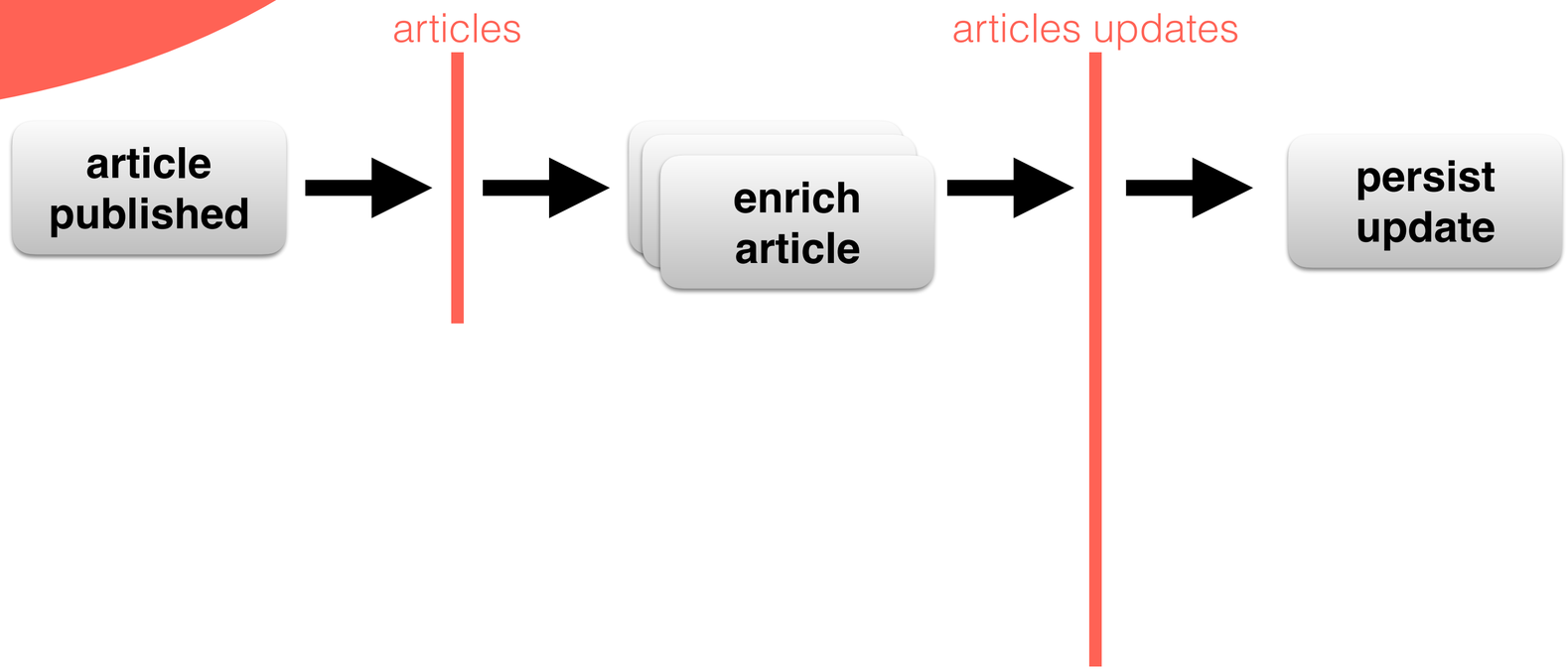
- Our editors wake up at 5am and are done reading at 8am
- Which is also when we want to send our newsletter
- We simply can’t wait for a batch process

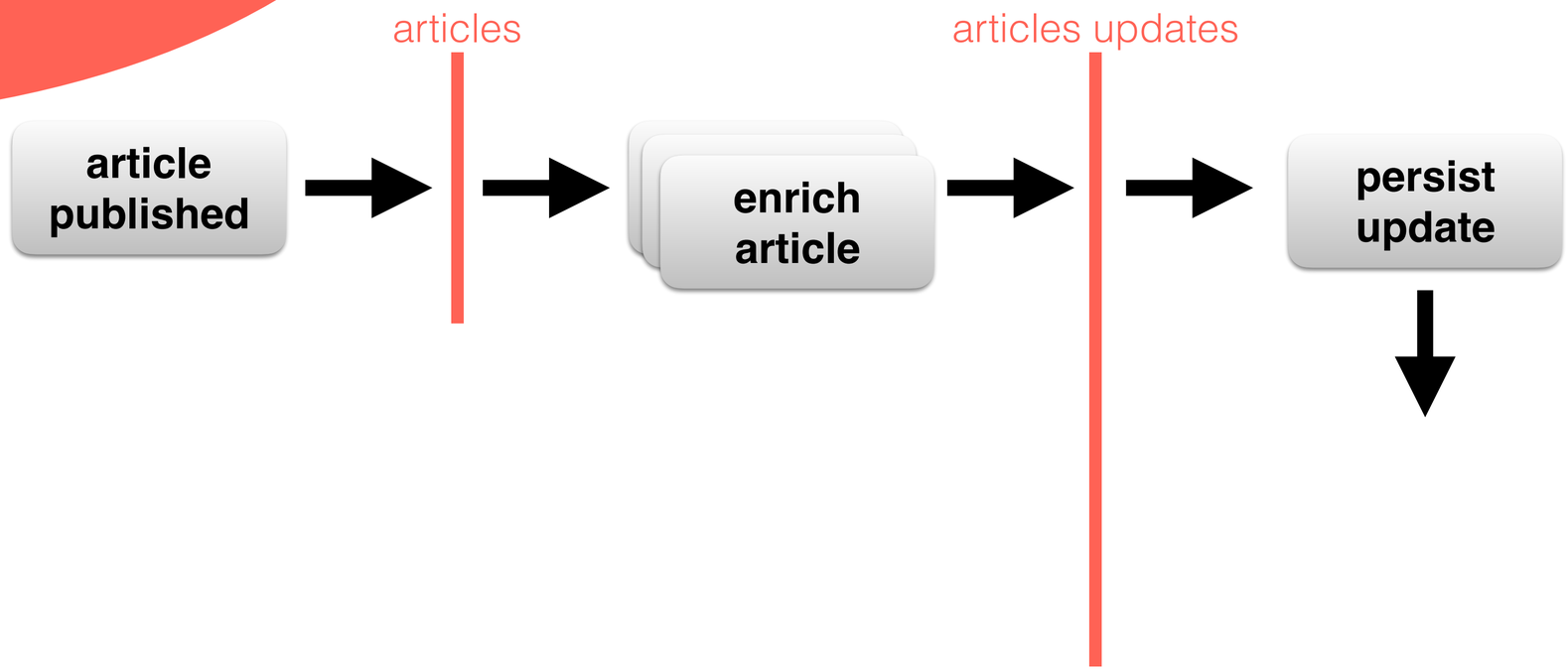

articles
article published

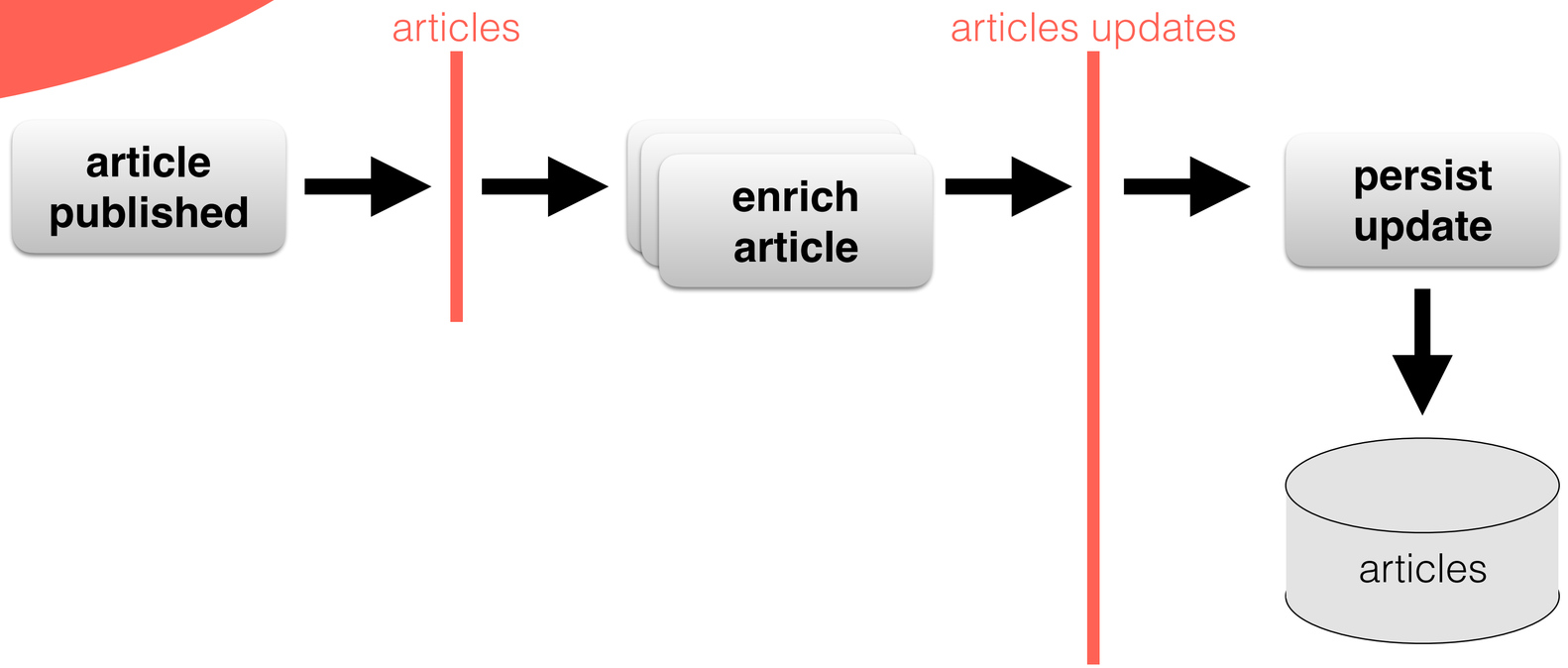

articles
article published
enrich article

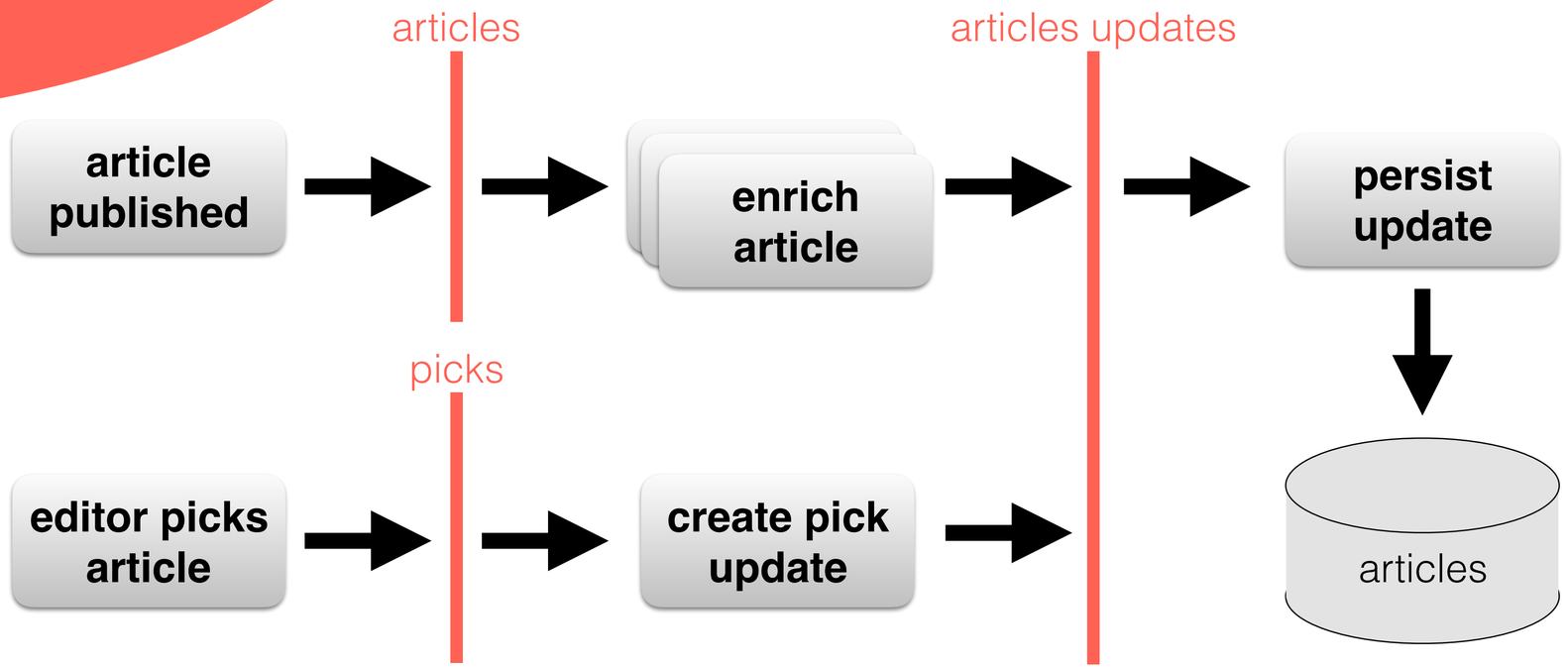

articles
article published
enrich article

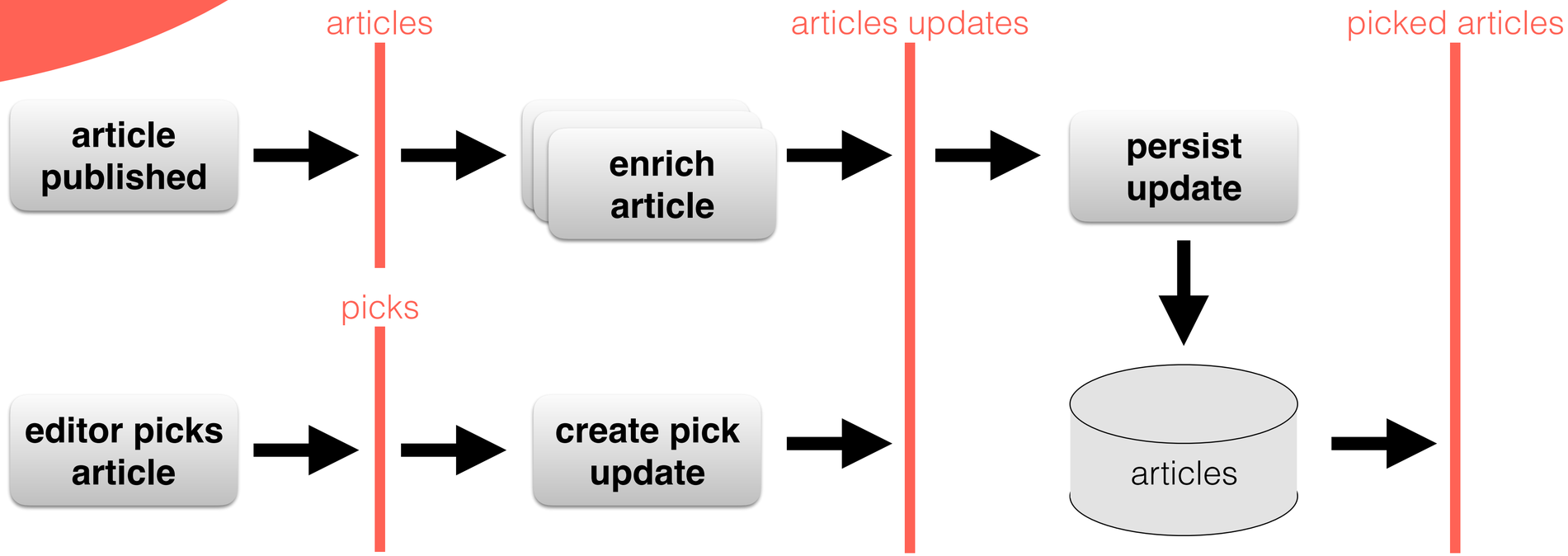

articles
articles updates
article published
enrich article

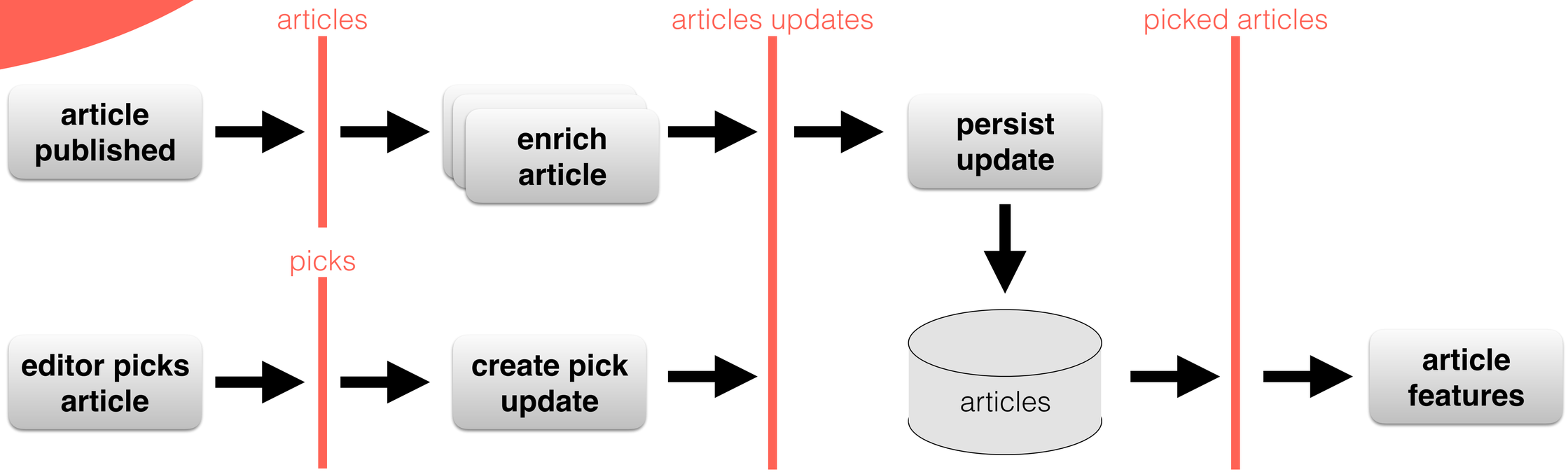

articles
articles updates
article published
enrich article
persist update

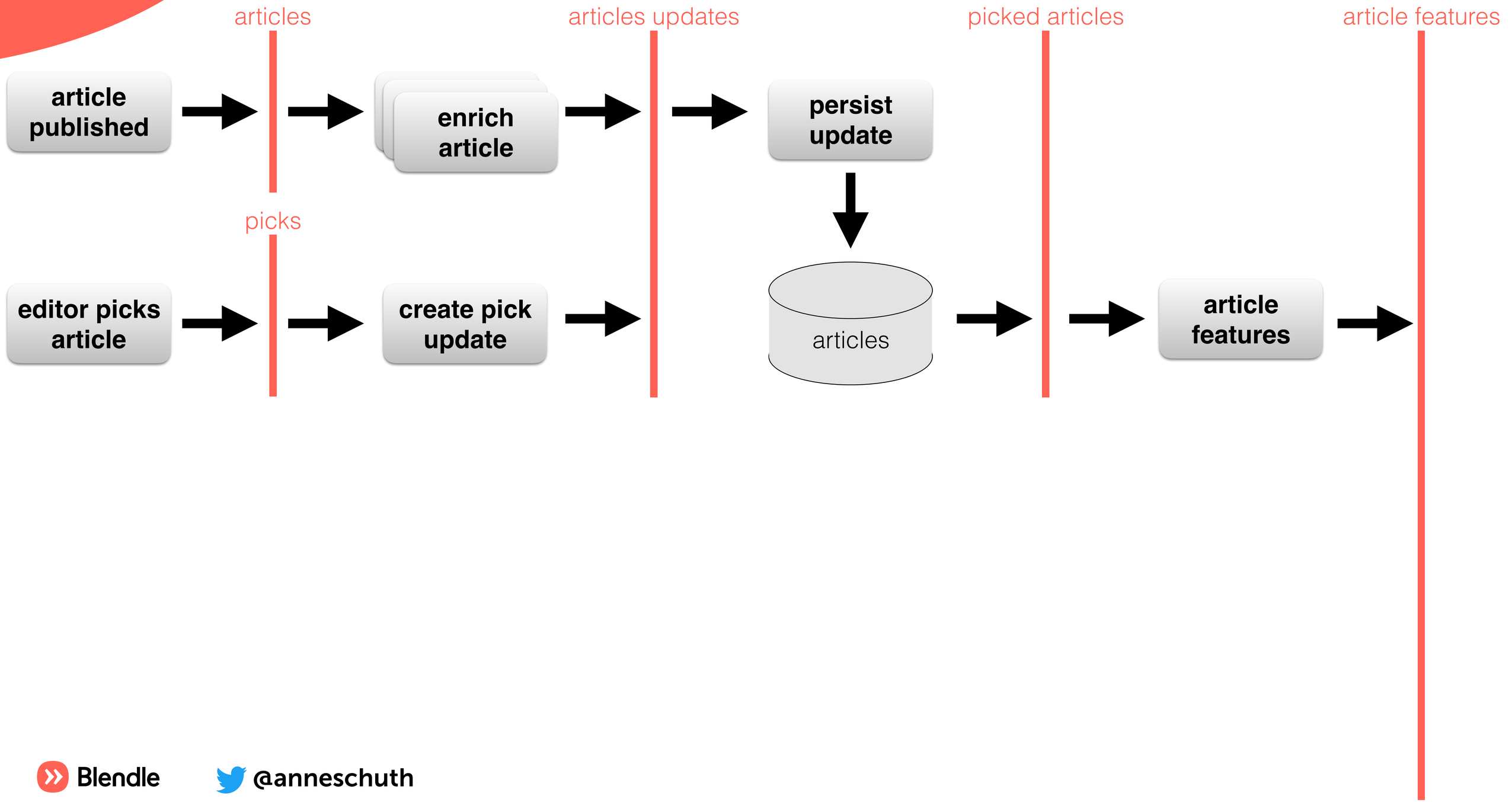

articles
articles updates
article published
enrich article
persist update


articles
articles updates
article published
enrich article
persist update
articles

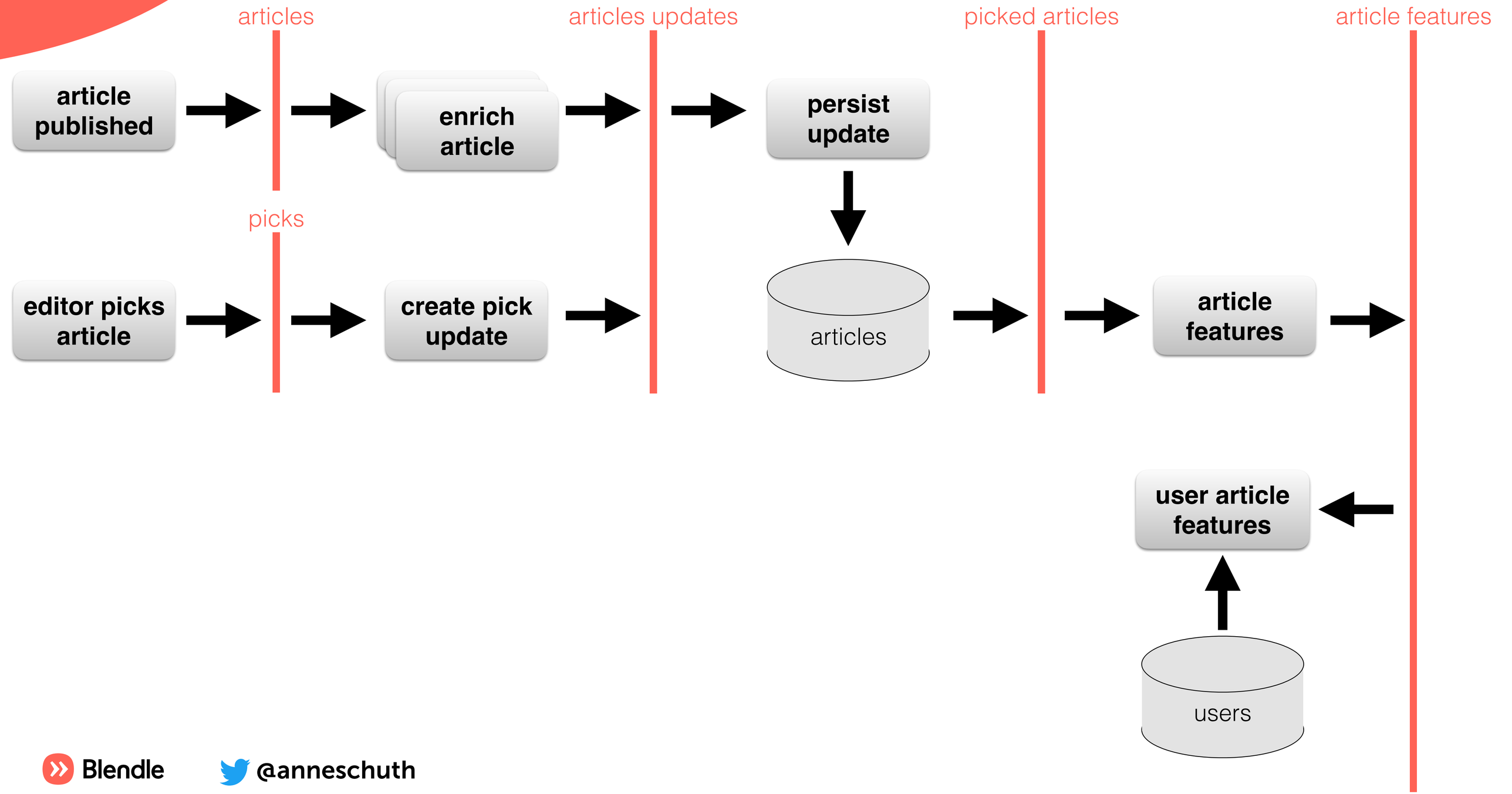

articles
articles updates
article published
enrich article
persist update
picks
editor picks article
create pick update
articles


articles
articles updates
picked articles
article published
enrich article
persist update
picks
editor picks article
create pick update
articles

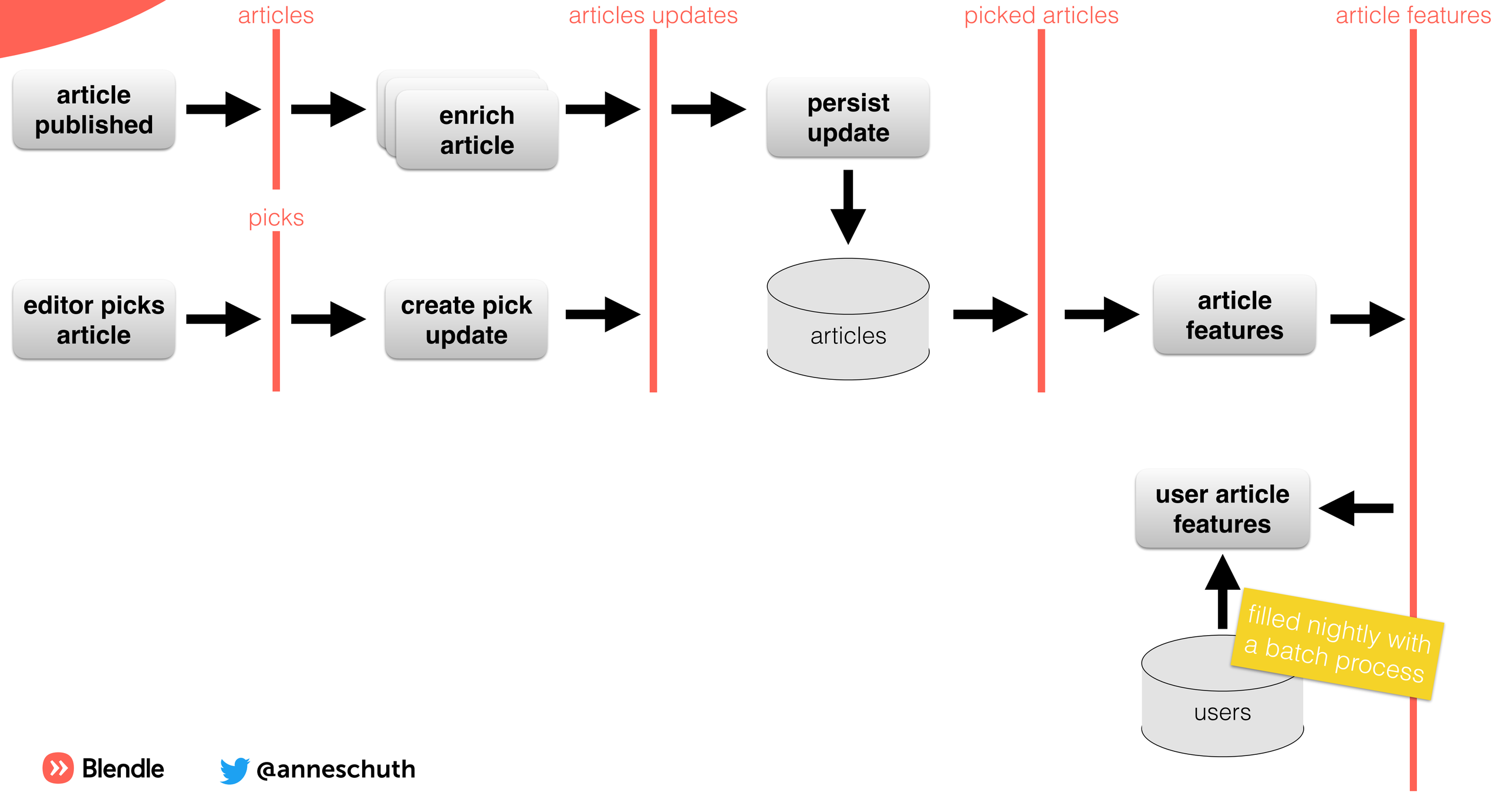

articles
articles updates
picked articles
article published
enrich article
persist update
picks
editor picks article
create pick update
articles
article features



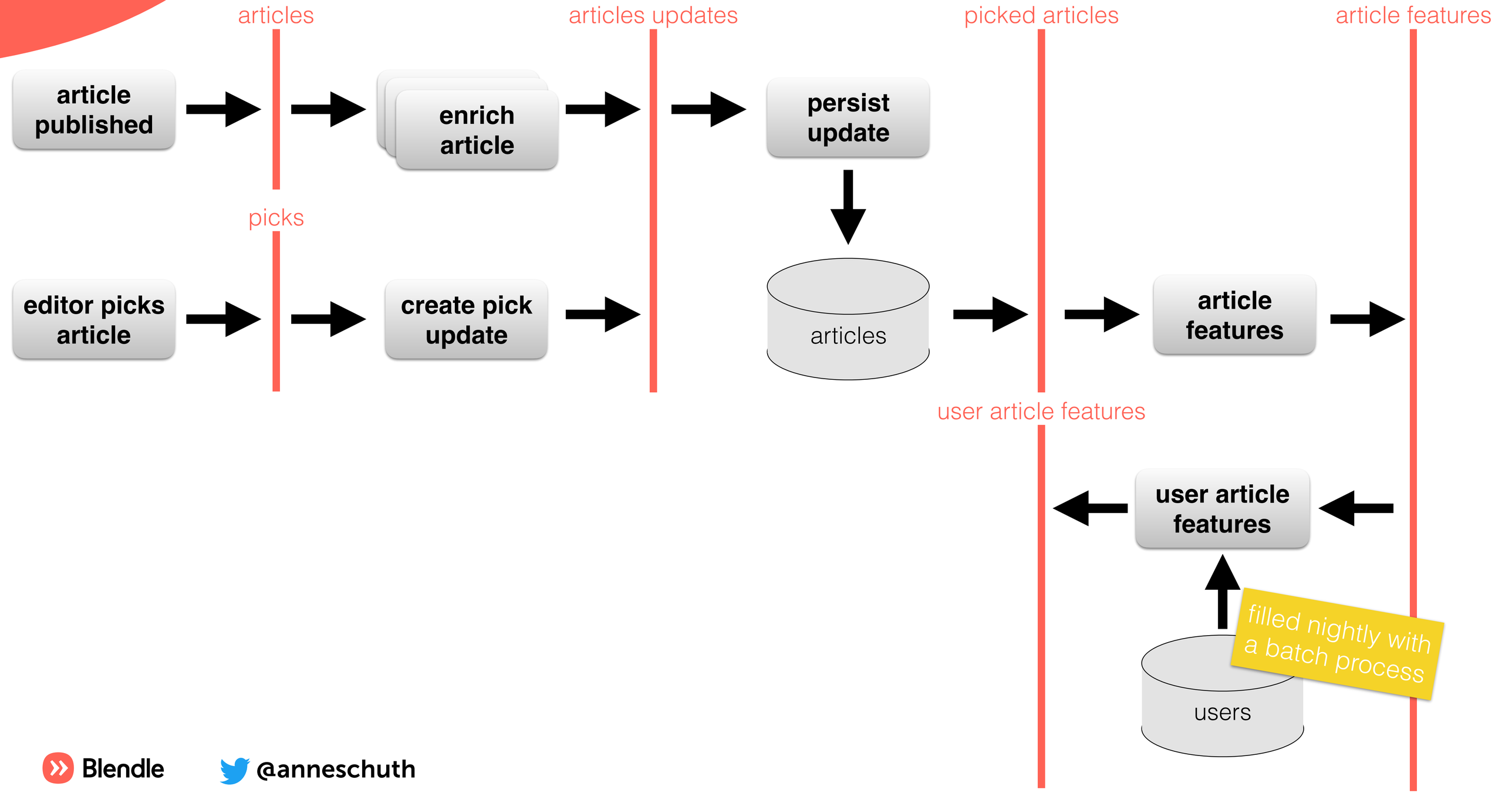
Blendle @anneschuth



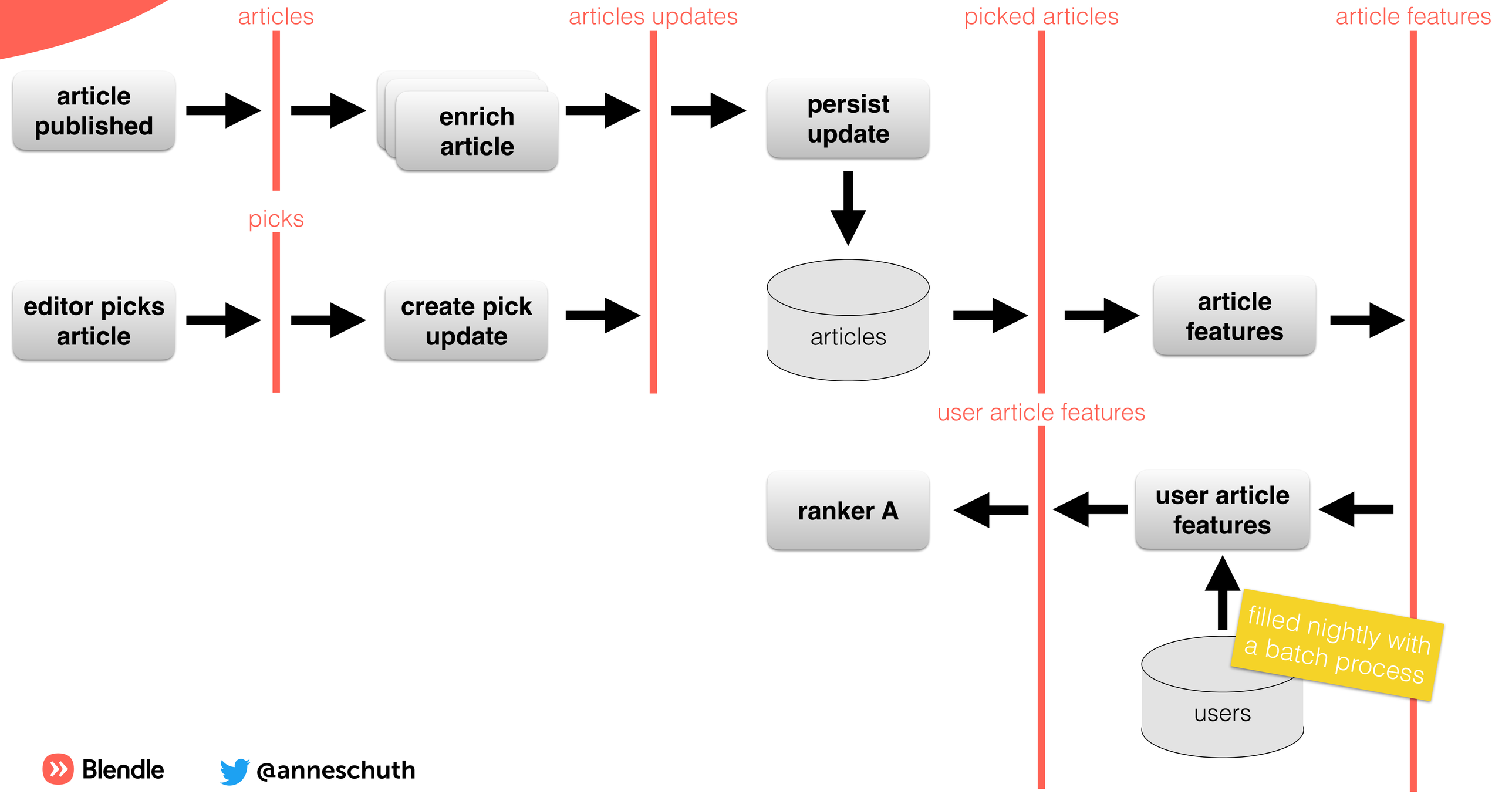
Blendle @anneschuth



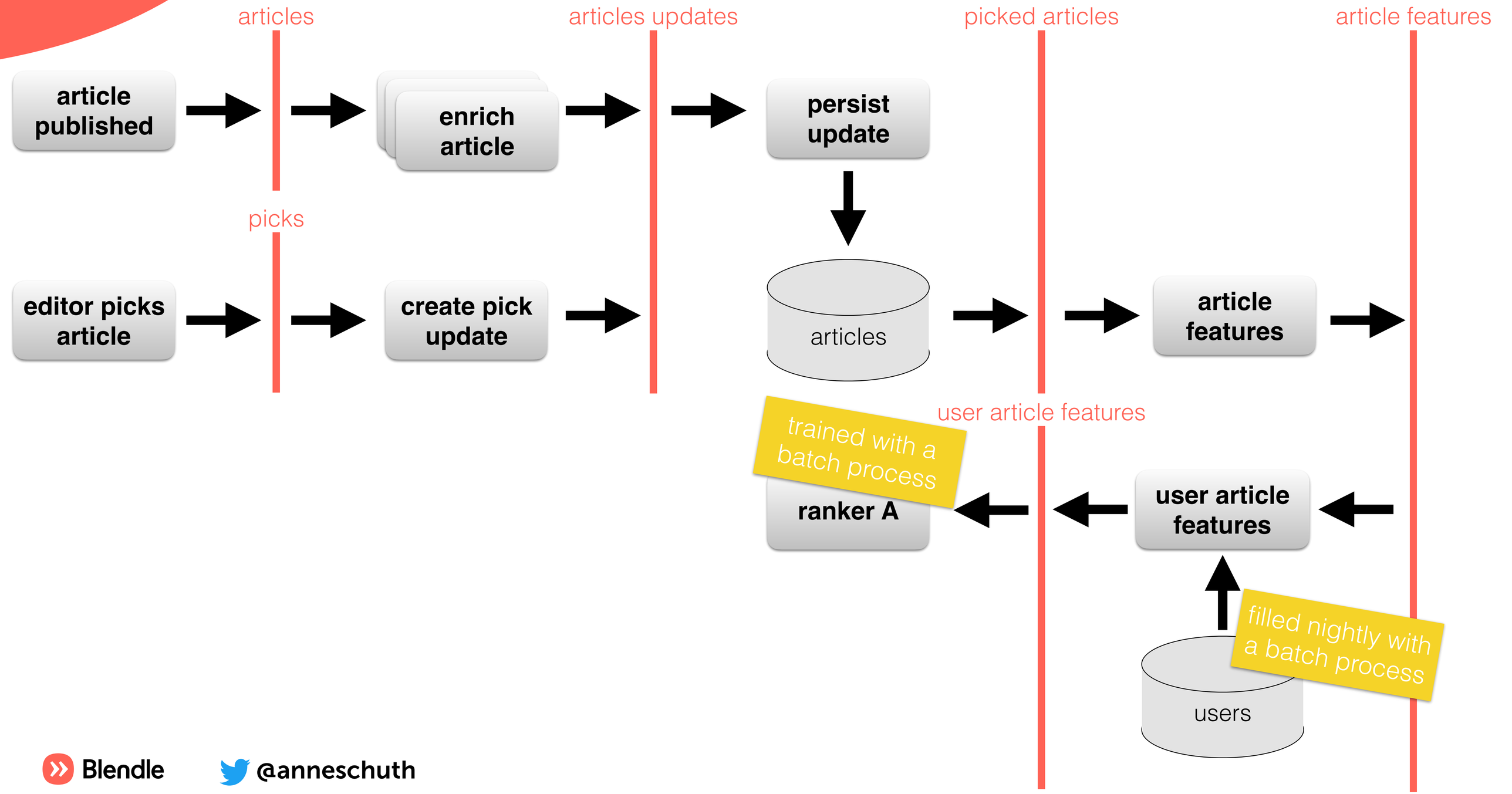
Blendle @anneschuth

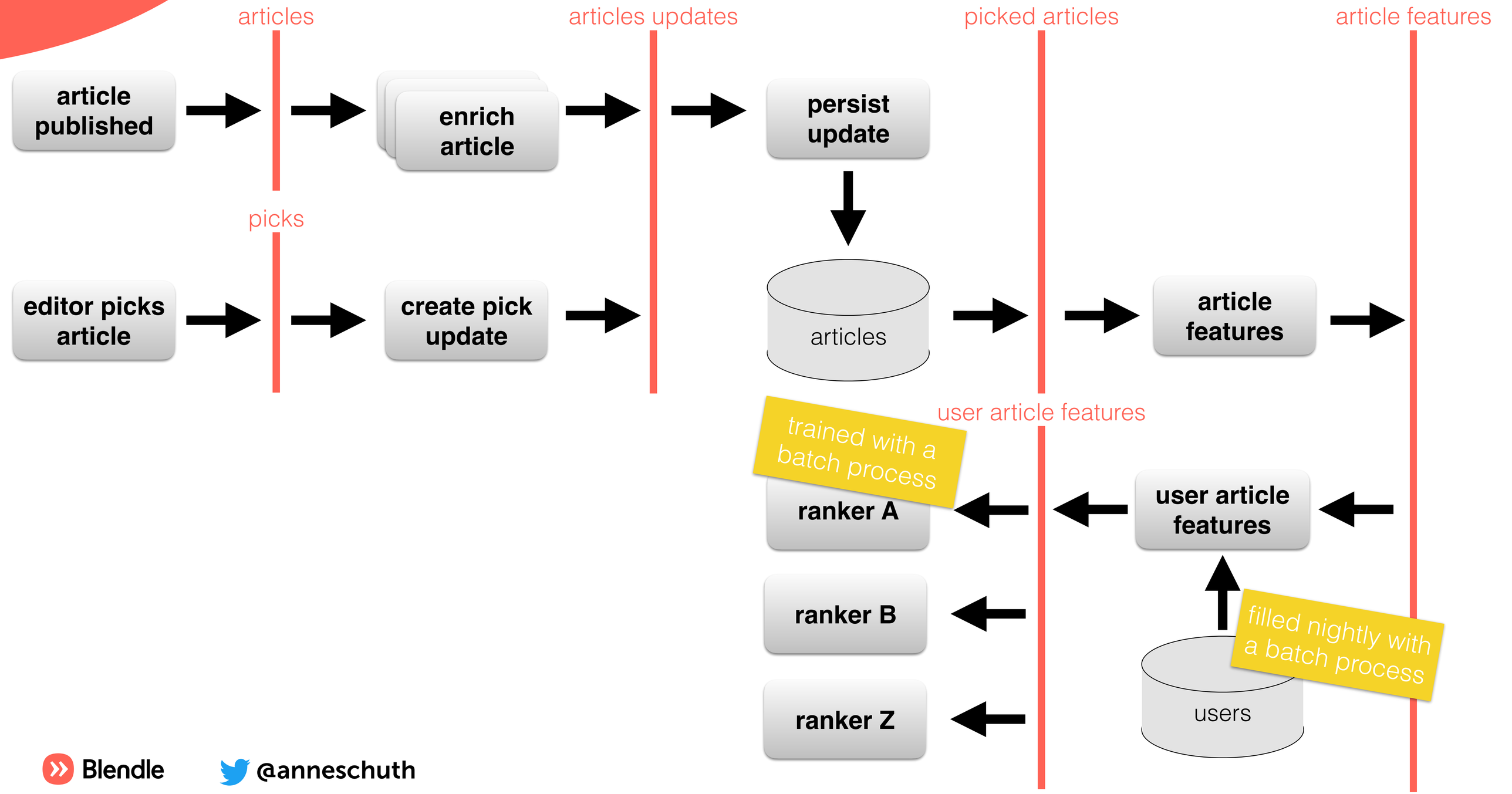


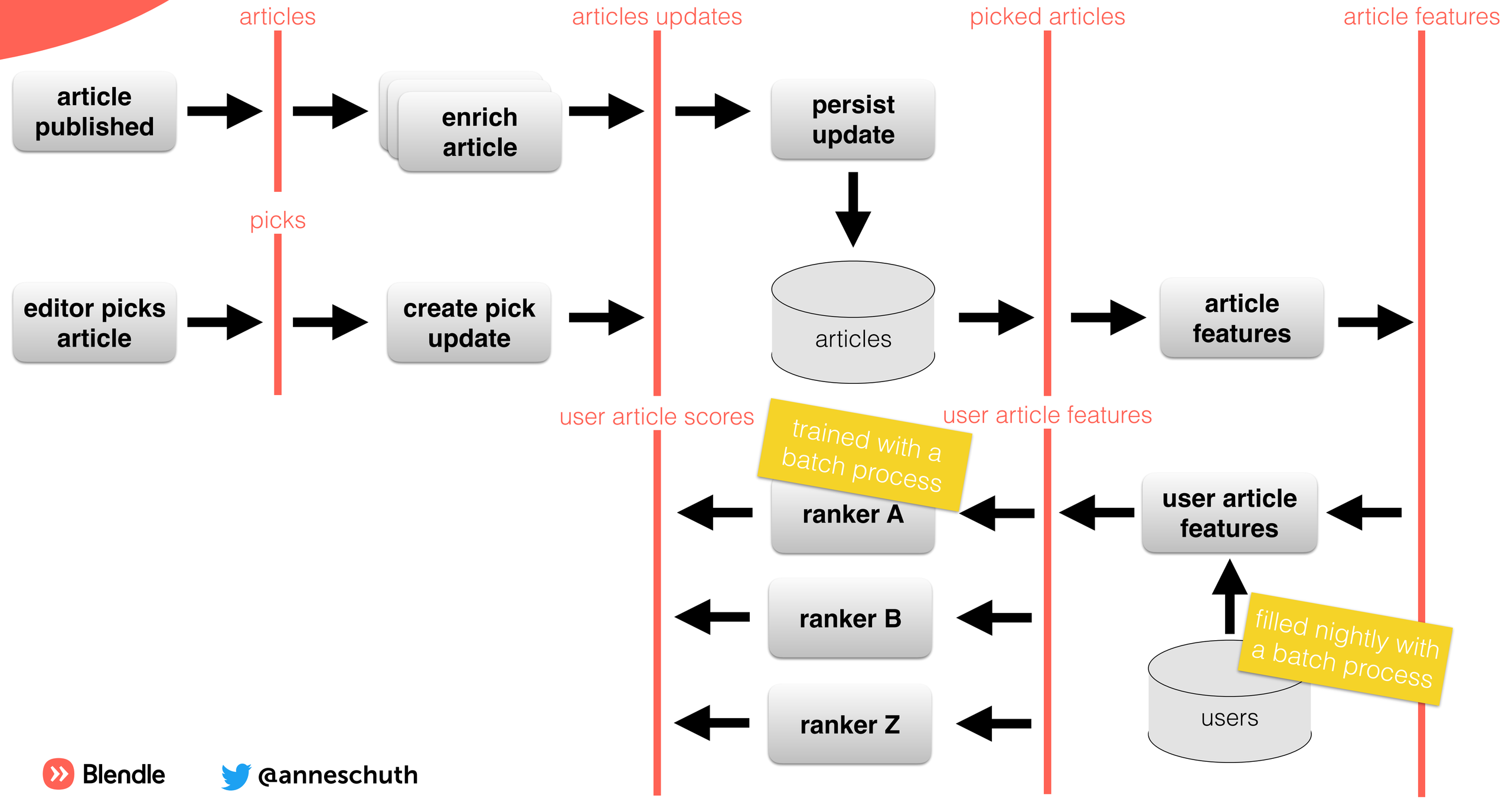
Blendle @anneschuth

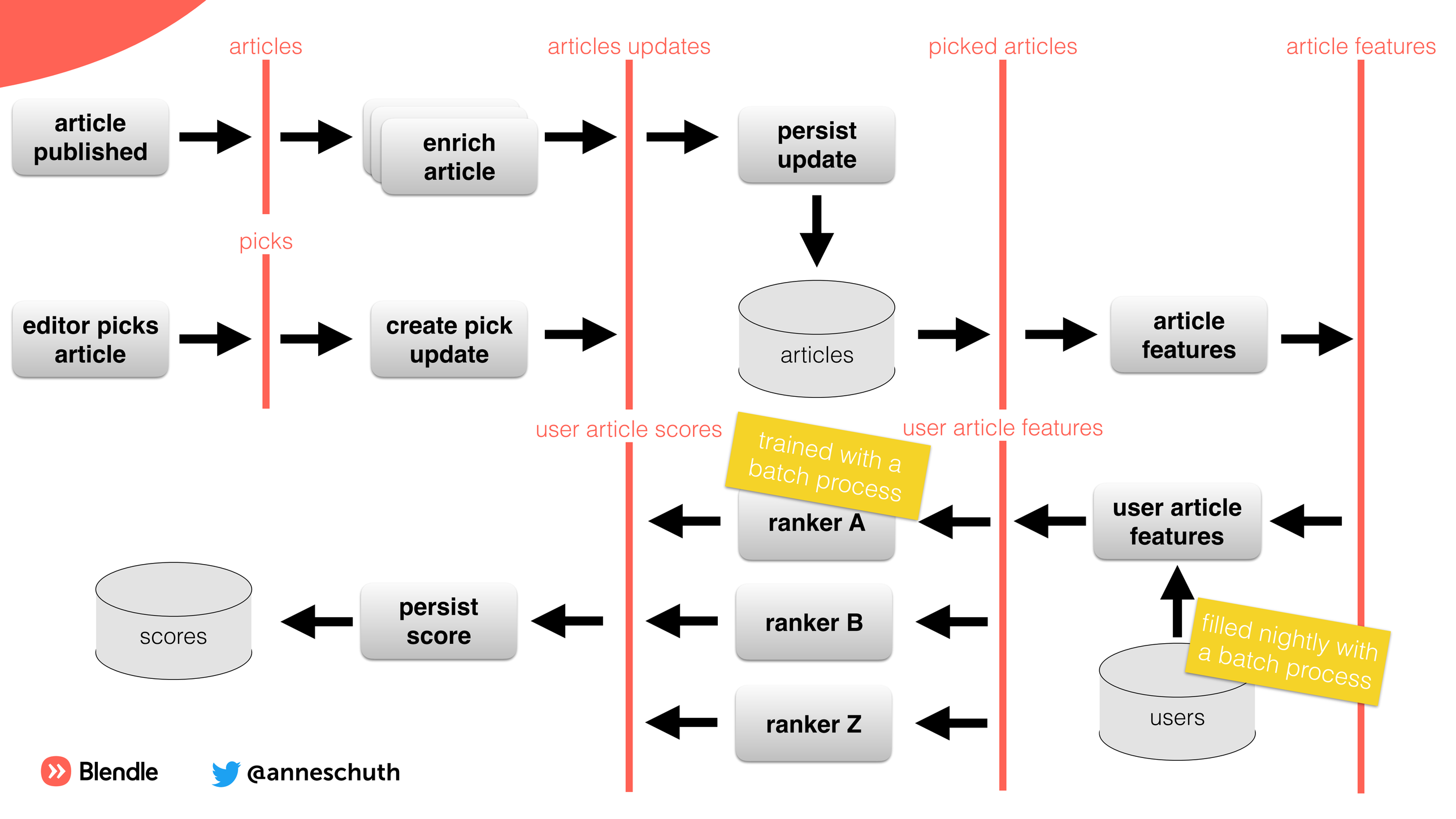


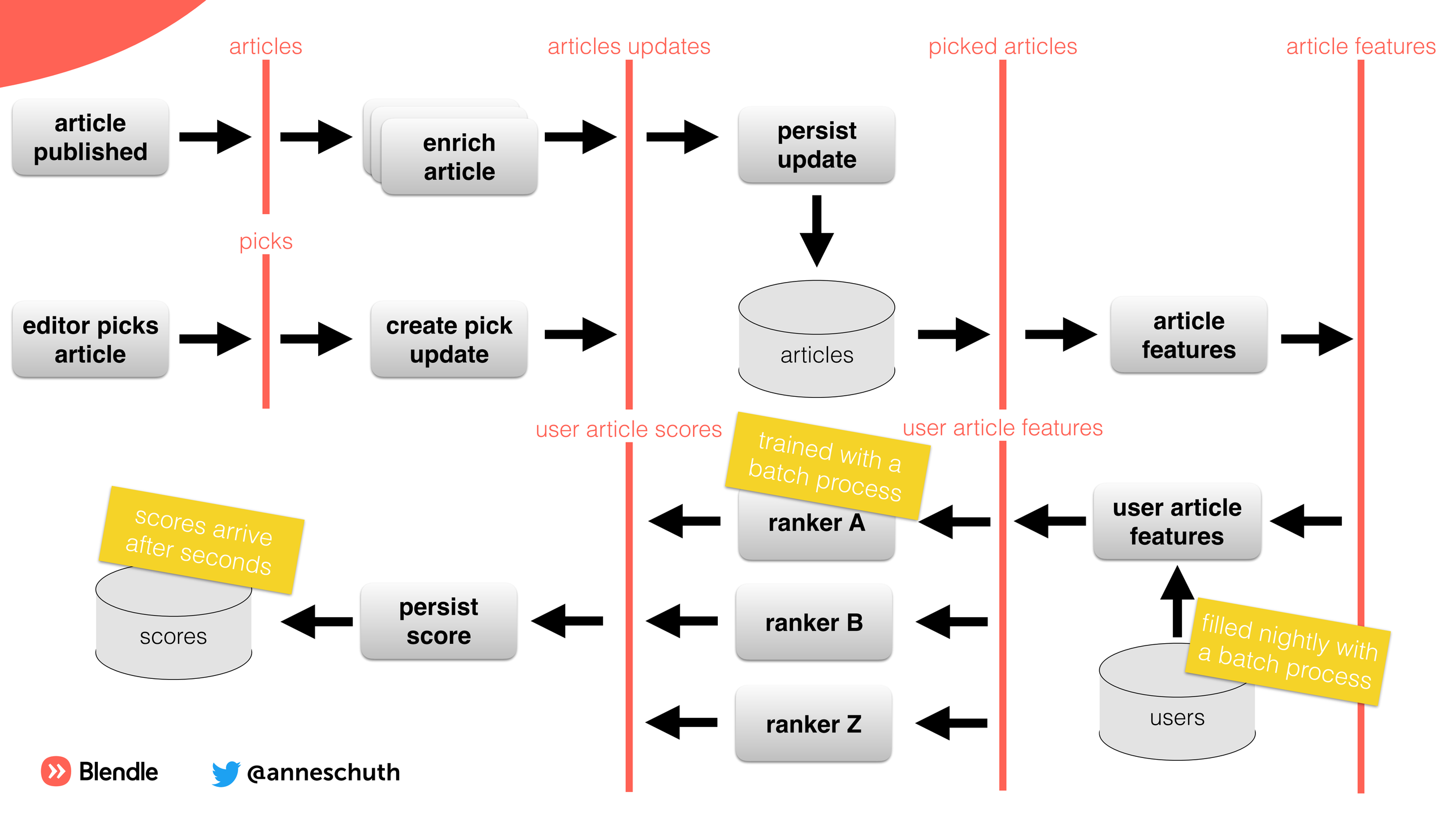
Blendle @anneschuth

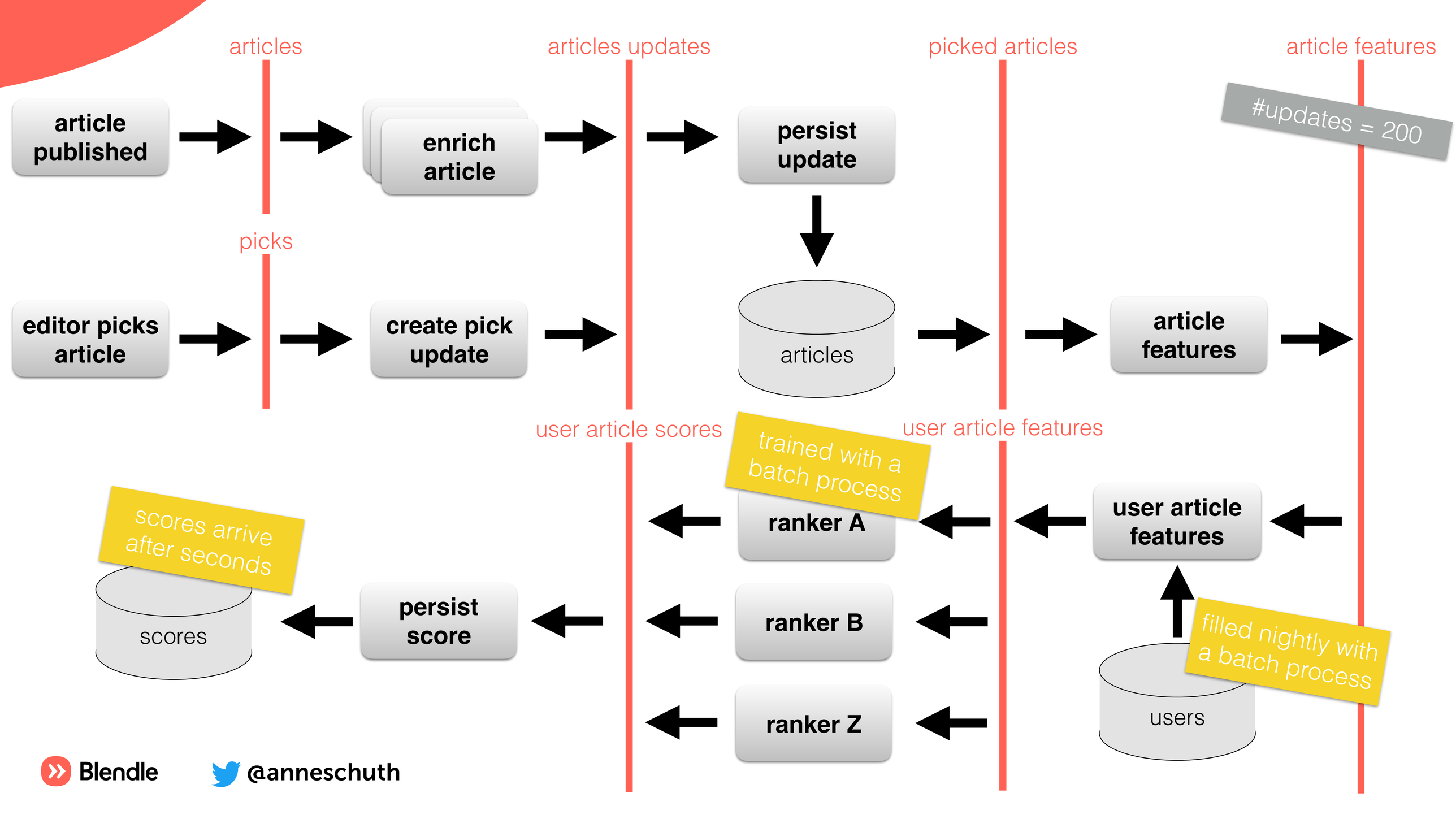

articles
articles updates
picked articles
article features
article published
enrich article
persist update
picks
editor picks article
create pick update
articles
article features
user article features
trained with a batch process
ranker A
user article features
filled nightly with a batch process
users

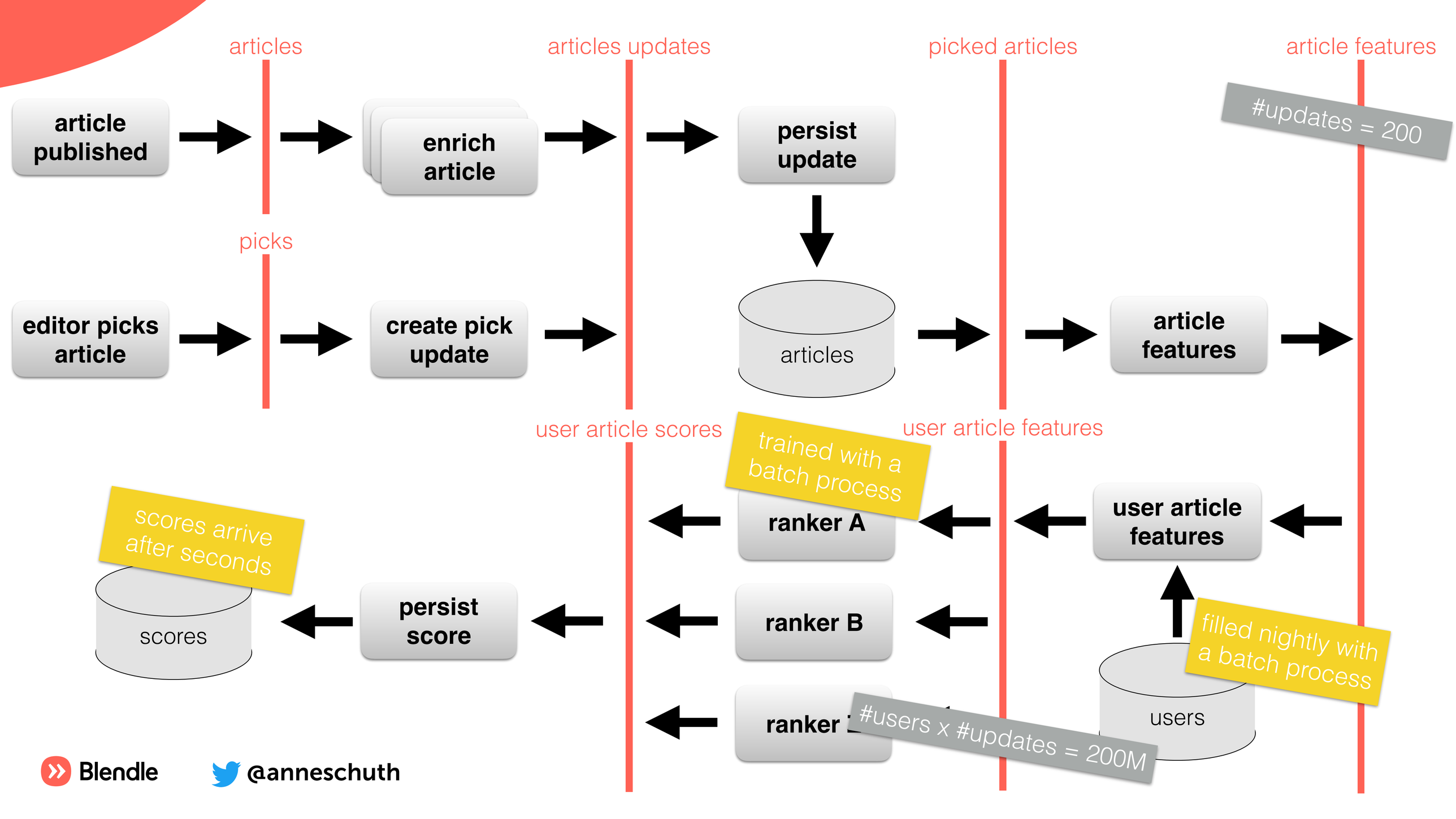

articles
articles updates
picked articles
article features
article published
enrich article
persist update
picks
editor picks article
create pick update
articles
article features
user article features
trained with a batch process
ranker A
user article features
ranker B
filled nightly with a batch process
ranker Z
users

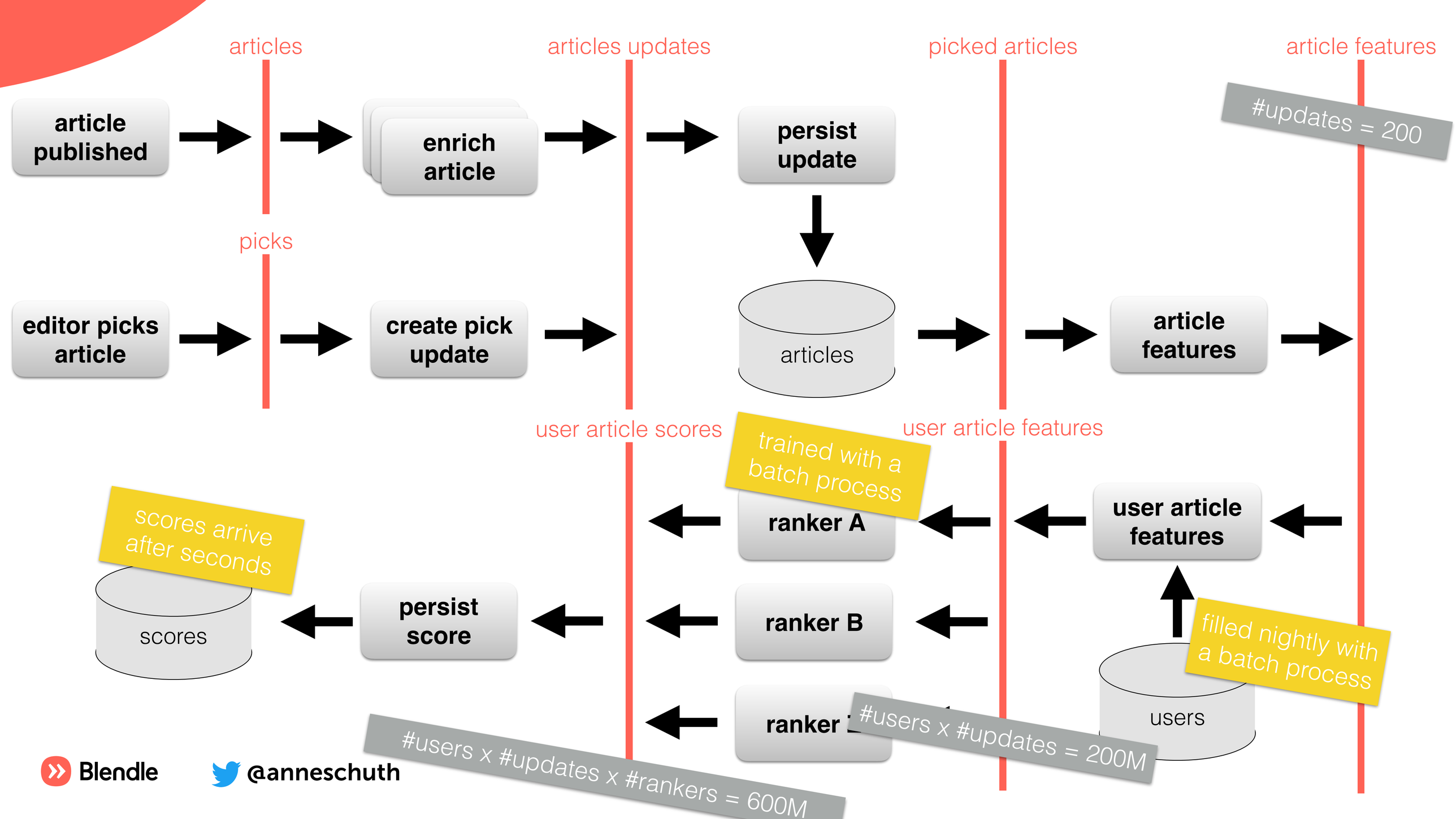

articles
articles updates
picked articles
article features
article published
enrich article
persist update
picks
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
ranker B
ranker Z
filled nightly with a batch process
users

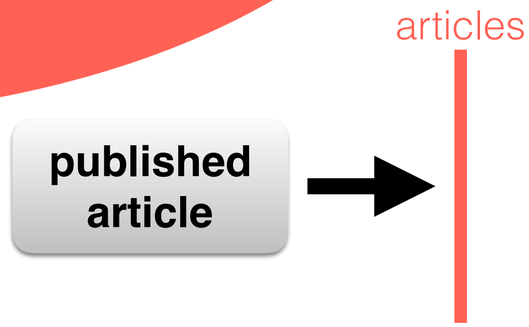

articles
articles updates
picked articles
article features
article published
enrich article
persist update
picks
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores
persist score
ranker B
filled nightly with a batch process
ranker Z
users
Blendle
@anneschuth


articles
articles updates
picked articles
article features
article published
enrich article
persist update
picks
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
scores
persist score
ranker B
ranker Z
filled nightly with a batch process
users
Blendle
@anneschuth


articles
articles updates
picked articles
article features
article published
enrich article
persist update
#updates = 200
picks
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
scores
persist score
ranker B
filled nightly with a batch process
ranker Z
users
Blendle
@anneschuth


articles
articles updates
picked articles
article features
article published
enrich article
persist update
#updates = 200
picks
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
scores
persist score
ranker B
filled nightly with a batch process
users
#users x #updates = 200M
Blendle
@anneschuth


articles
articles updates
picked articles
article features
article published
enrich article
persist update
#updates = 200
picks
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
scores
persist score
ranker B
filled nightly with a batch process
users
#users x #updates = 200M
#users x #updates x #rankers = 600M
Blendle
@anneschuth


articles
published article

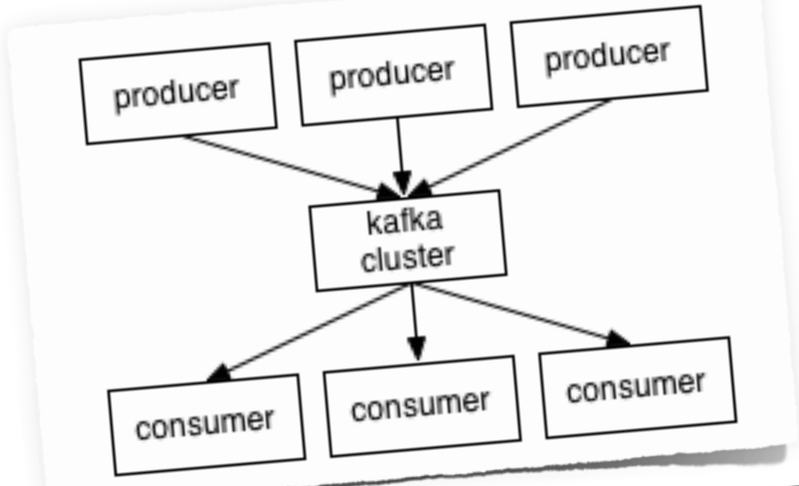
articles

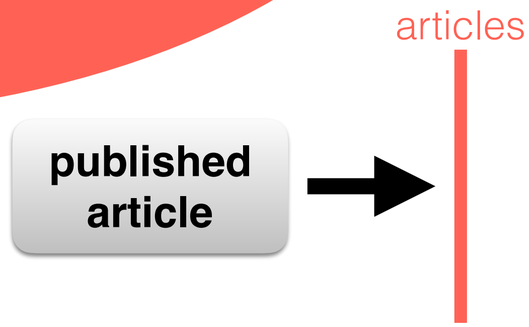
published article

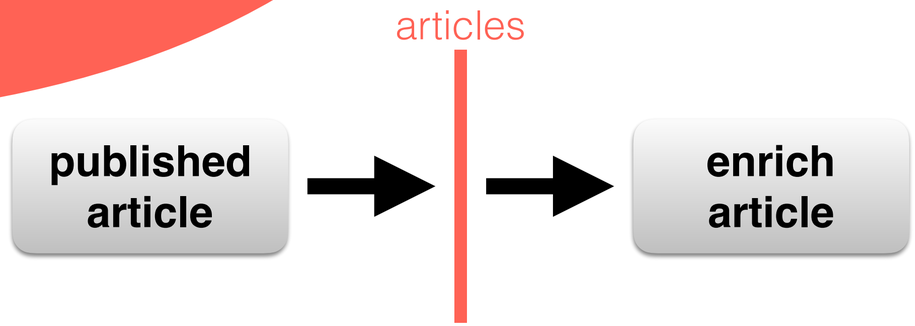
Kafka

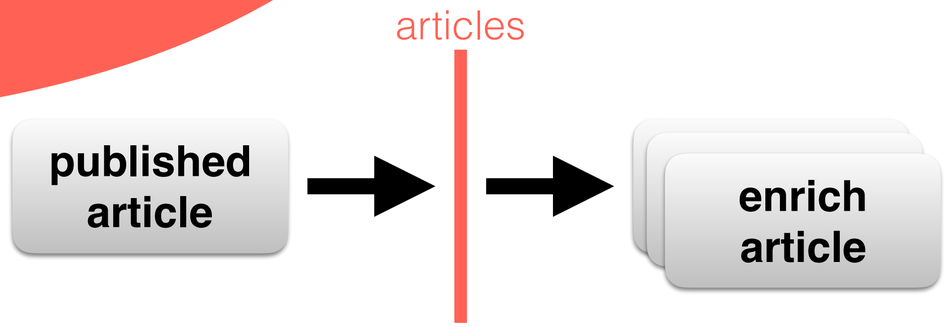
1.1 Introduction

Kafka is a distributed, partitioned, replicated commit log service. It provides the functionality of a messaging system, but with a unique design.




articles
published article


articles
published article
enrich article


articles
published article
enrich article

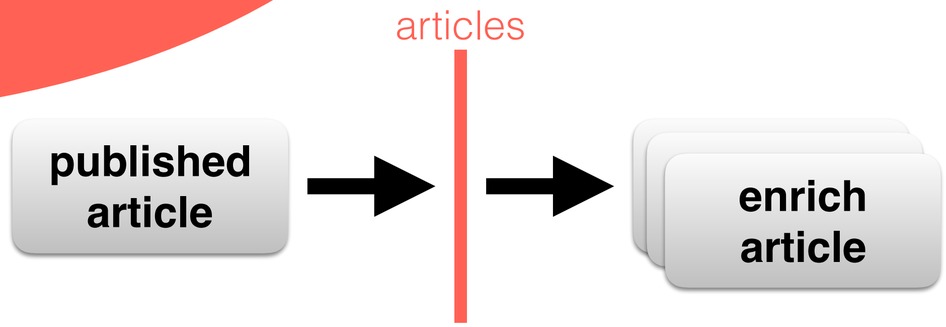
articles

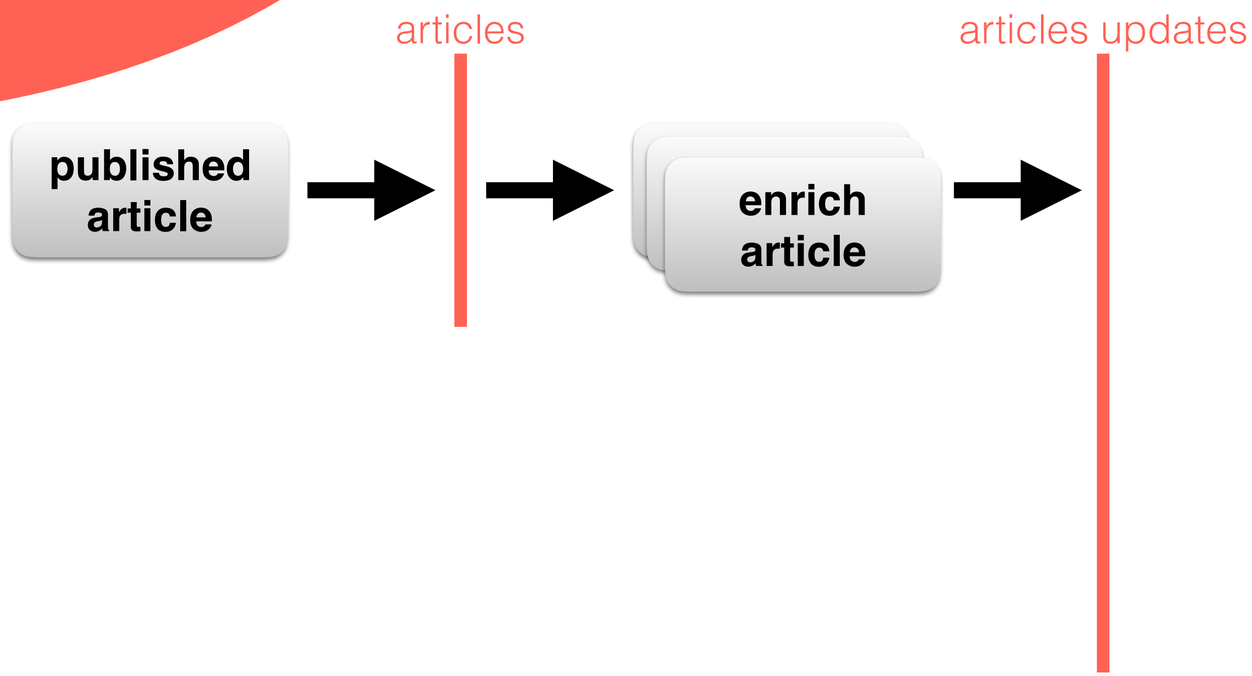
published article

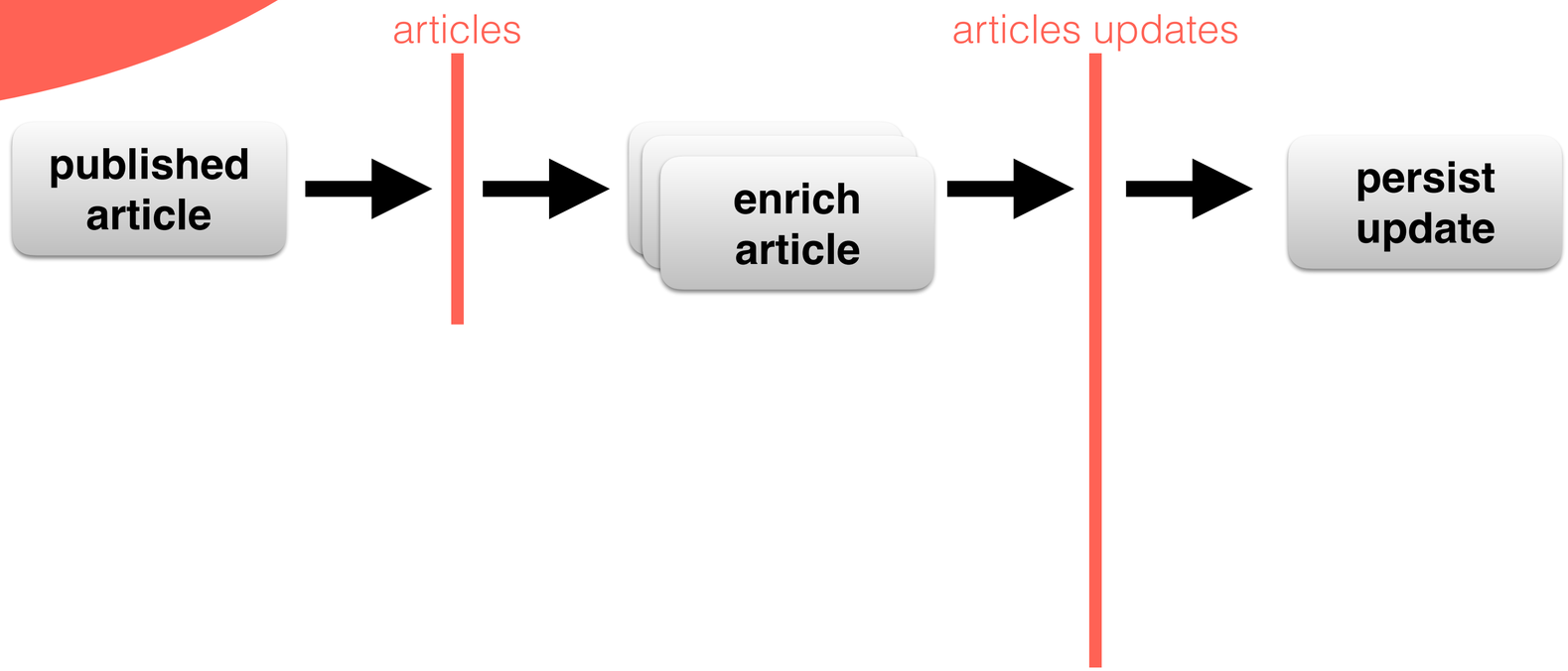
```python
import argparse
import os
import sys

from blendle.pipes.task import Task
from blendle.tasks.enrich.itemprofiles import Pipeline

ENRICH_PIPELINE = os.environ.get('ENRICH_PIPELINE',
                                 'language,tokenizer,postagger,ner,stylometry,author').split(',')


class Enrich(Task):
    def __init__(self, in_file, out_file, error_file, **kwargs):
        self.enricher = kwargs.get('enricher', ENRICH_PIPELINE)
        self.pipeline = Pipeline(self.enricher)
        super().__init__(in_file, out_file, error_file, **kwargs)

    def process(self, item):
        enrichments = self.pipeline(item).get('enrichments', {})
        yield {'id': item['id'],
               'item-metadata': {'enrichments': enrichments}}
```

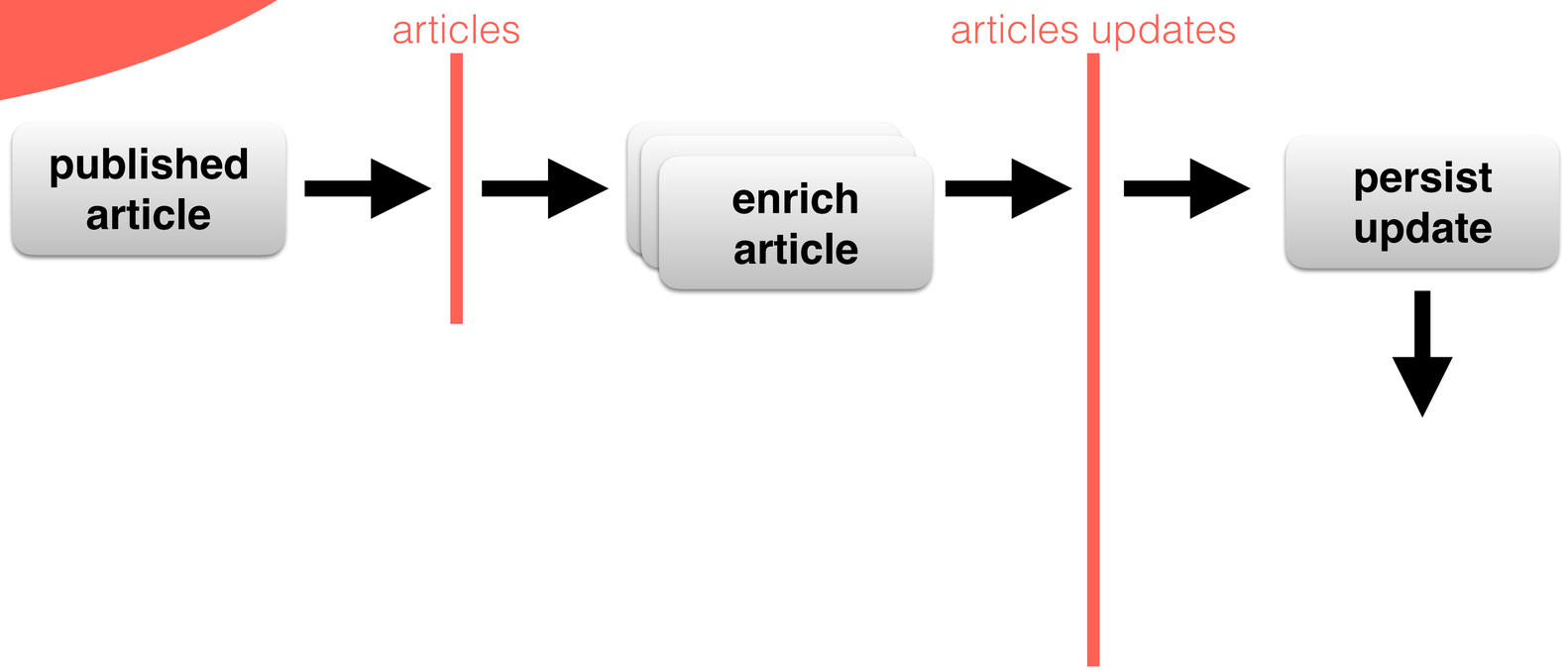
Blendle

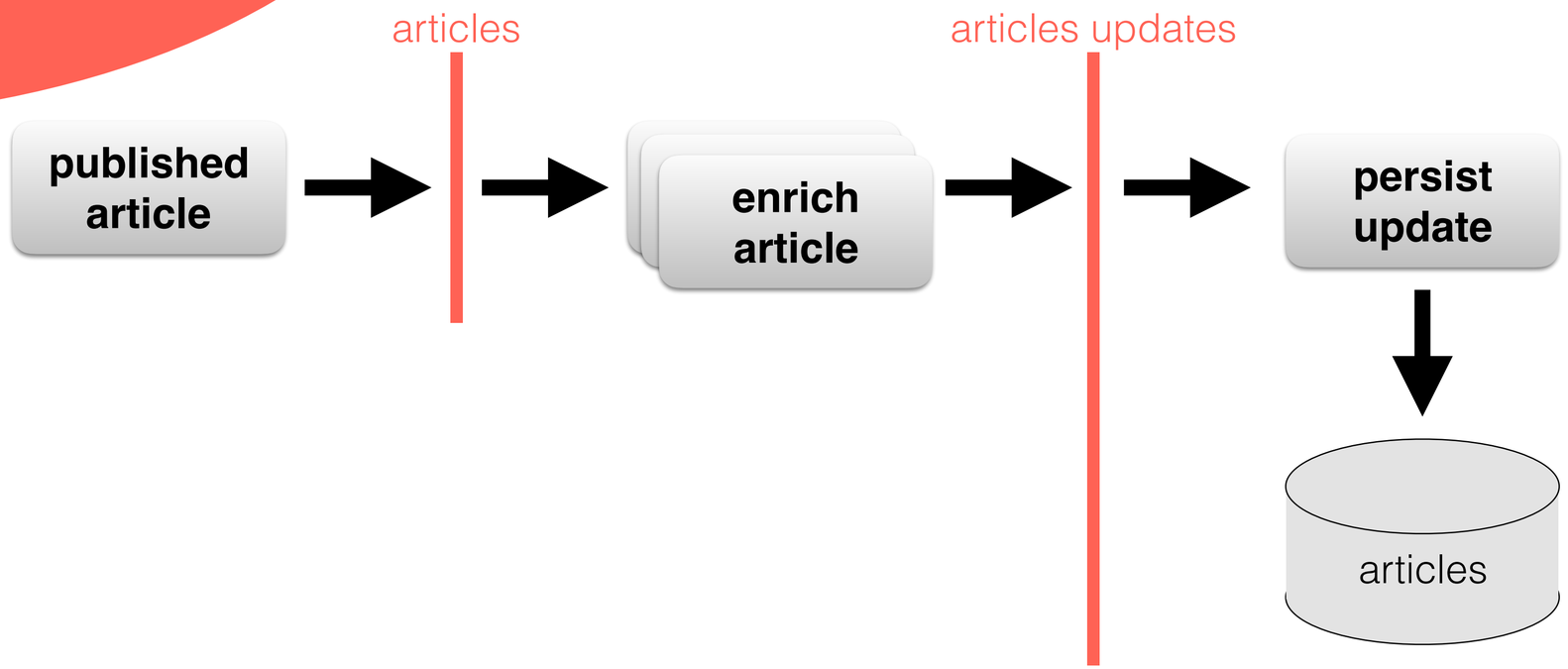
@anneschuth

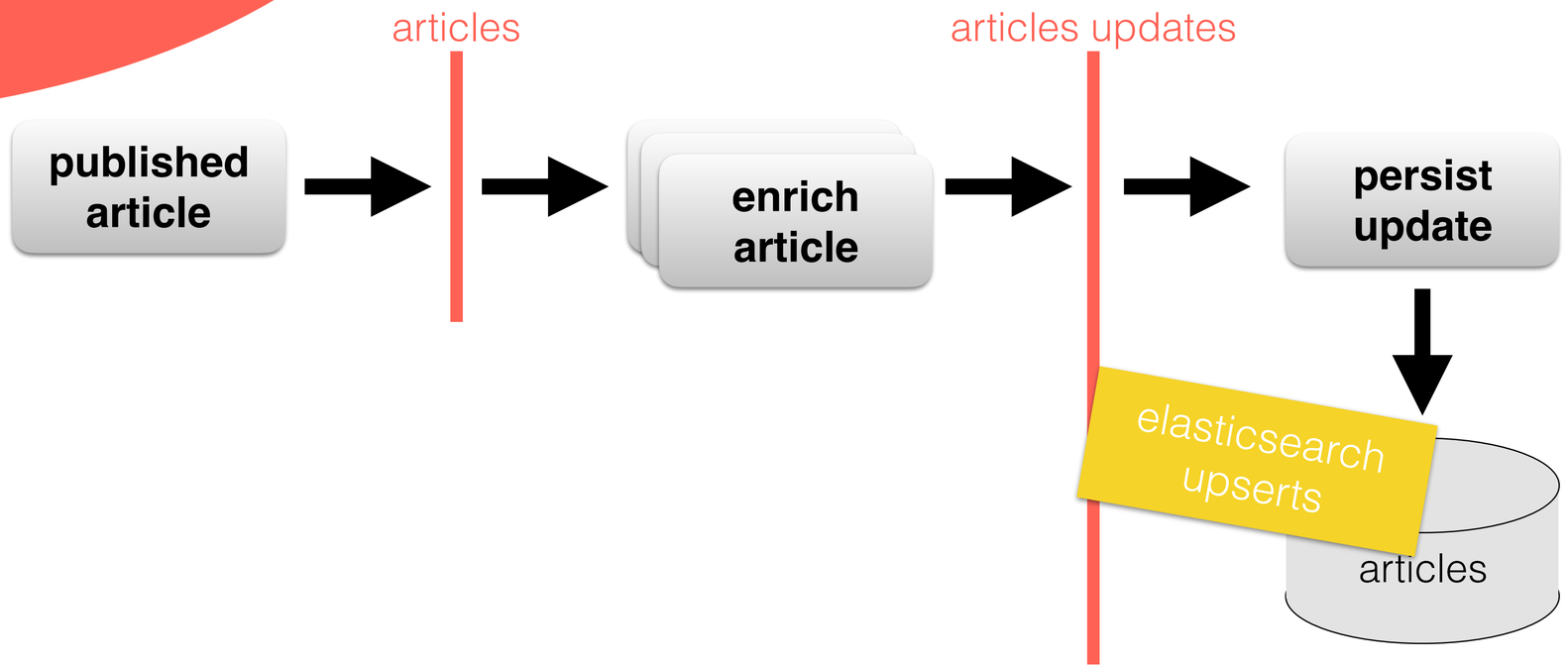

articles
published article
enrich article

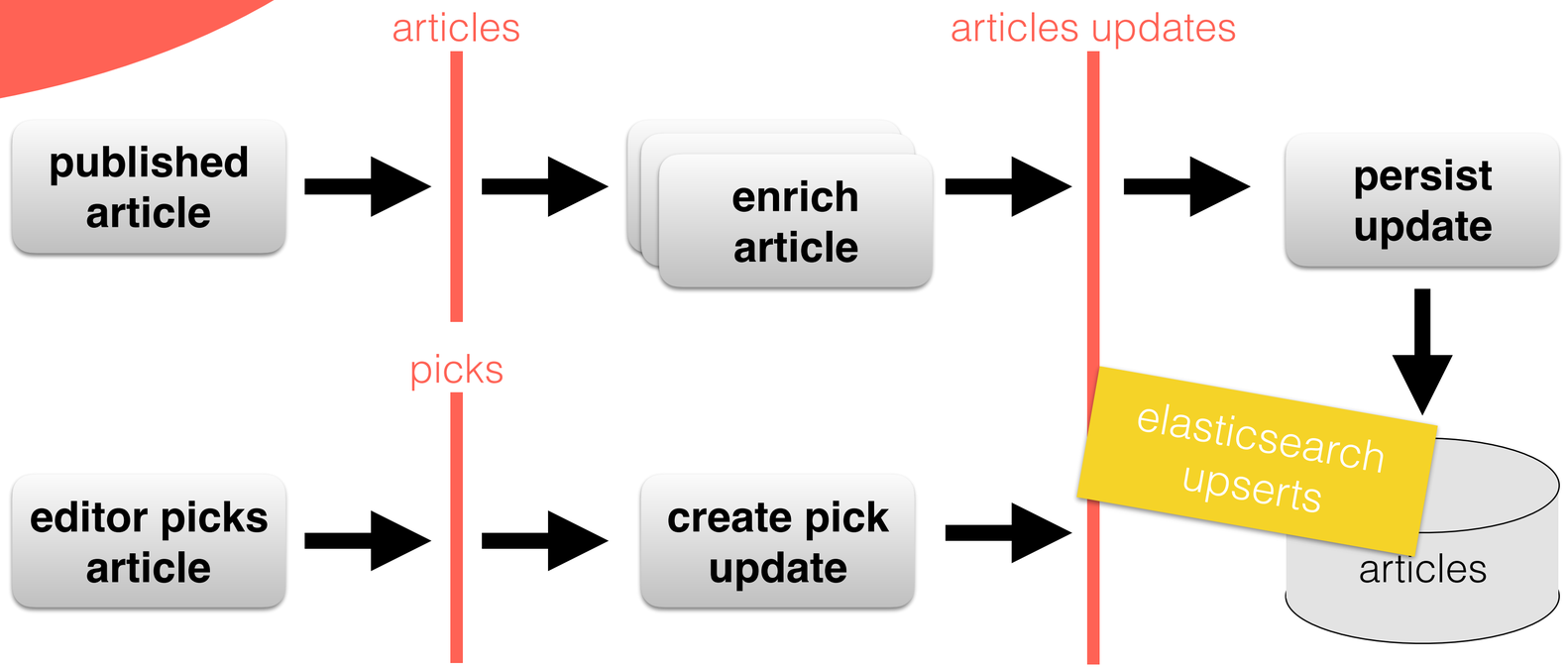

articles
articles updates
published article
enrich article

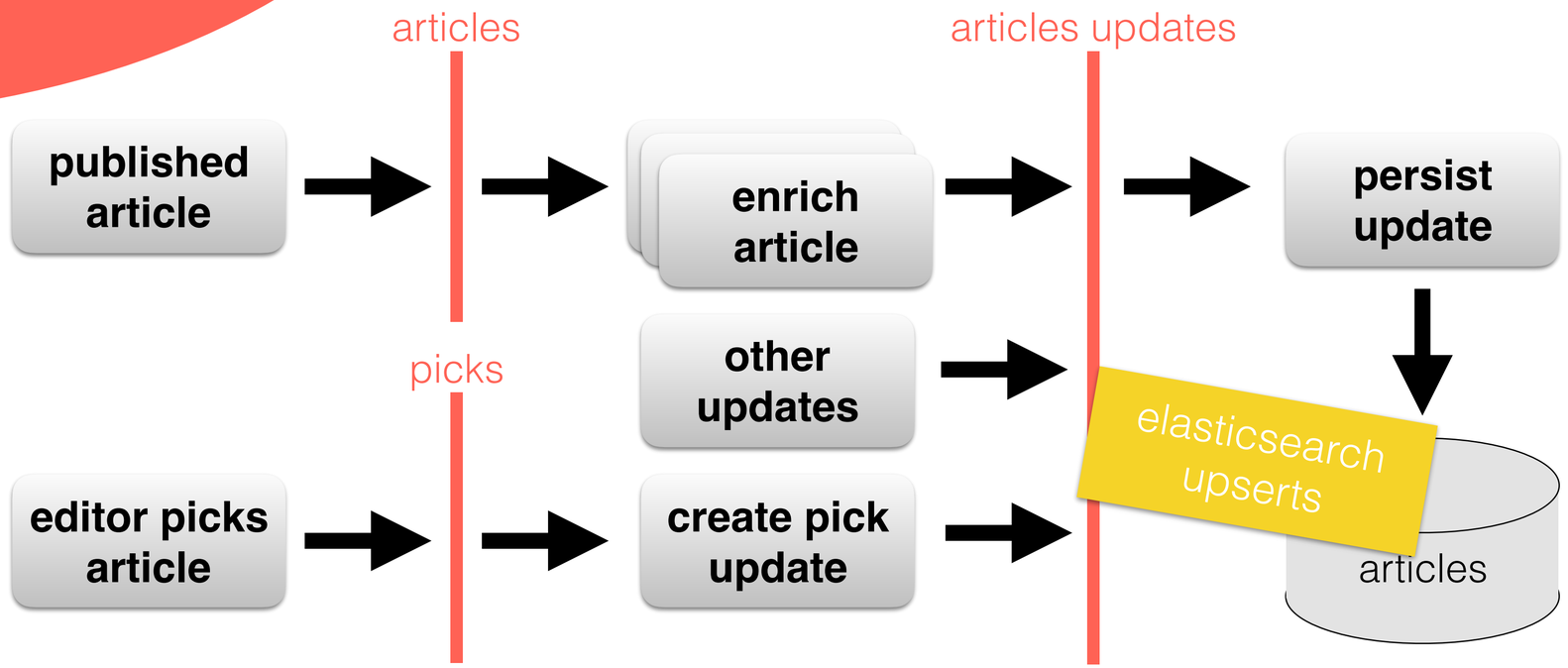

articles
articles updates
published article
enrich article
persist update

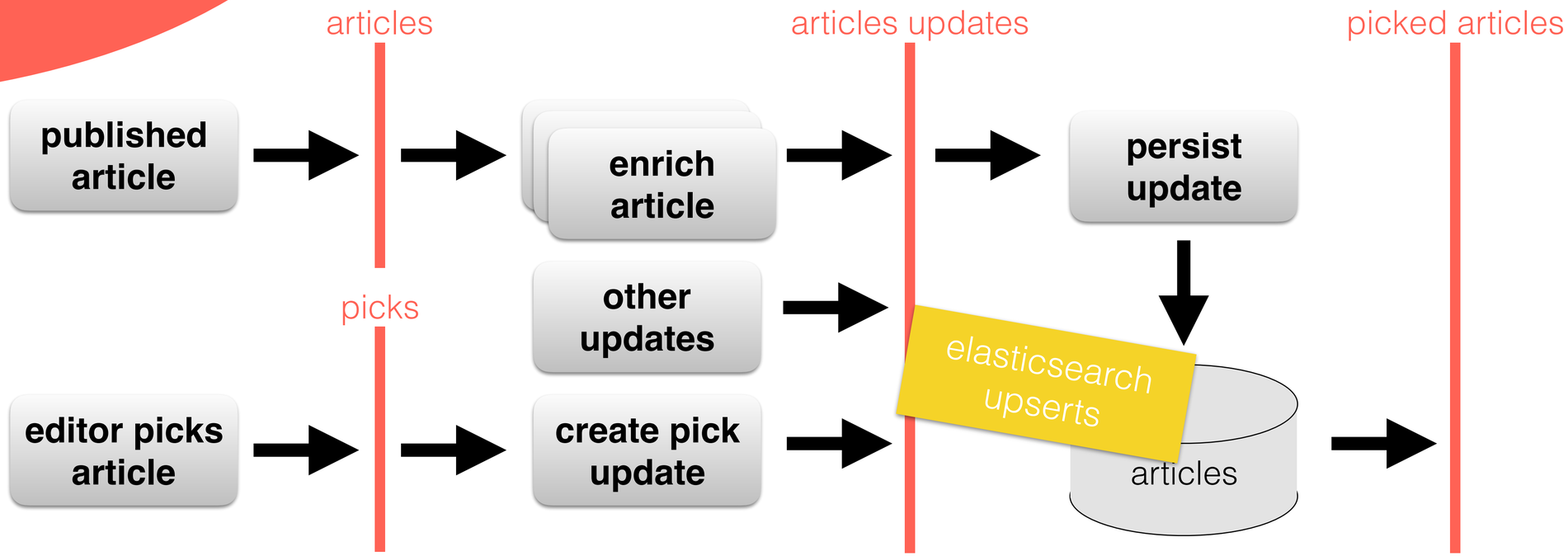

articles
articles updates
published article
enrich article
persist update

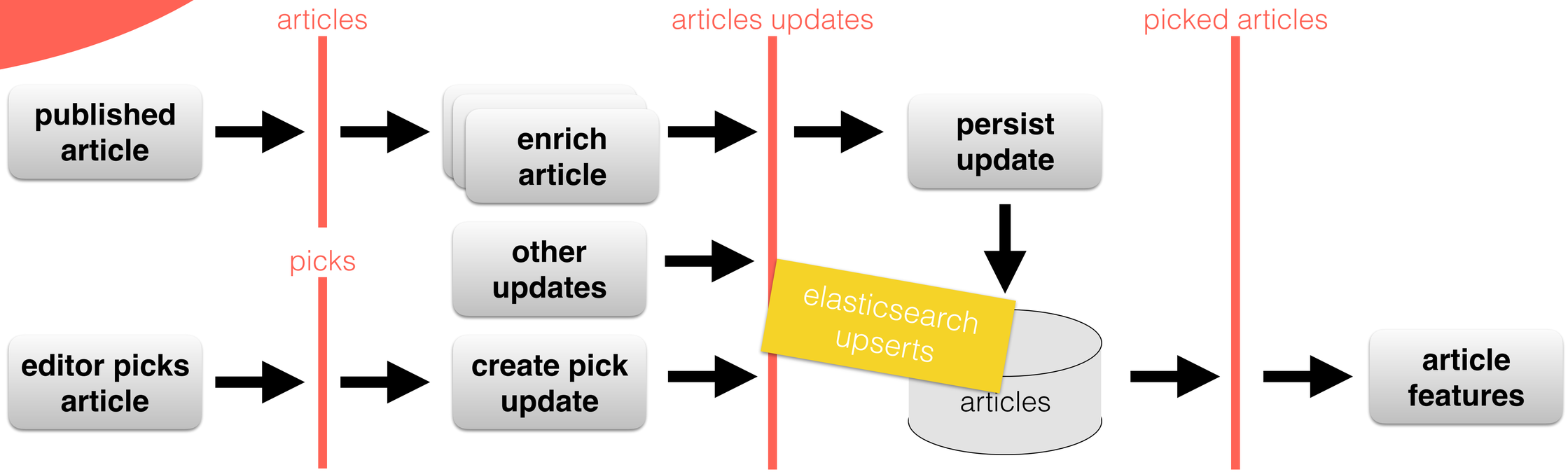

articles
articles updates
published article
enrich article
persist update
articles

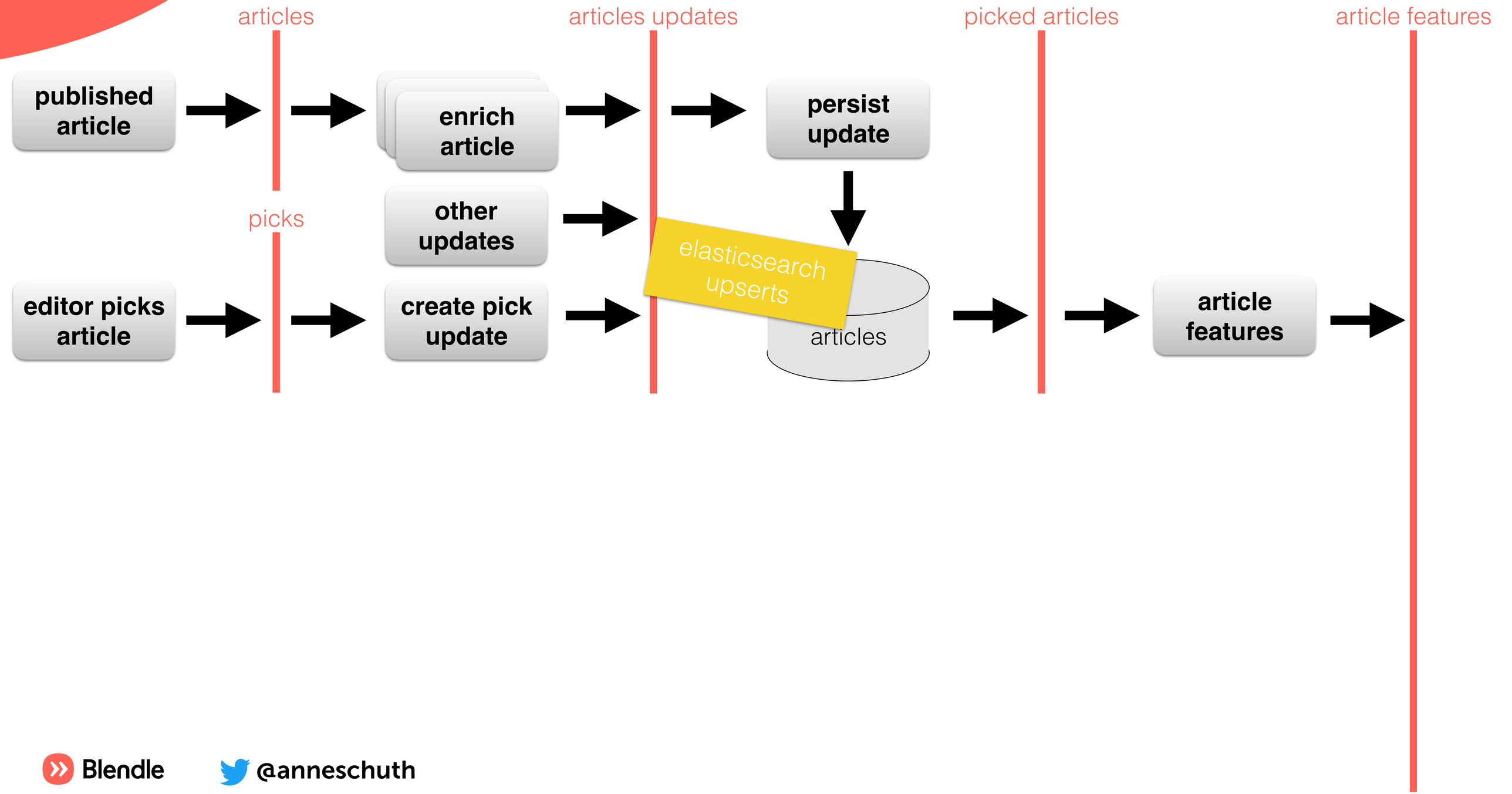

articles
articles updates
published article
enrich article
persist update
elasticsearch upserts
articles


articles
articles updates
published article
enrich article
persist update
picks
elasticsearch upserts
editor picks article
create pick update
articles

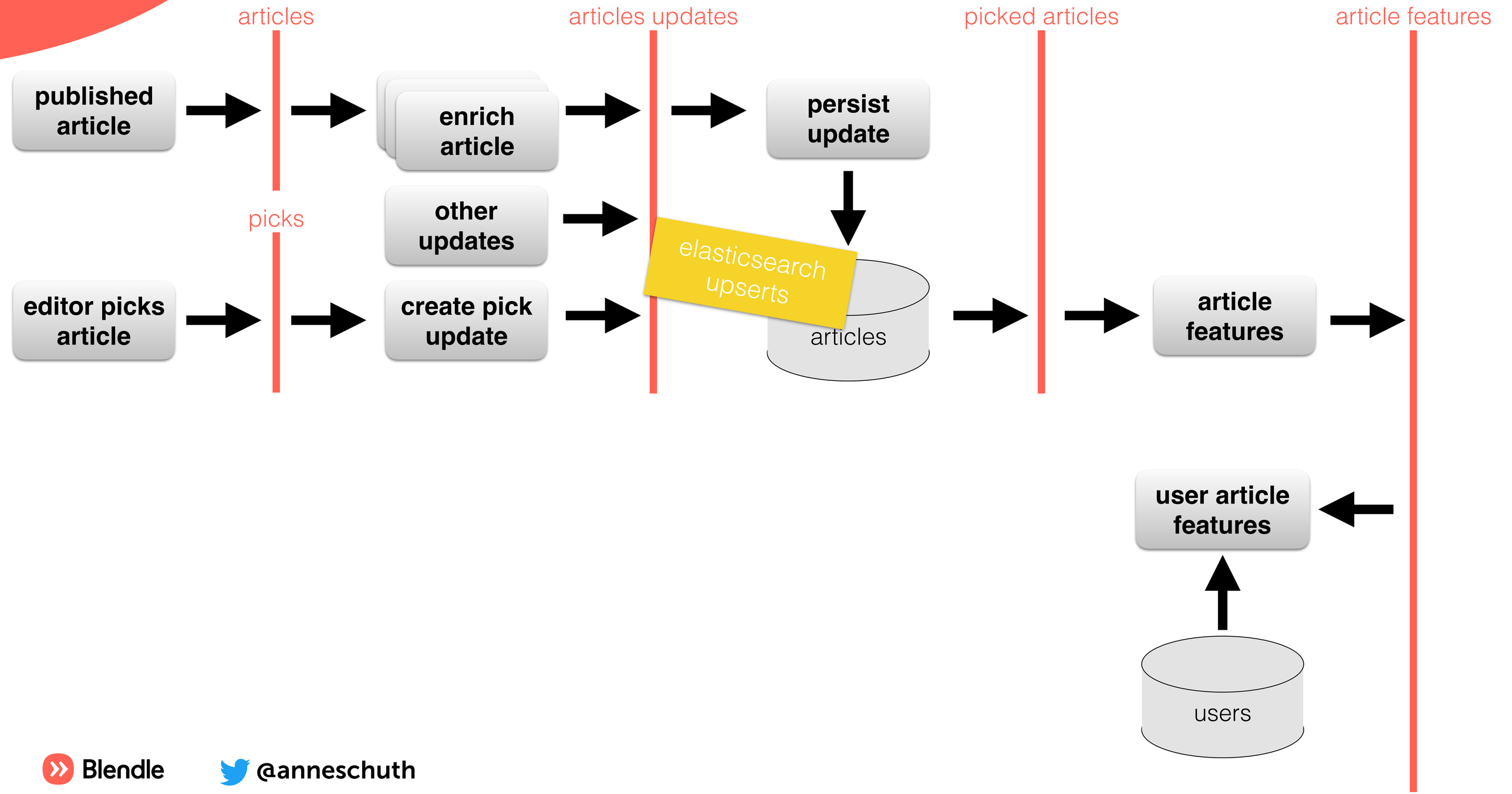

articles
articles updates
published article
enrich article
persist update
picks
other updates
elasticsearch upserts
editor picks article
create pick update
articles


articles
articles updates
picked articles
published article
enrich article
persist update
other updates
picks
elasticsearch upserts
editor picks article
create pick update
articles

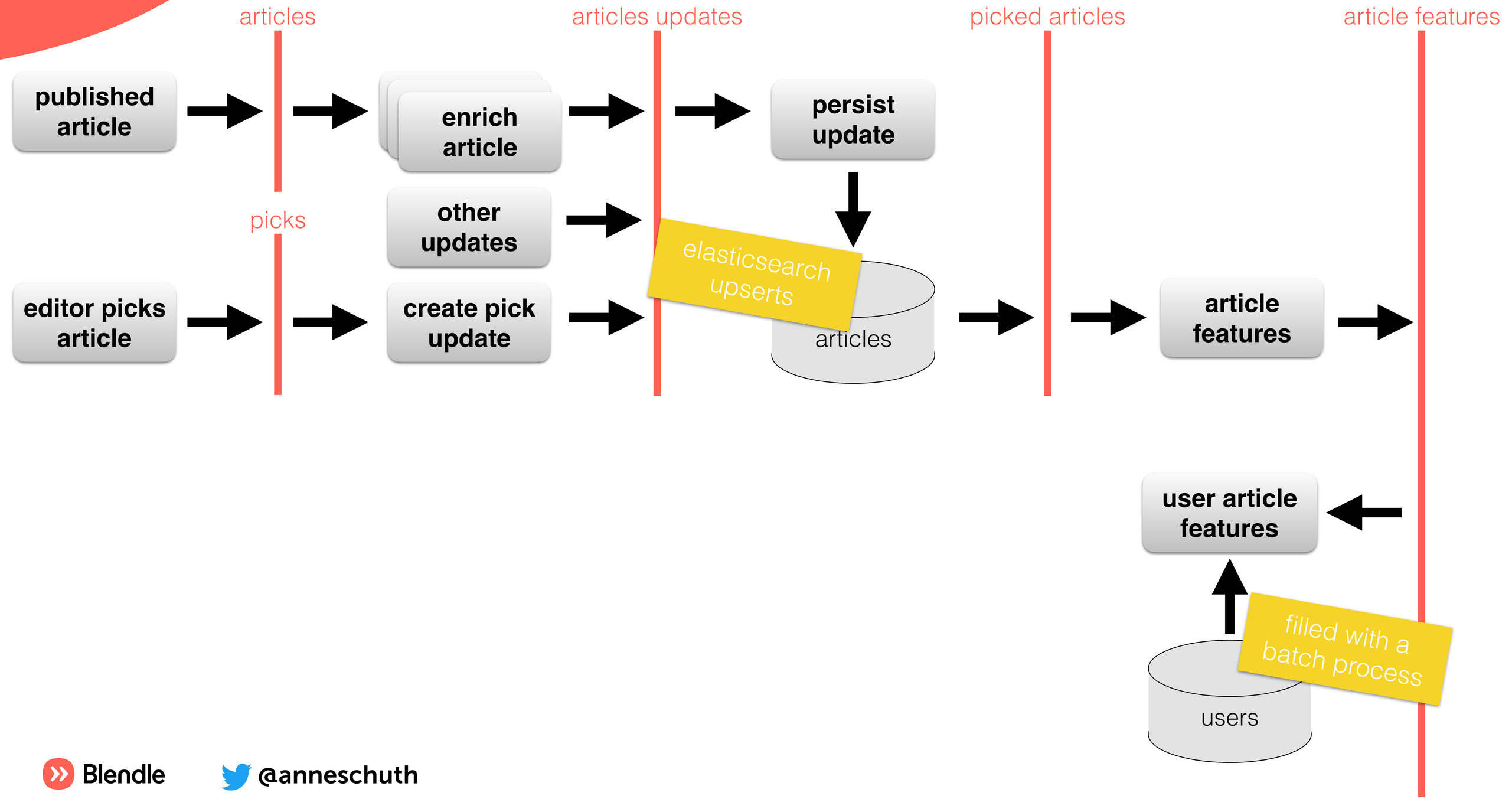

articles
articles updates
picked articles
published article
enrich article
persist update
picks
other updates
elasticsearch upserts
editor picks article
create pick update
articles
article features



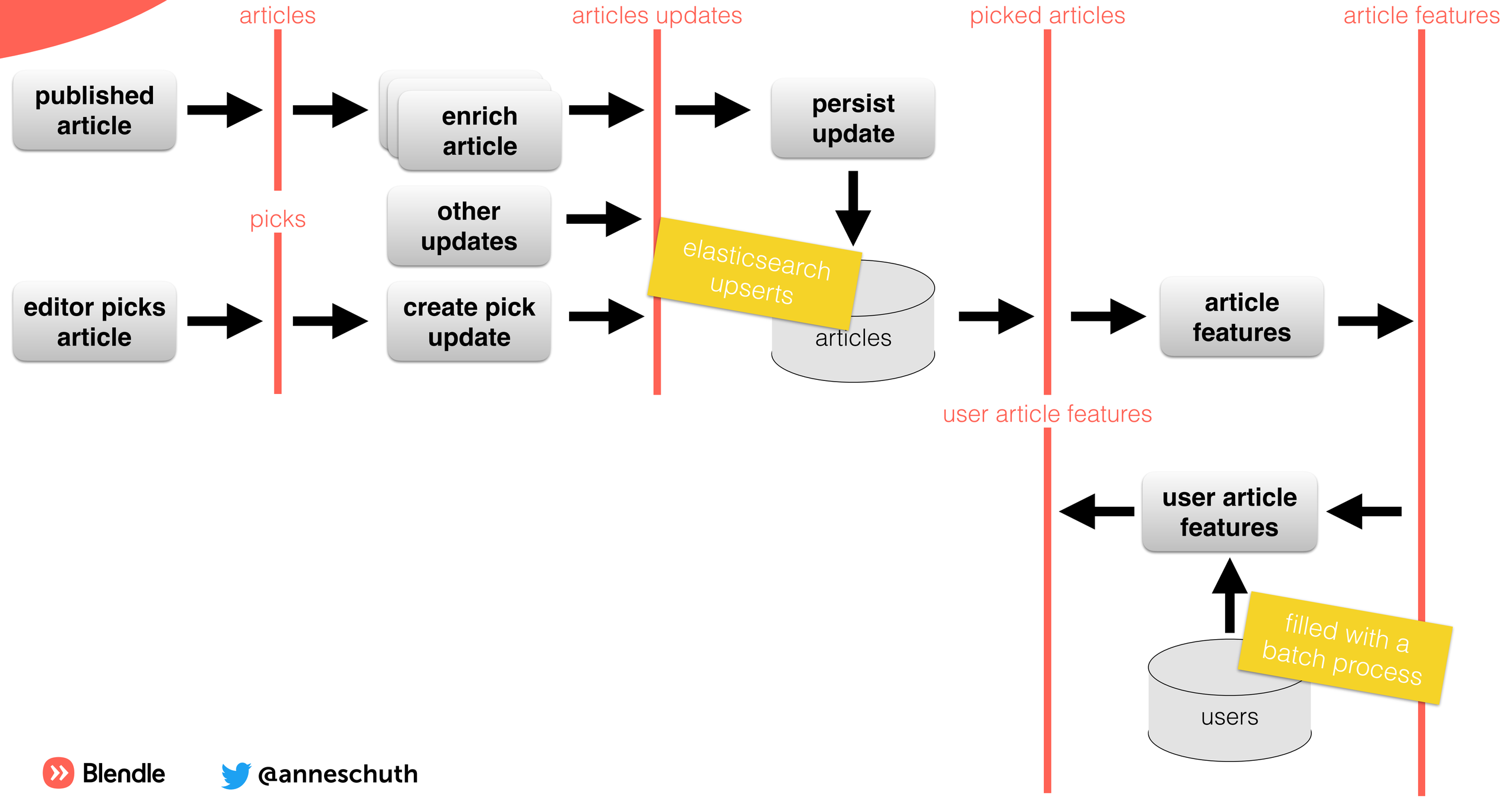
Blendle @anneschuth



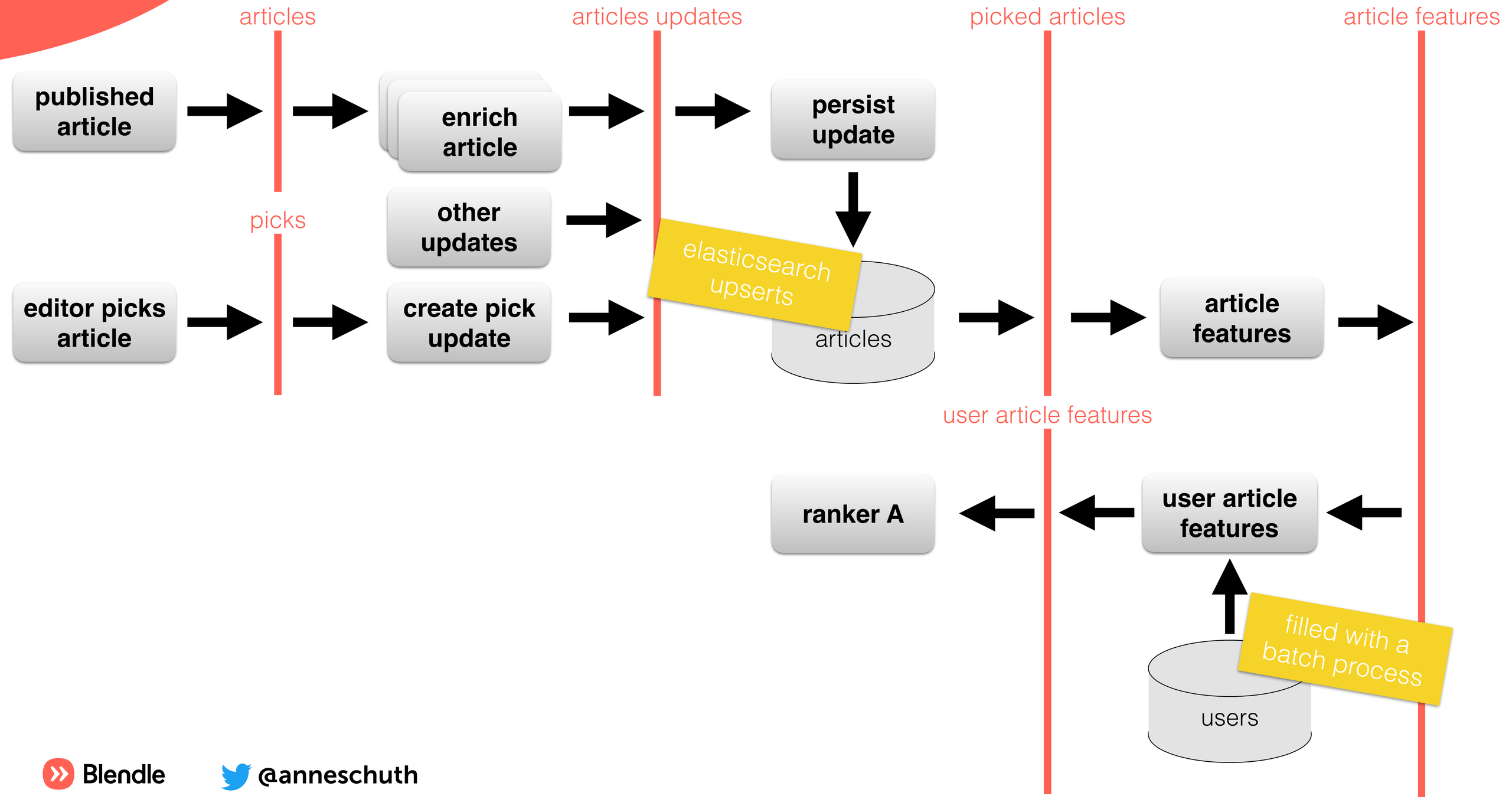
Blendle @anneschuth



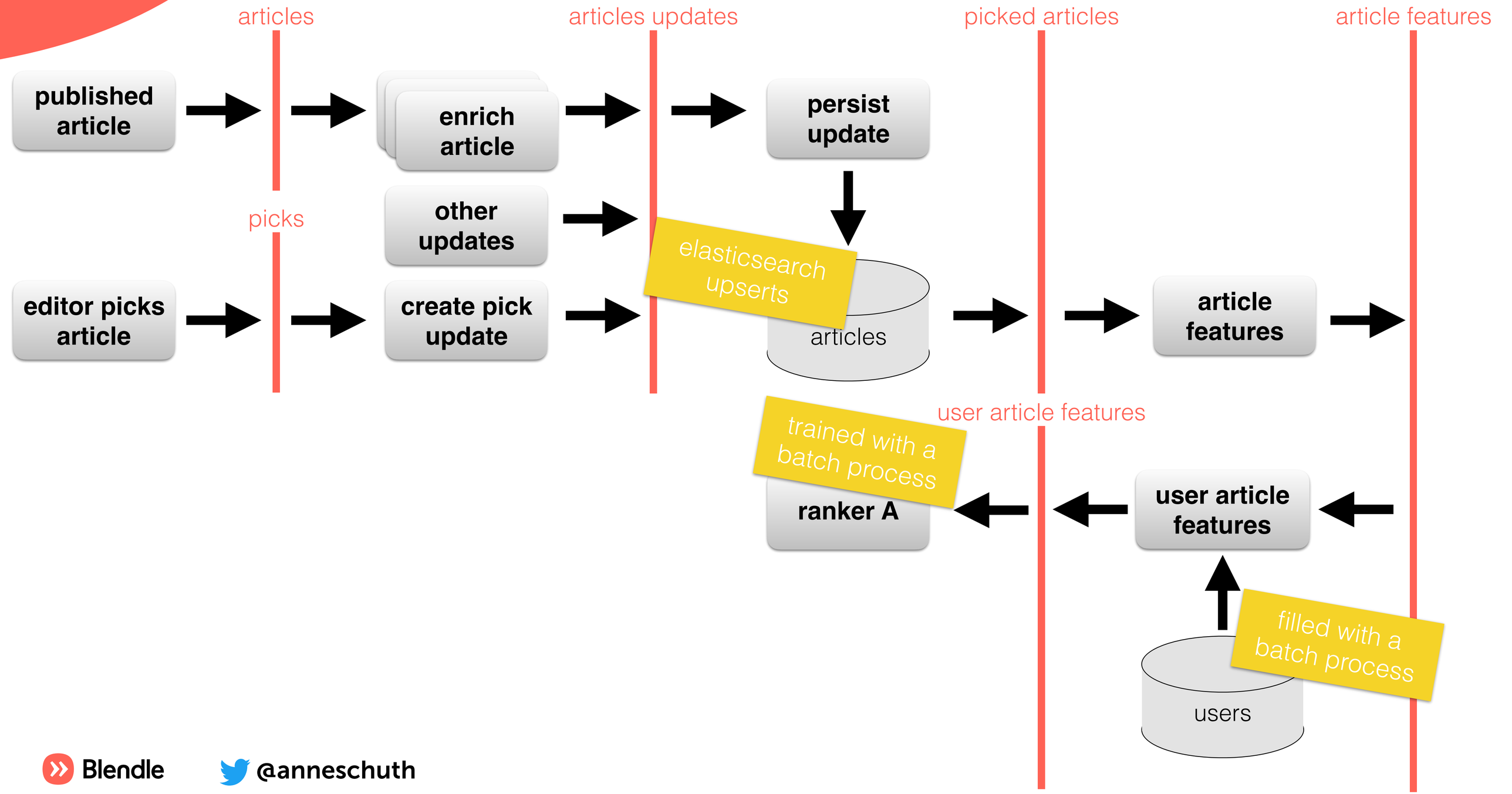
Blendle @anneschuth



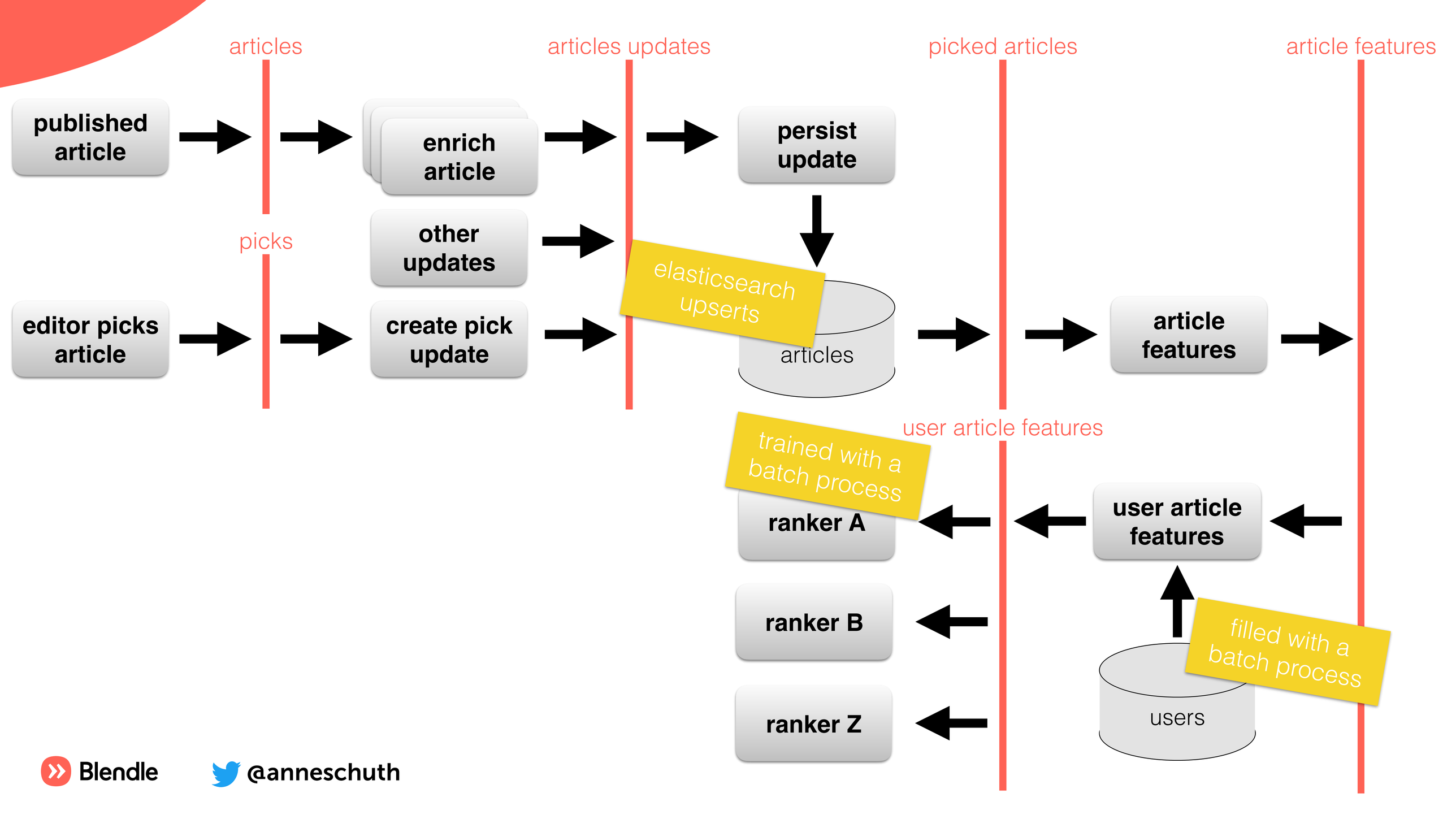
Blendle @anneschuth

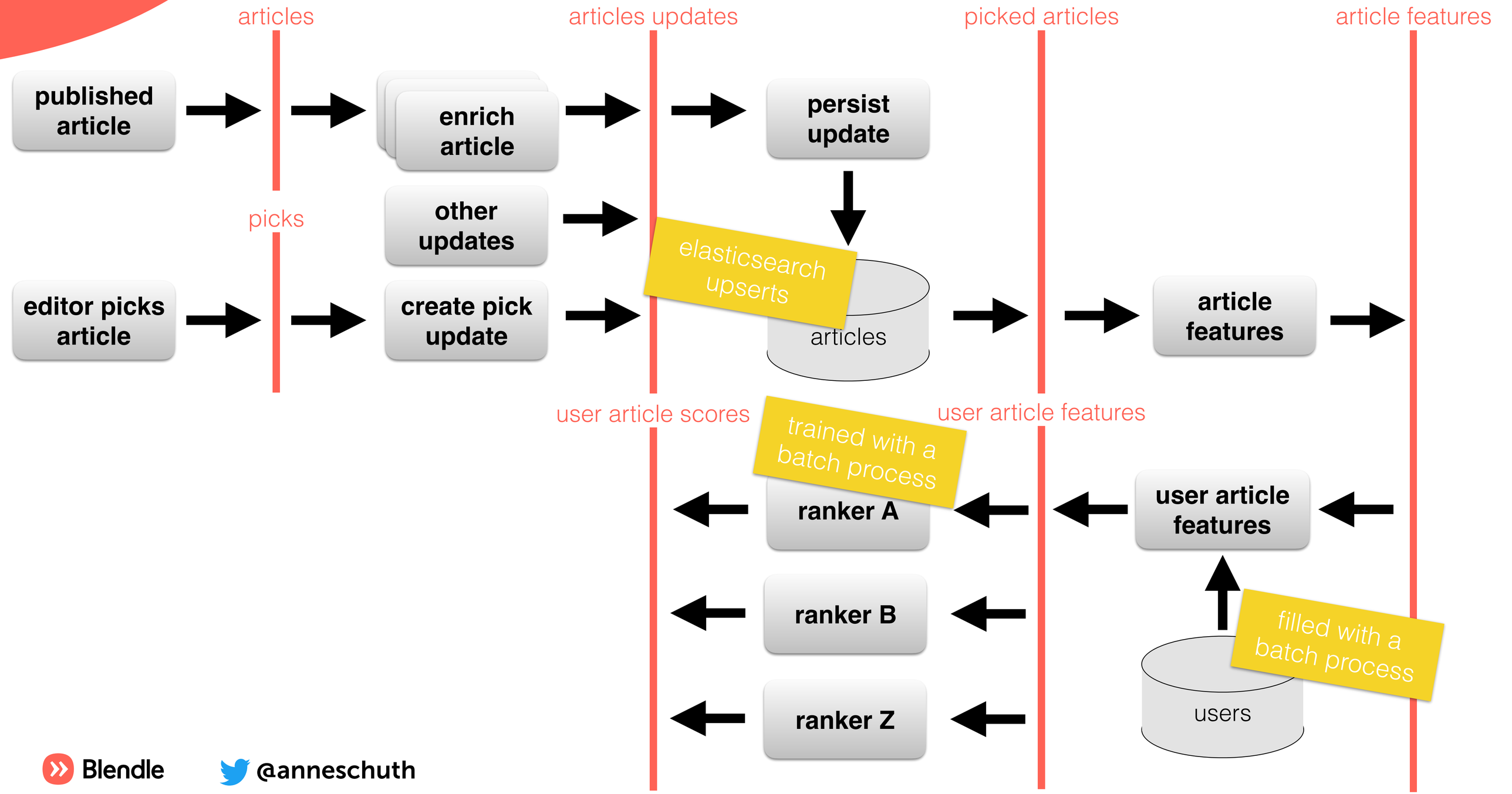


Blendle @anneschuth

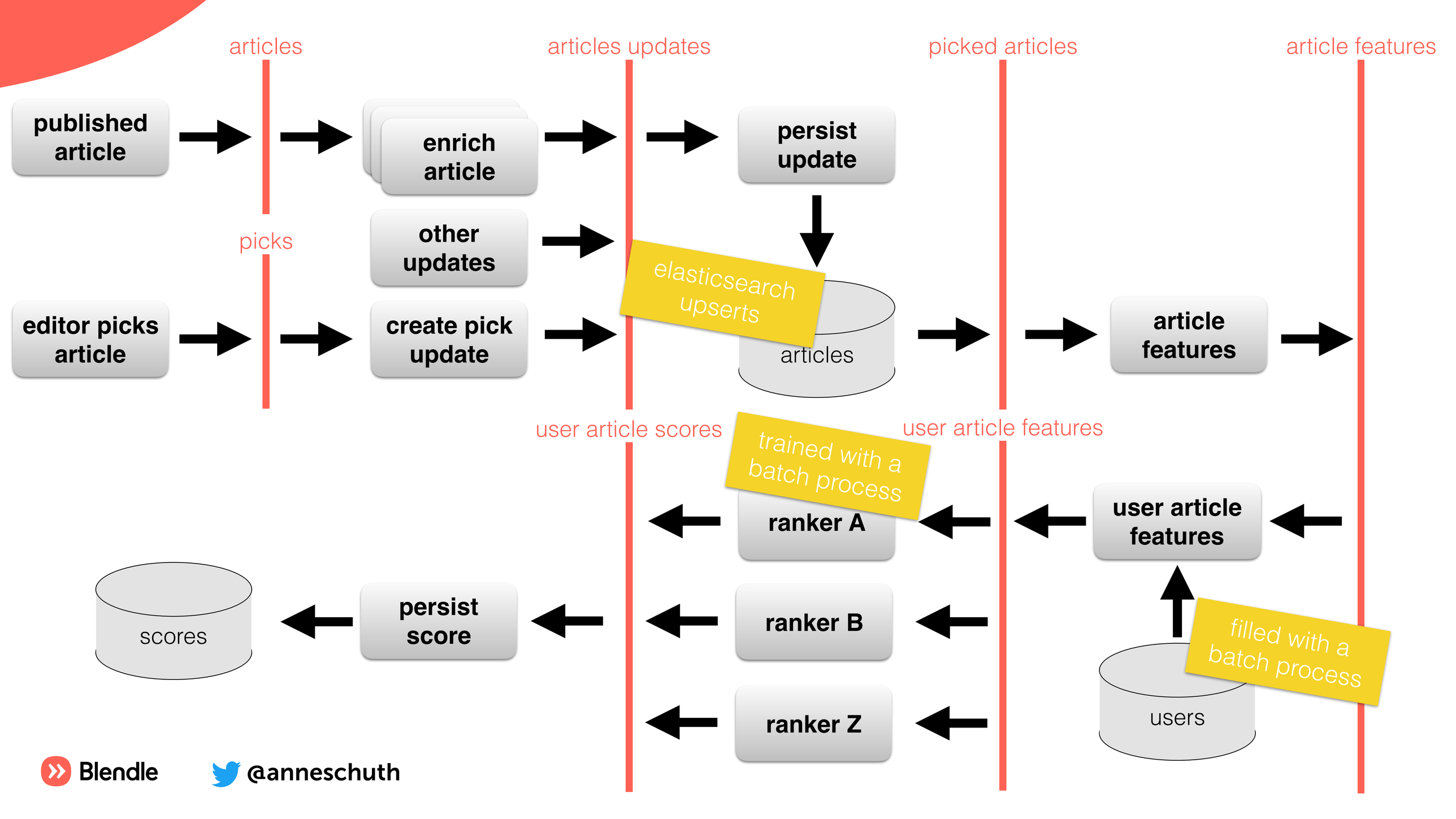


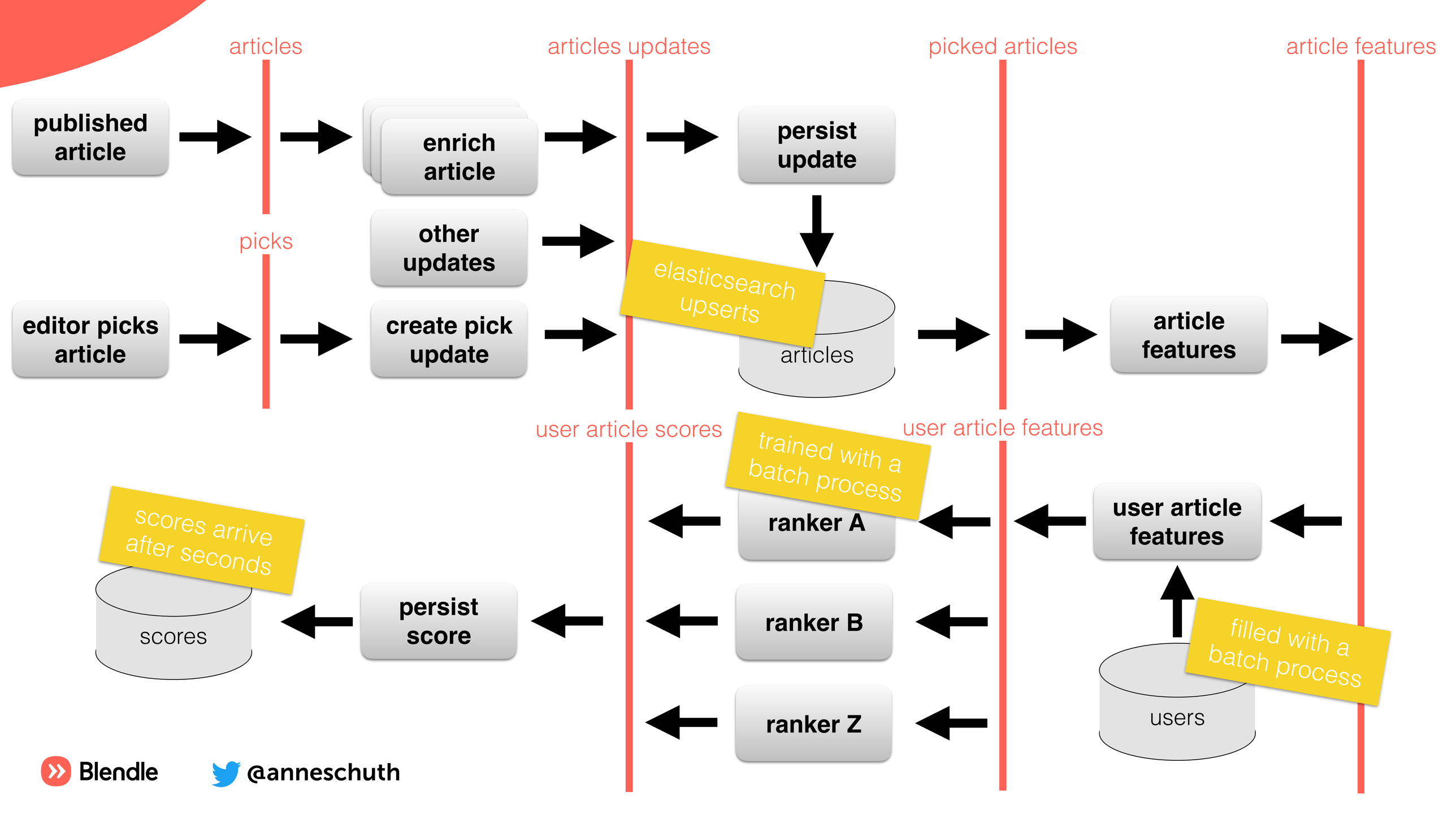
Blendle @anneschuth

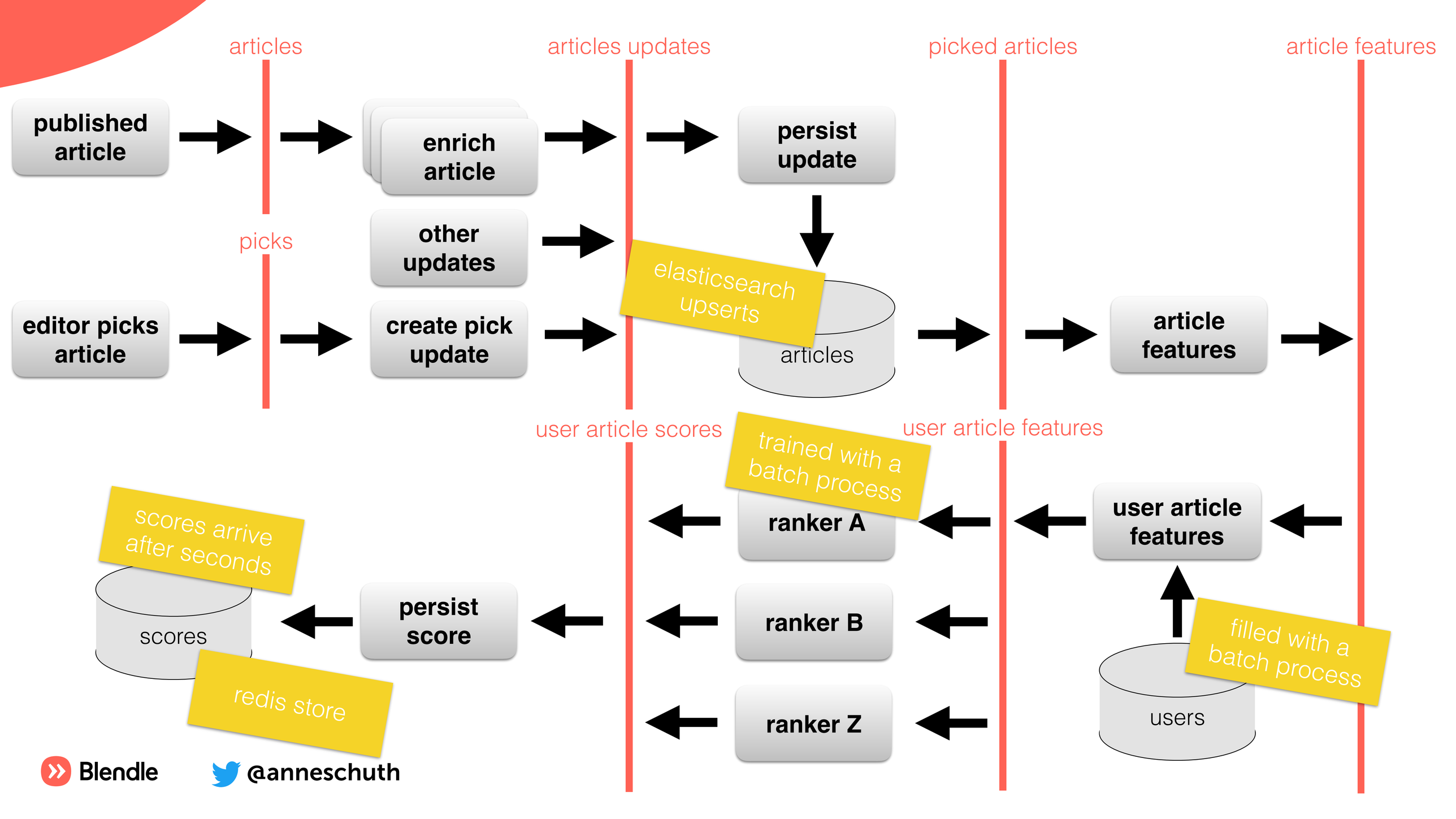

articles
articles updates
picked articles
article features
published article
enrich article
persist update
picks
other updates
editor picks article
create pick update
elasticsearch upserts
articles
article features
user article features
trained with a batch process
ranker A
user article features
ranker B
filled with a batch process
ranker Z
users
Blendle
@anneschuth

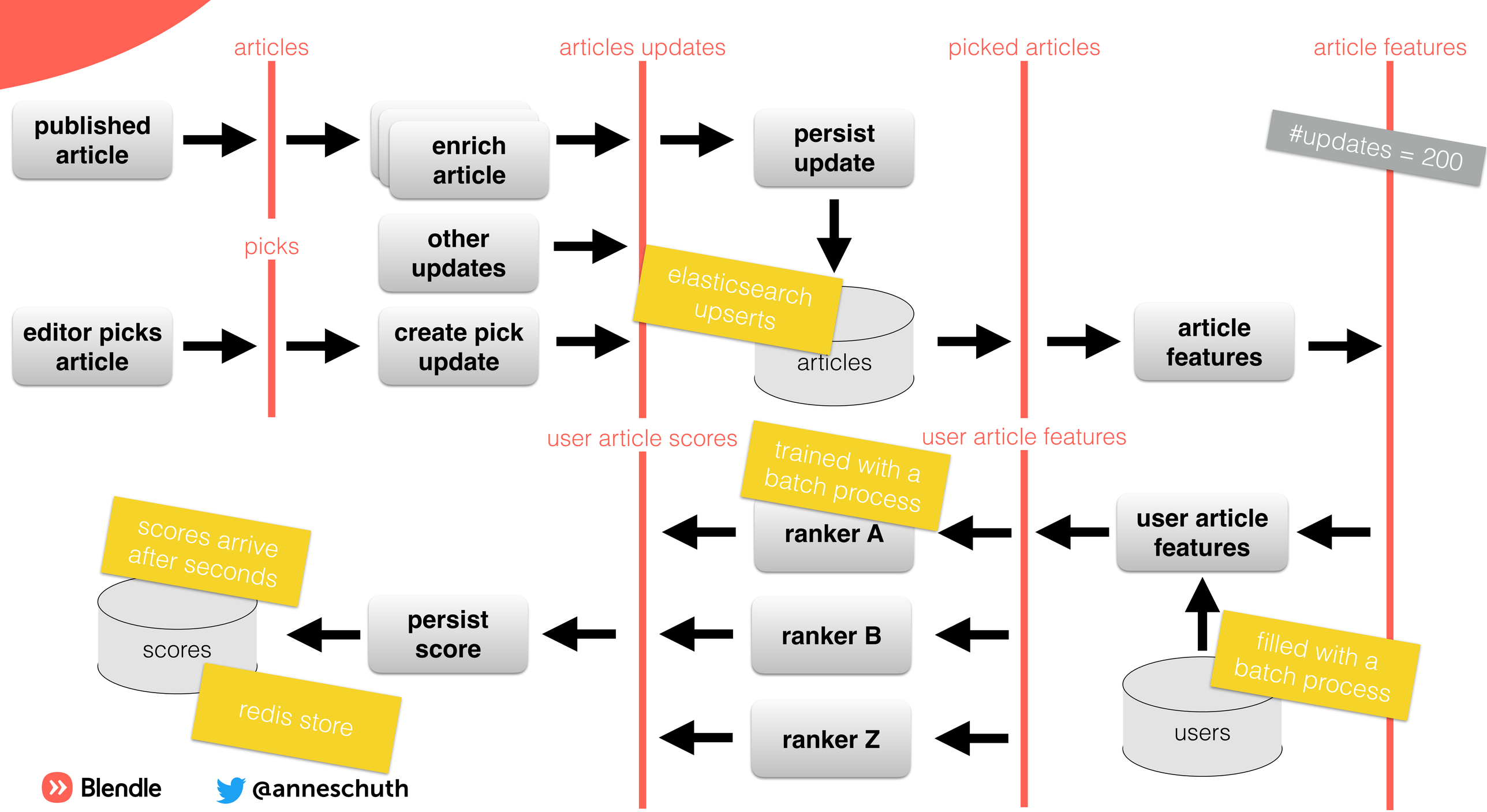


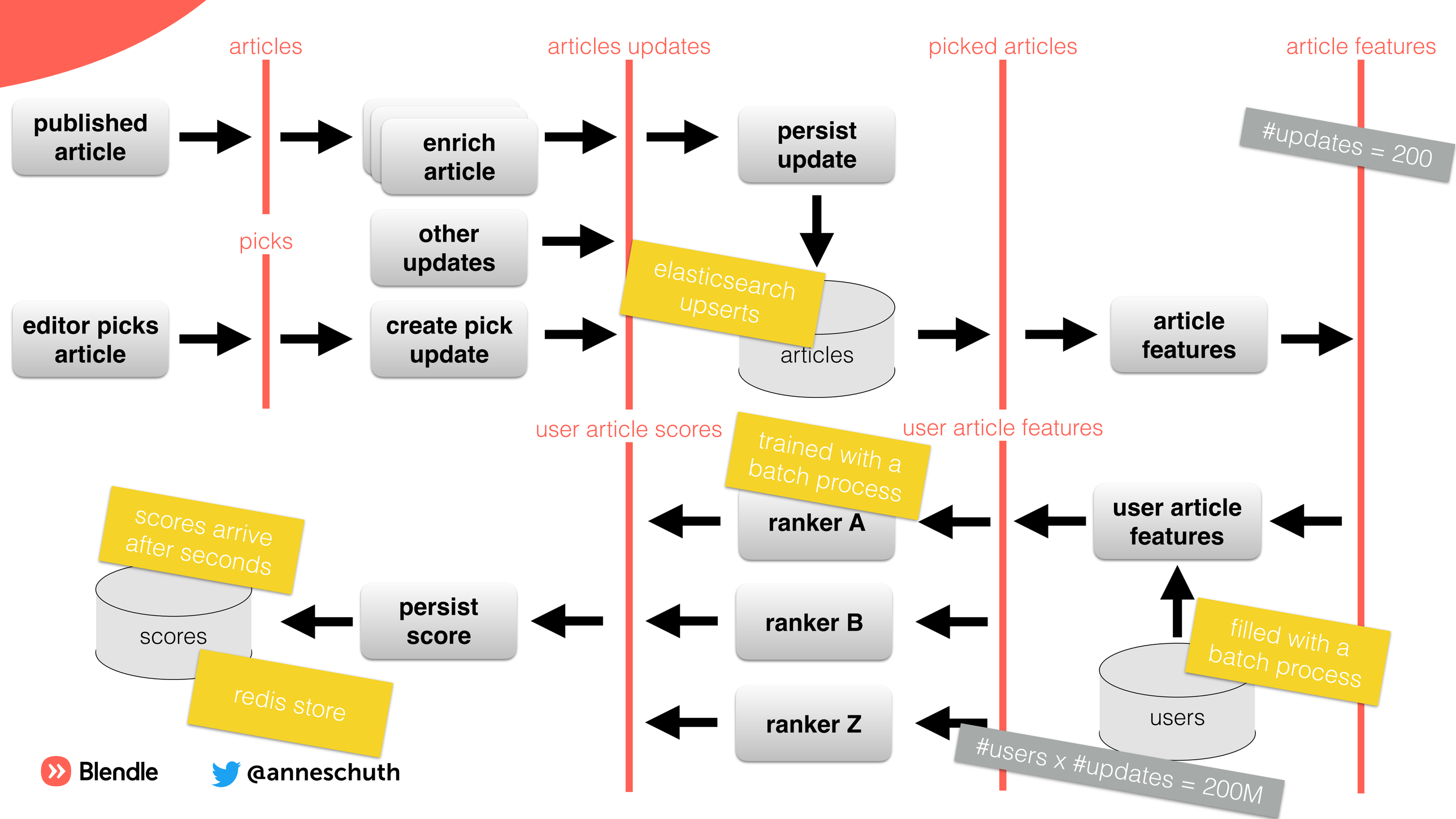
Blendle @anneschuth

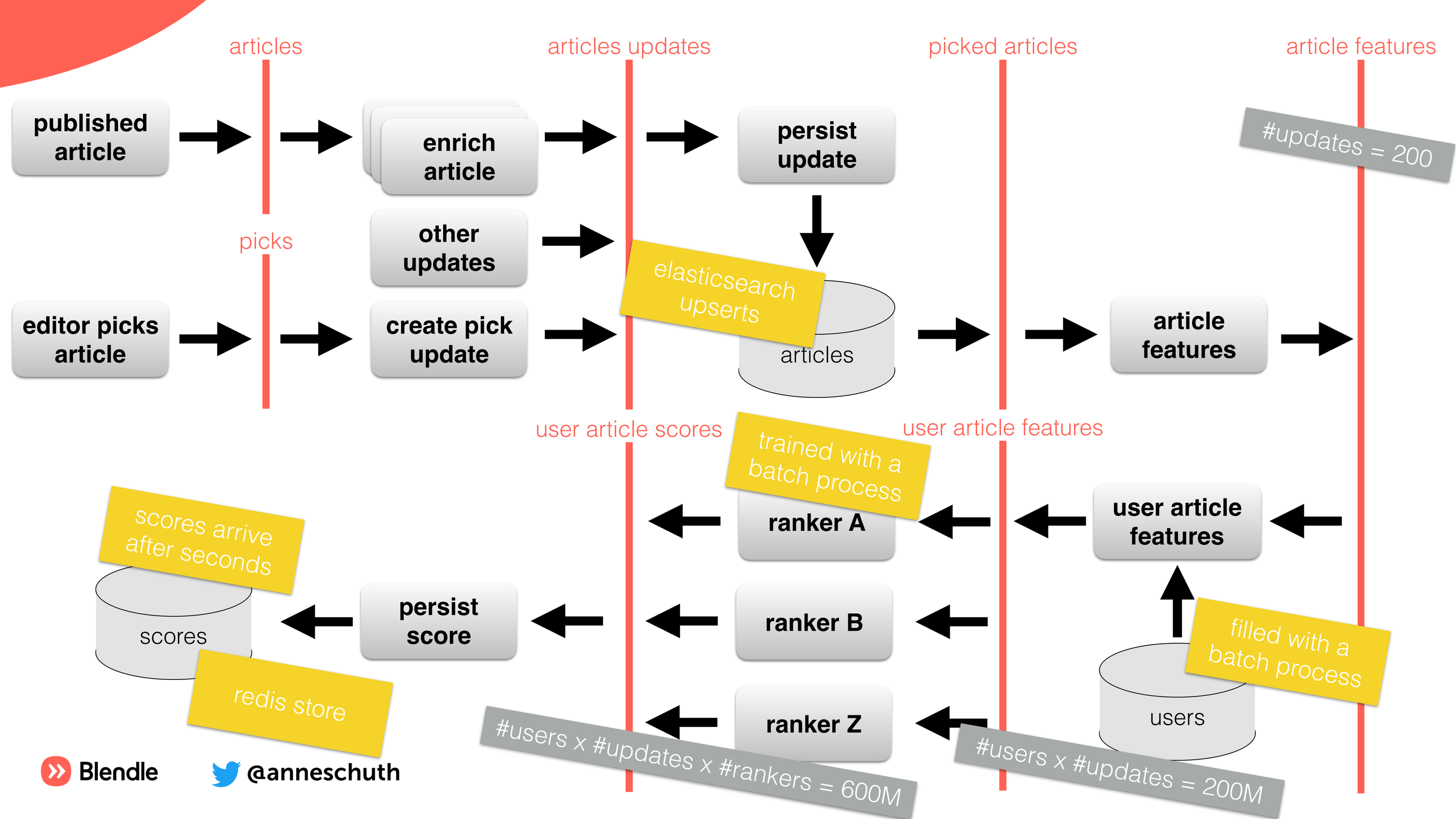

articles
articles updates
picked articles
article features
published article
enrich article
persist update
picks
other updates
editor picks article
create pick update
elasticsearch upserts
articles
article features
user article scores
trained with a batch process
user article features
ranker A
user article features
ranker B
ranker Z
scores
persist score
filled with a batch process
users
Blendle
@anneschuth

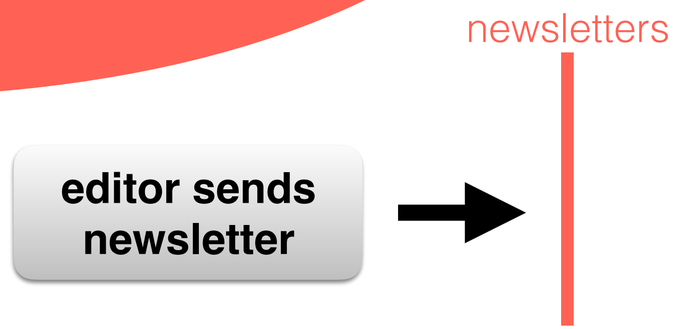

articles
articles updates
picked articles
article features
published article
enrich article
persist update
picks
other updates
editor picks article
create pick update
elasticsearch upserts
articles
article features
user article scores
trained with a batch process
user article features
scores arrive after seconds
ranker A
user article features
persist score
ranker B
scores
filled with a batch process
ranker Z
users
Blendle
@anneschuth

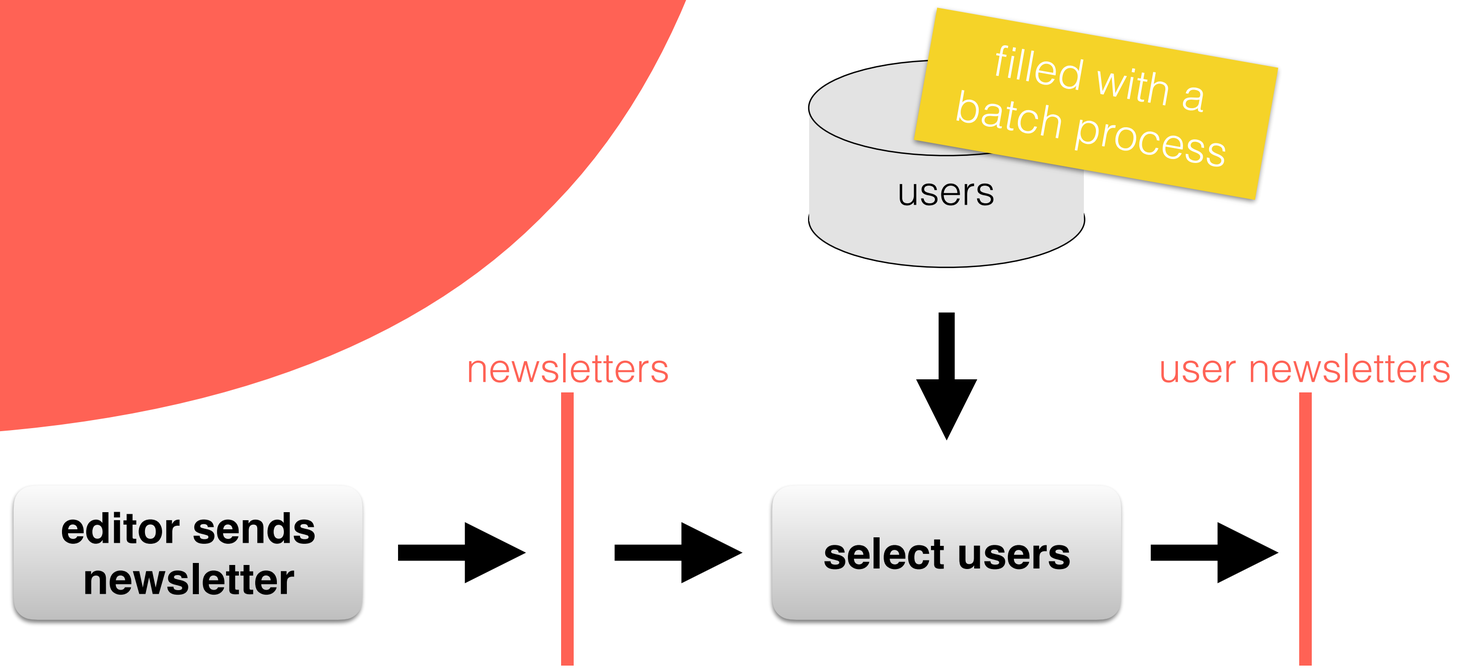

articles
articles updates
picked articles
article features
published article
enrich article
persist update
picks
other updates
editor picks article
create pick update
elasticsearch upserts
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
persist score
ranker B
filled with a batch process
scores
ranker Z
users
redis store
Blendle
@anneschuth

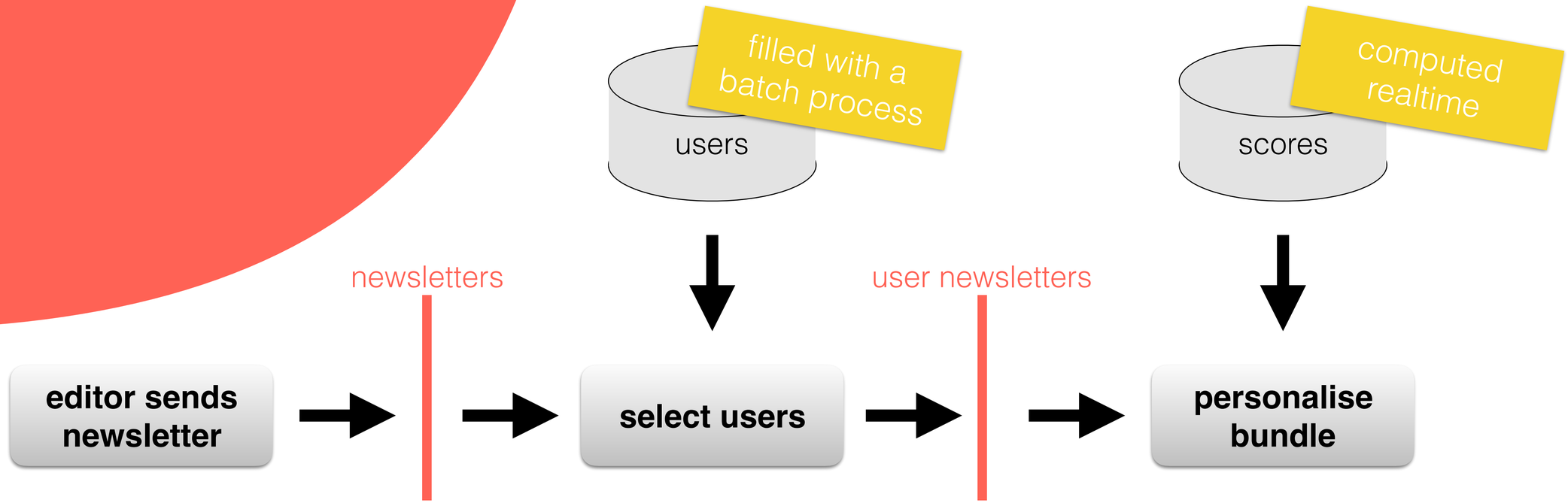

articles
articles updates
picked articles
article features
published article
enrich article
persist update
#updates = 200
picks
other updates
elasticsearch upserts
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
persist score
ranker B
filled with a batch process
scores
ranker Z
users
redis store
Blendle
@anneschuth

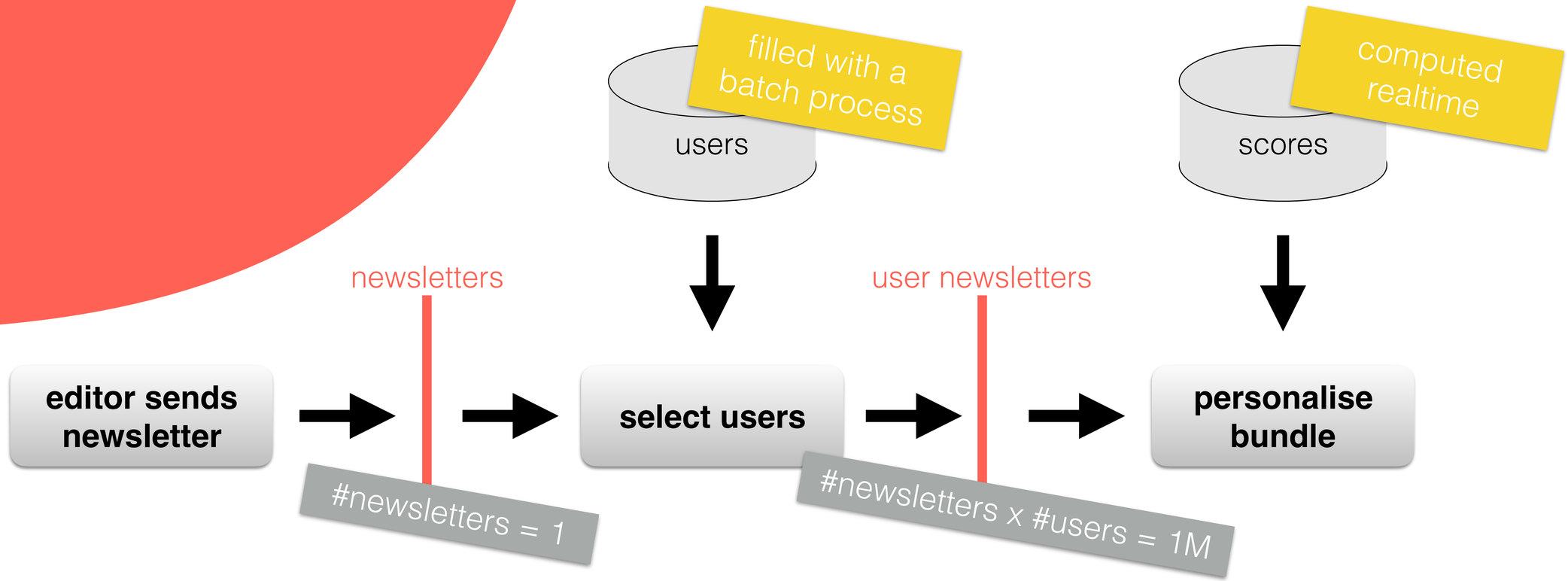

articles
articles updates
picked articles
article features
published article
enrich article
persist update
#updates = 200
picks
other updates
elasticsearch upserts
editor picks article
create pick update
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
persist score
ranker B
filled with a batch process
scores
ranker Z
users
redis store
#users x #updates = 200M
Blendle
@anneschuth


articles
articles updates
picked articles
article features
published article
enrich article
persist update
#updates = 200
picks
other updates
editor picks article
create pick update
elasticsearch upserts
articles
article features
user article scores
user article features
trained with a batch process
ranker A
user article features
scores arrive after seconds
scores
persist score
ranker B
filled with a batch process
redis store
ranker Z
users
#users x #updates x #rankers = 600M
#users x #updates = 200M
Blendle
@anneschuth


newsletters
editor sends newsletter


filled with a batch process
users
newsletters
user newsletters
editor sends newsletter
select users


users
filled with a batch process
scores
computed realtime
newsletters
user newsletters
editor sends newsletter
select users
personalise bundle


users
filled with a batch process
scores
computed realtime
newsletters
user newsletters
editor sends newsletter
select users
personalise bundle
#newsletters = 1
#newsletters x #users = 1M

Bundler

editor sends newsletter

Blendle

@annes

Bundler

Filter

editor sends newsletter

Bundler

Filter

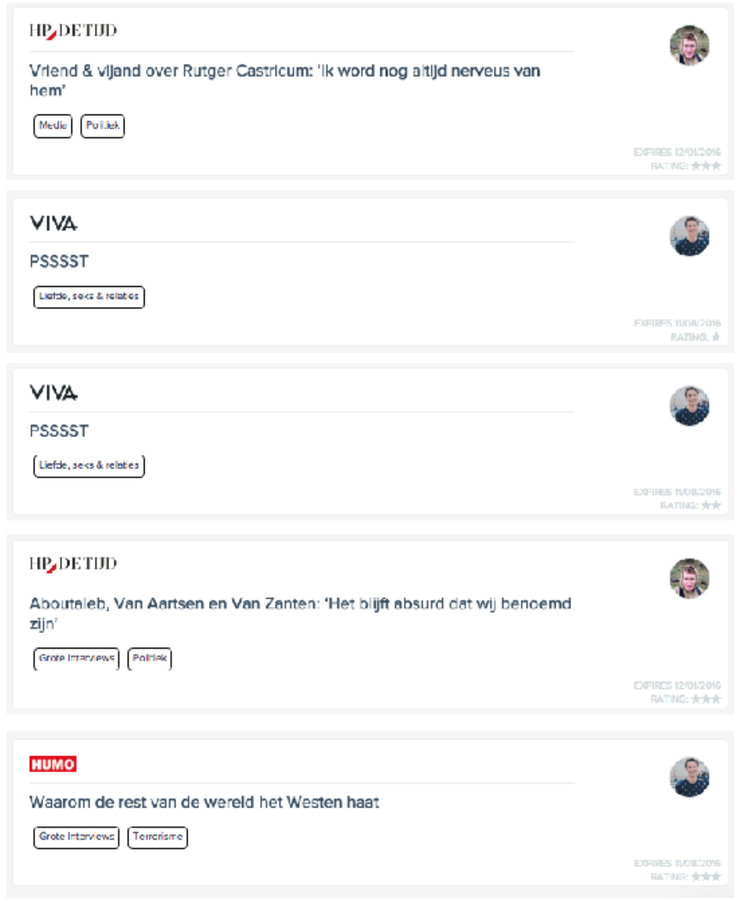
editor sends newsletter

Blendle

@annes

Bundler

Filter



editor sends newsletter

Blendle

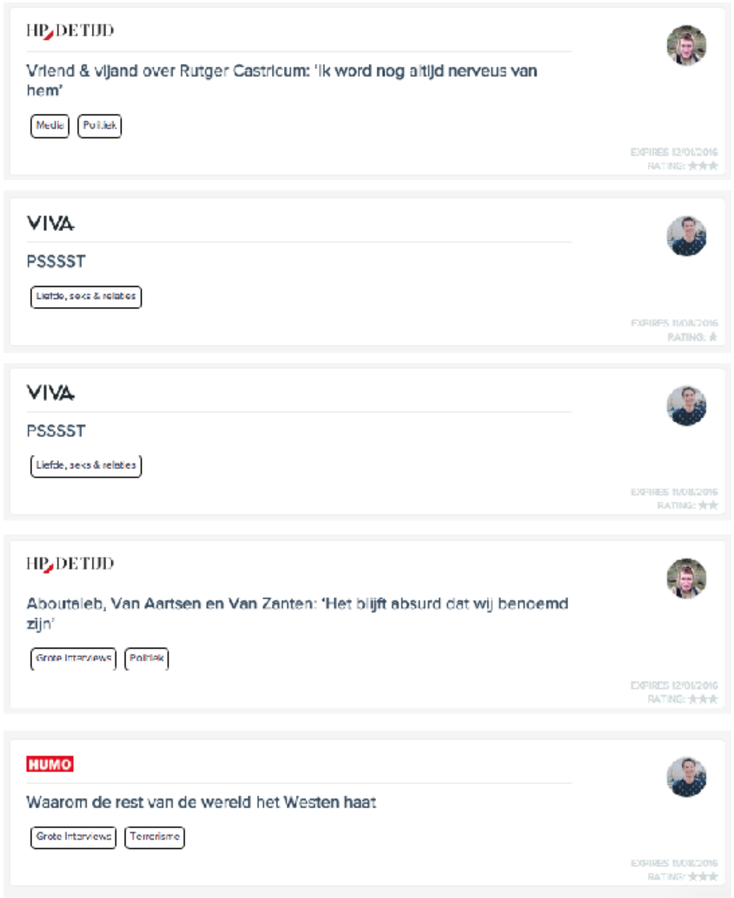
@annes

editor sends newsletter

Bundler

Filter

Score



Blendle

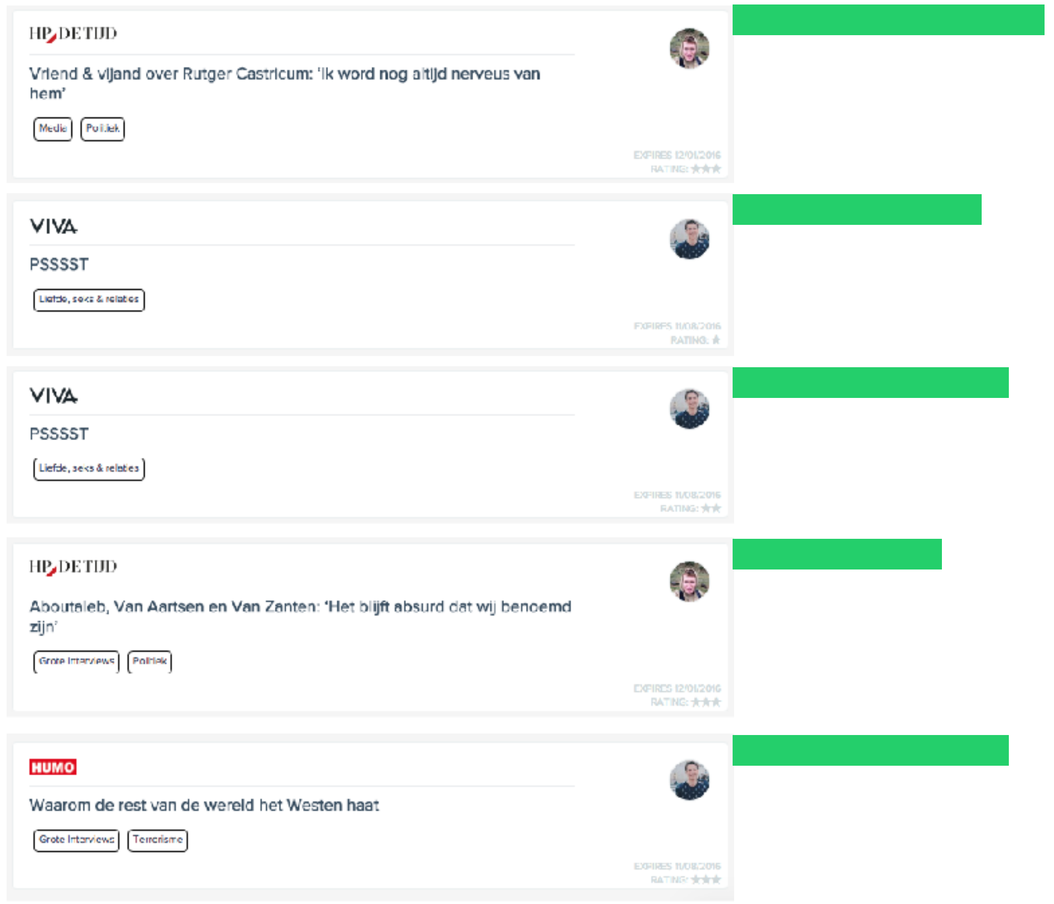
@annes

editor sends newsletter

Bundler

Filter

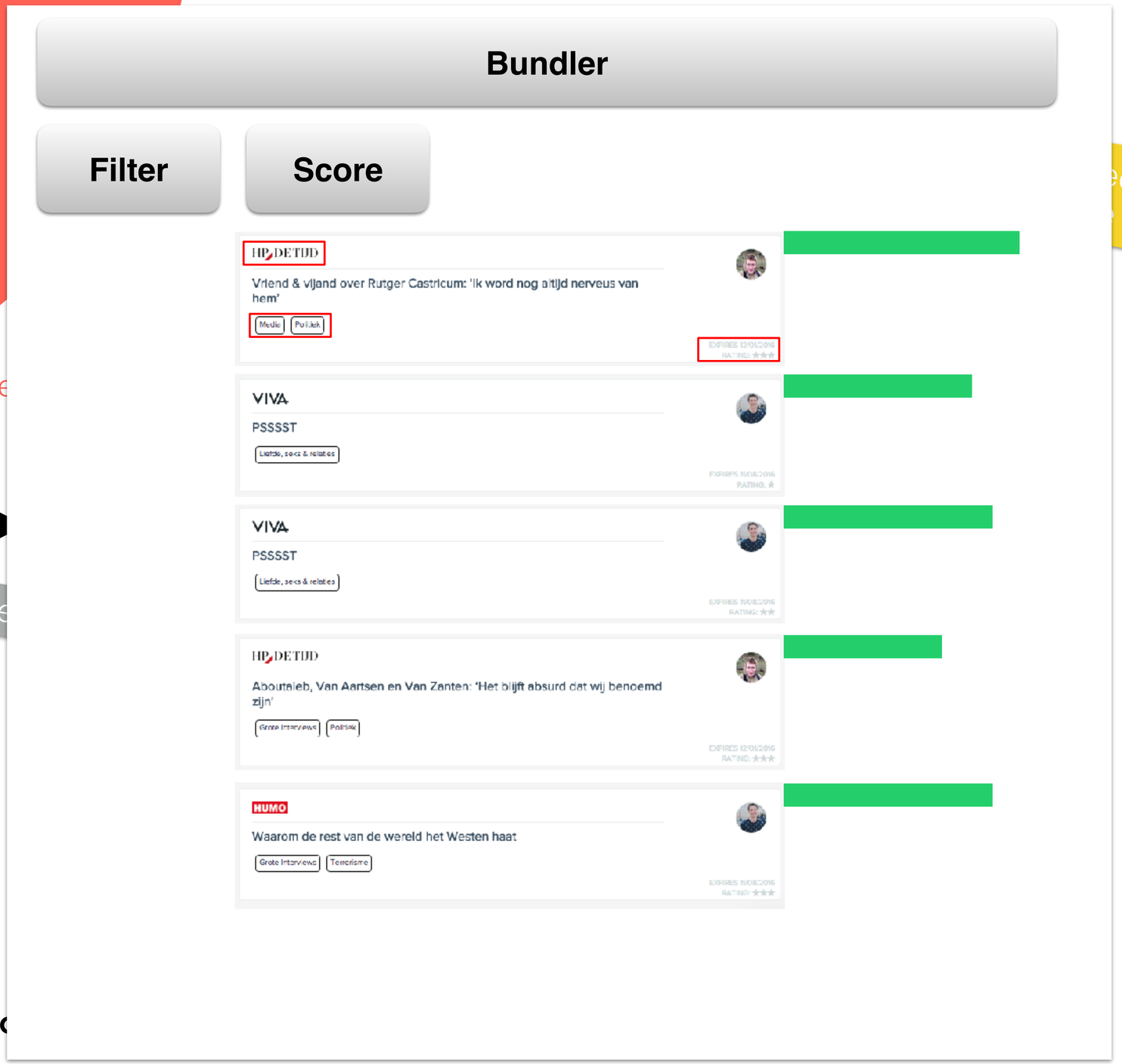
Score



Blendle

@annes

editor sends newsletter



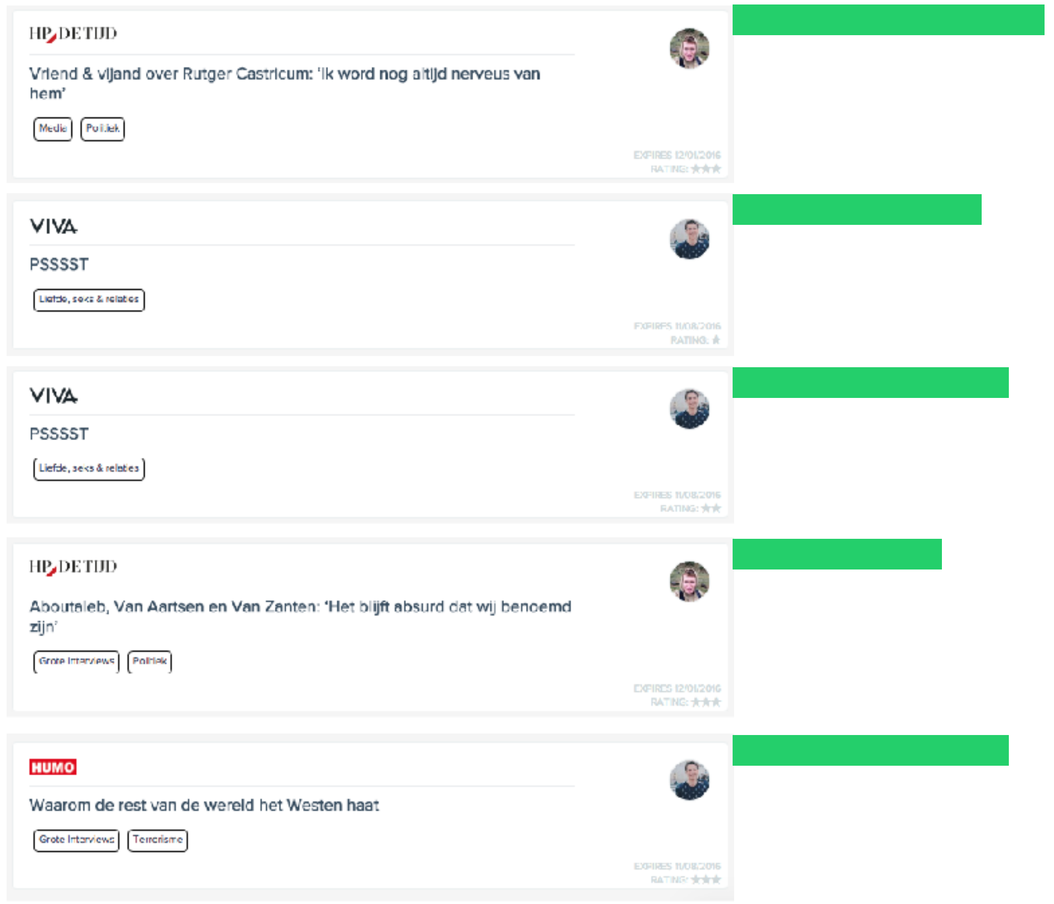
Blendle @annesc

editor sends newsletter

Bundler

Filter

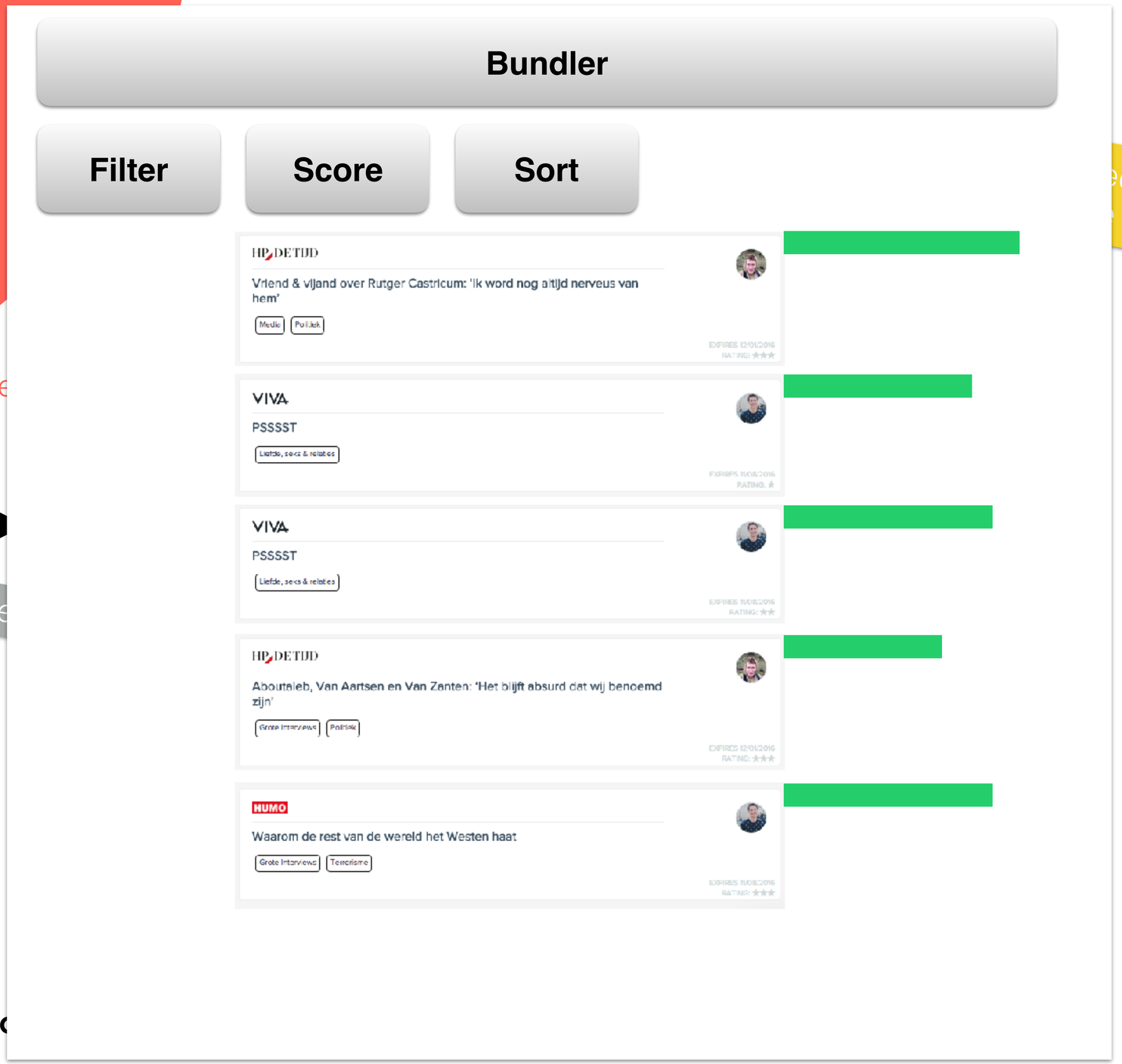
Score



Blendle

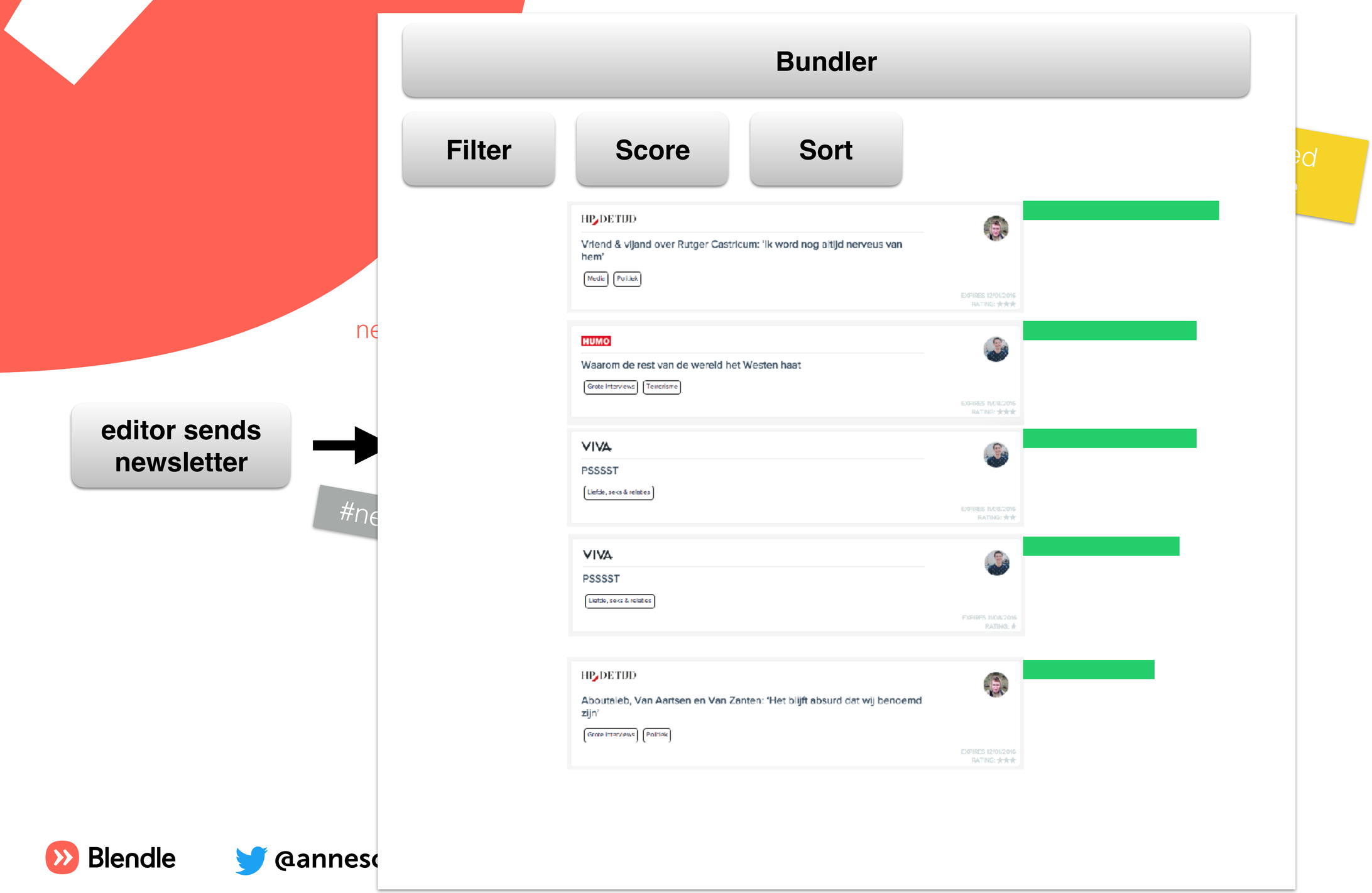
@annes

editor sends newsletter



Blendle

@annesc


Bundler
Filter
Score
Sort
editor sends newsletter
HP DE TIJD
Vriend & vijand over Rutger Castricum: 'Ik word nog altijd nerveus van hem'
Media
Politiek
HUMO
Waarom de rest van de wereld het Westen haat
Grote Interviews
Terrorisme
VIVA
PSSSST
Liefde, seks & relaties
VIVA
PSSSST
Liefde, seks & relaties
HP DE TIJD
Aboutaleb, Van Aartsen en Van Zanten: 'Het blijft absurd dat wij benoemd zijn'
Grote Interviews
Politiek
Blendle
@annesc

editor sends newsletter

#ne

Bundler

Filter | **Score** | **Sort** | **Diversify**

HP/DE TIJD

Vriend & vijand over Rutger Castricum: 'Ik word nog altijd nerveus van hem'

Media | Politiek

EXPIRES 12/01/2016
RATING: ★★★

HUMO

Waarom de rest van de wereld het Westen haat

Grote interviews | Terrorisme

EXPIRES 11/01/2016
RATING: ★★★

VIVA

PSSSST

Liefde, sex & relaties

EXPIRES 11/01/2016
RATING: ★★

VIVA

PSSSST

Liefde, sex & relaties

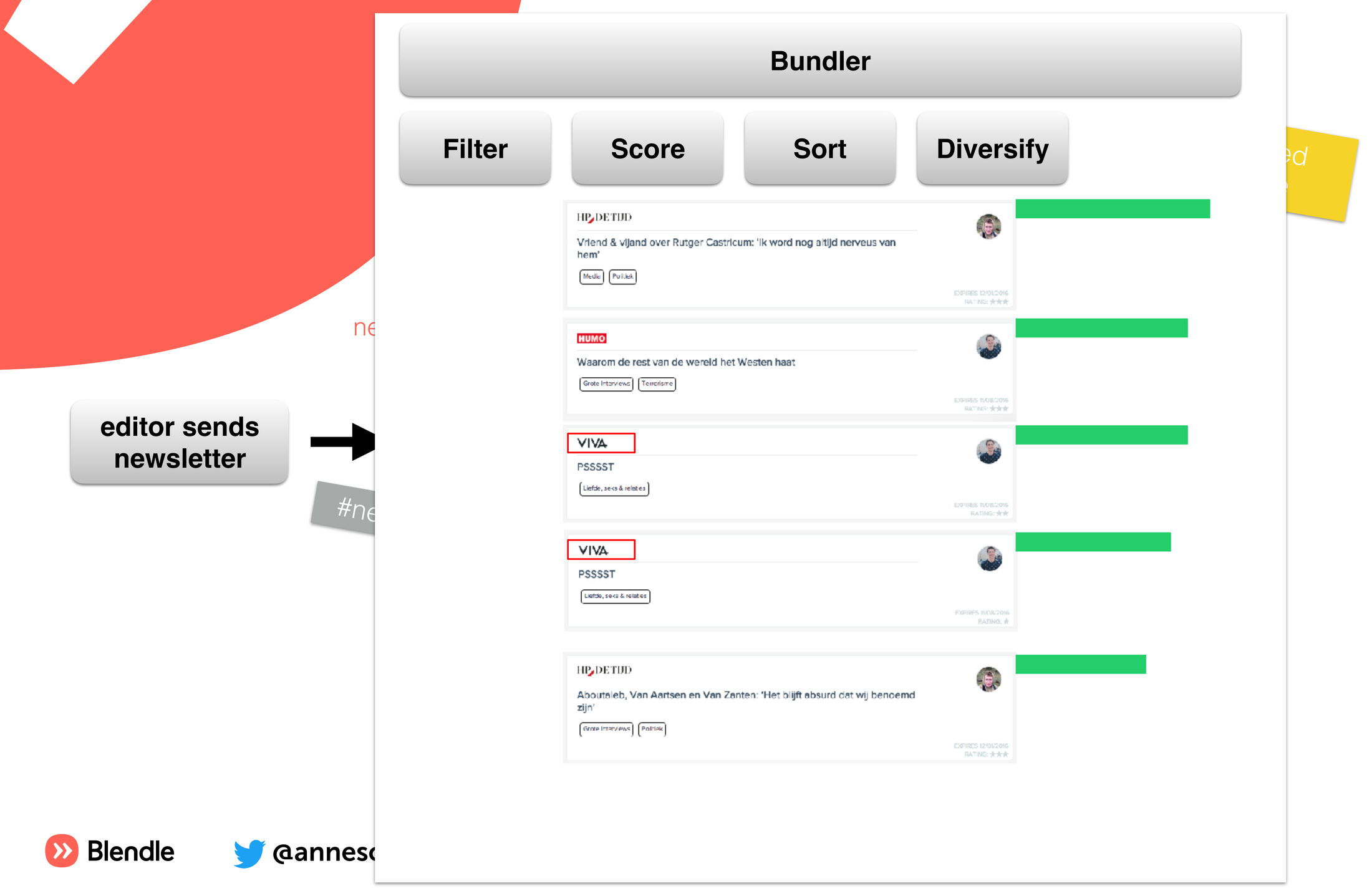
EXPIRES 11/01/2016
RATING: ★

HP/DE TIJD

Aboutaleb, Van Aartsen en Van Zanten: 'Het blijft absurd dat wij benoemd zijn'

Grote interviews | Politiek

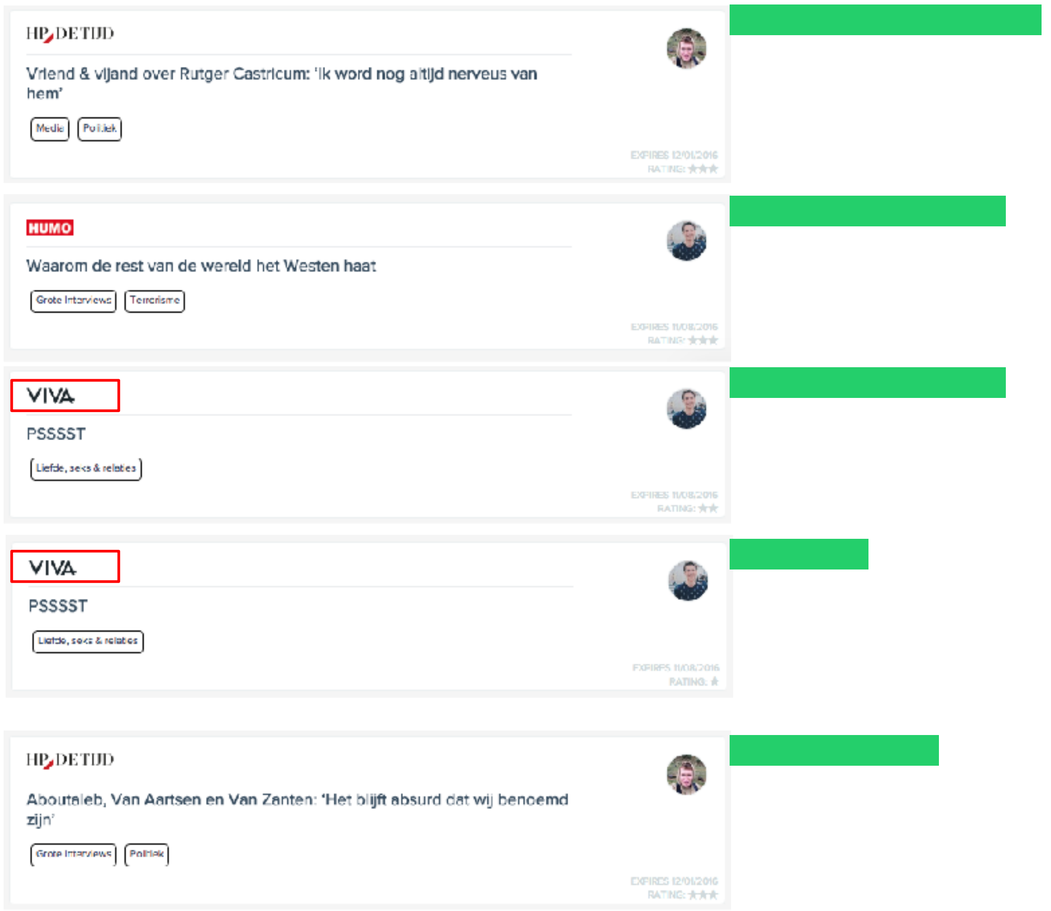
EXPIRES 12/01/2016
RATING: ★★★


Bundler
Filter
Score
Sort
Diversify
editor sends newsletter
HP DE TIJD
Vriend & vijand over Rutger Castricum: 'Ik word nog altijd nerveus van hem'
Media
Politiek
HUMO
Waarom de rest van de wereld het Westen haat
Grote Interviews
Terrorisme
VIVA
PSSSST
Liefde, seks & relaties
VIVA
PSSSST
Liefde, seks & relaties
HP DE TIJD
Aboutaleb, Van Aartsen en Van Zanten: 'Het blijft absurd dat wij benoemd zijn'
Grote Interviews
Politiek
Blendle
@annesc

editor sends newsletter

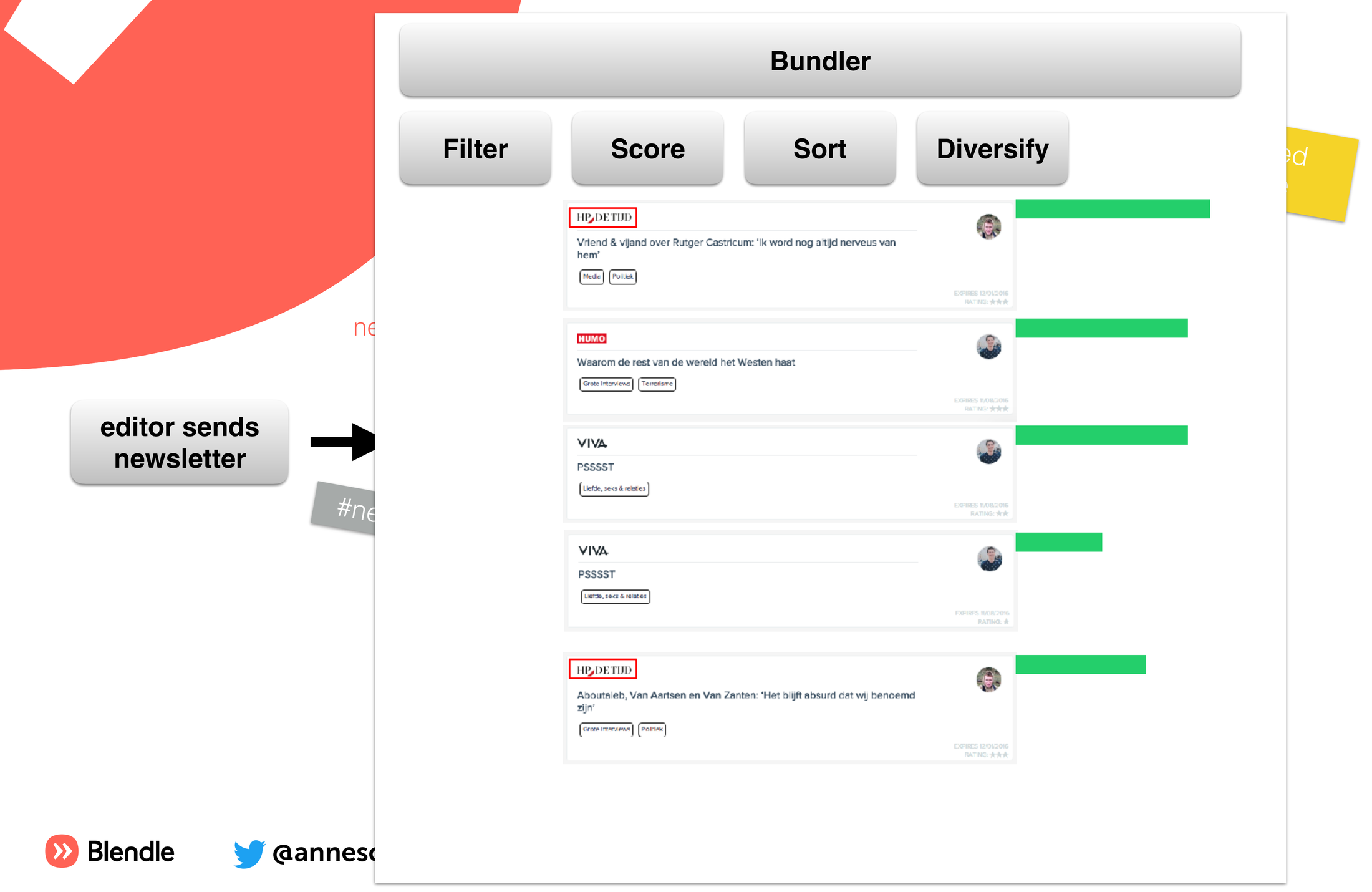
Bundler

Filter | **Score** | **Sort** | **Diversify**

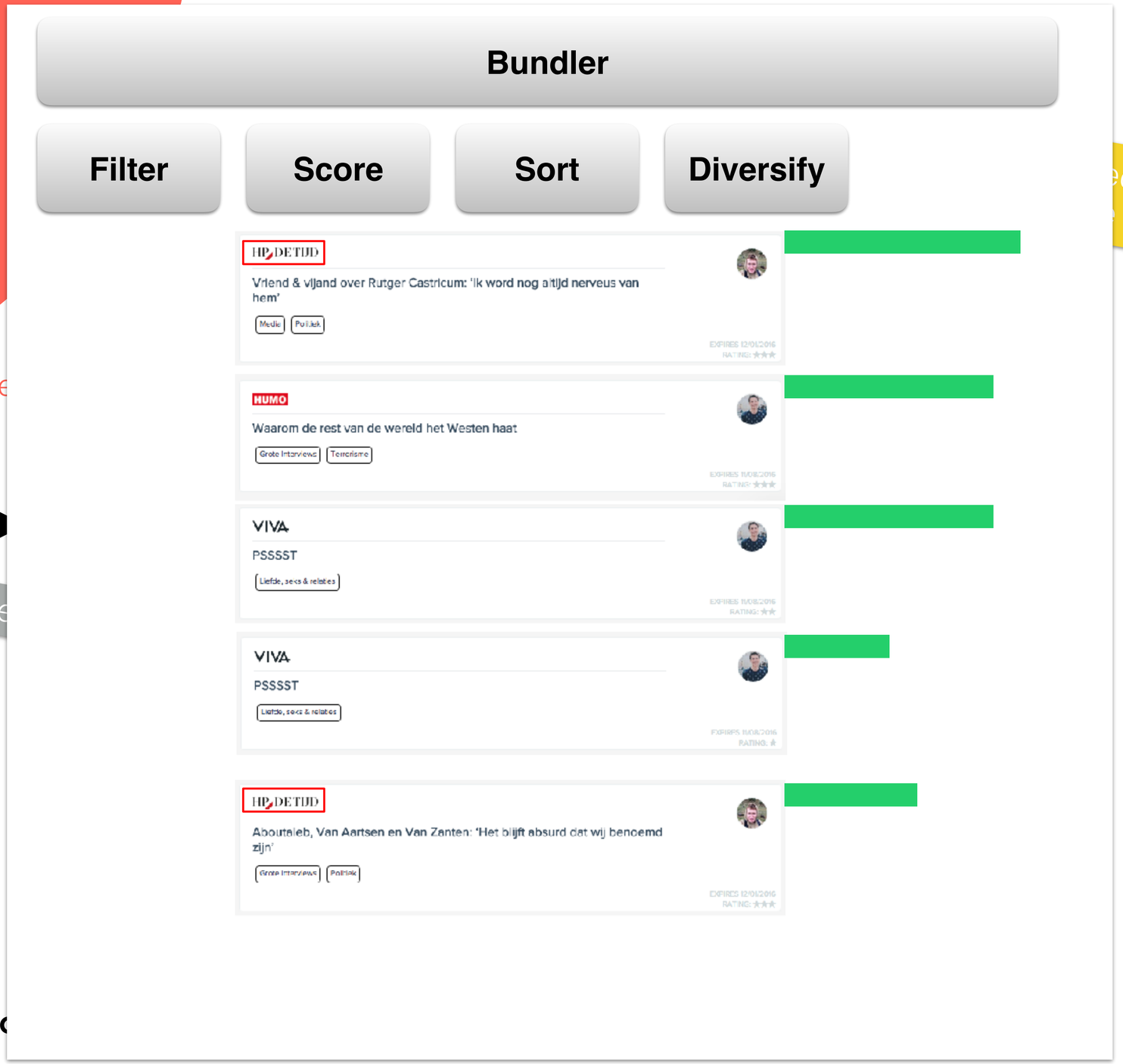


Blendle

@annes


Bundler
Filter
Score
Sort
Diversify
editor sends newsletter
HP DE TIJD
Vriend & vijand over Rutger Castricum: 'Ik word nog altijd nerveus van hem'
Media
Politiek
HUMO
Waarom de rest van de wereld het Westen haat
Grote Interviews
Terrorisme
VIVA
PSSSST
Liefde, seks & relaties
VIVA
PSSSST
Liefde, seks & relaties
HP DE TIJD
Aboutaleb, Van Aartsen en Van Zanten: 'Het blijft absurd dat wij benoemd zijn'
Grote Interviews
Politiek
Blendle
@annesc

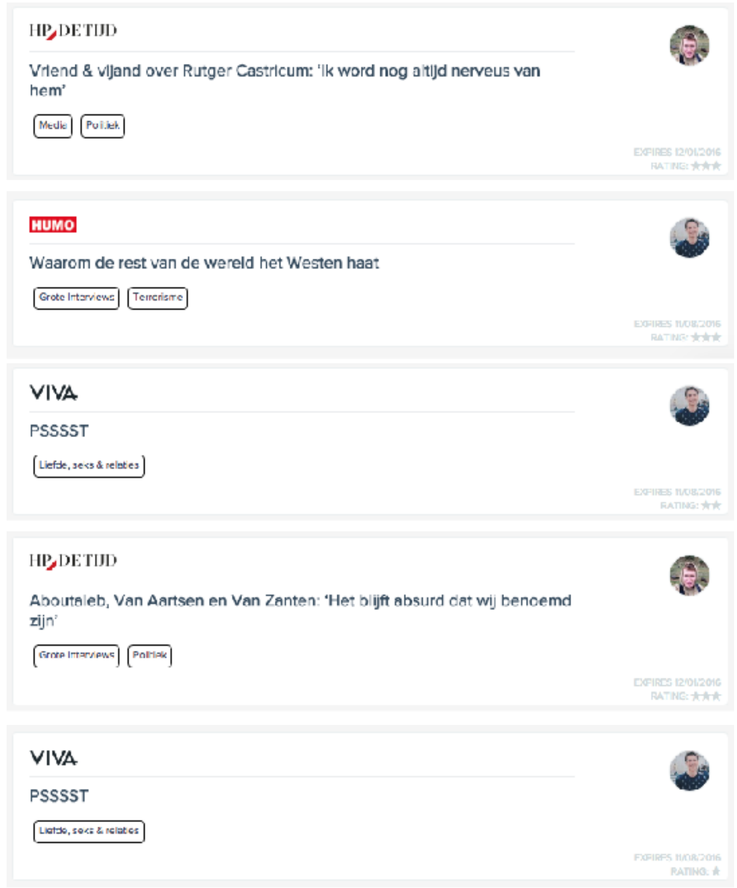
editor sends newsletter



Blendle @annesc

Bundler

Filter **Score** **Sort** **Diversify**

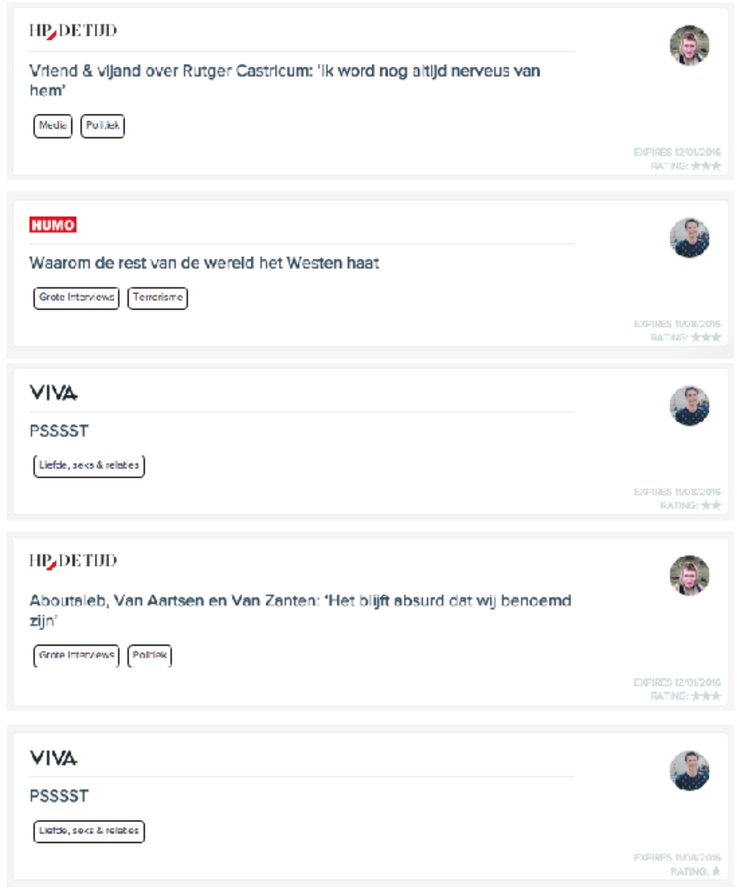


editor sends newsletter

Blendle @annes

Bundler

Filter | **Score** | **Sort** | **Diversify** | **Top**



editor sends newsletter

Blendle

@annes

editor sends newsletter

#ne

Bundler

Filter | **Score** | **Sort** | **Diversify** | **Top**

HP/DE TIJD

Vriend & vijand over Rutger Castricum: 'Ik word nog altijd nerveus van hem'

Media | Politiek

EXPIRES 12/01/2016
RATING: ★★★

HUMO

Waarom de rest van de wereld het Westen haat

Grote interviews | Terrorisme

EXPIRES 11/01/2016
RATING: ★★★

VIVA

PSSSST

Liefde, seks & relaties

EXPIRES 11/01/2016
RATING: ★★

HP/DE TIJD

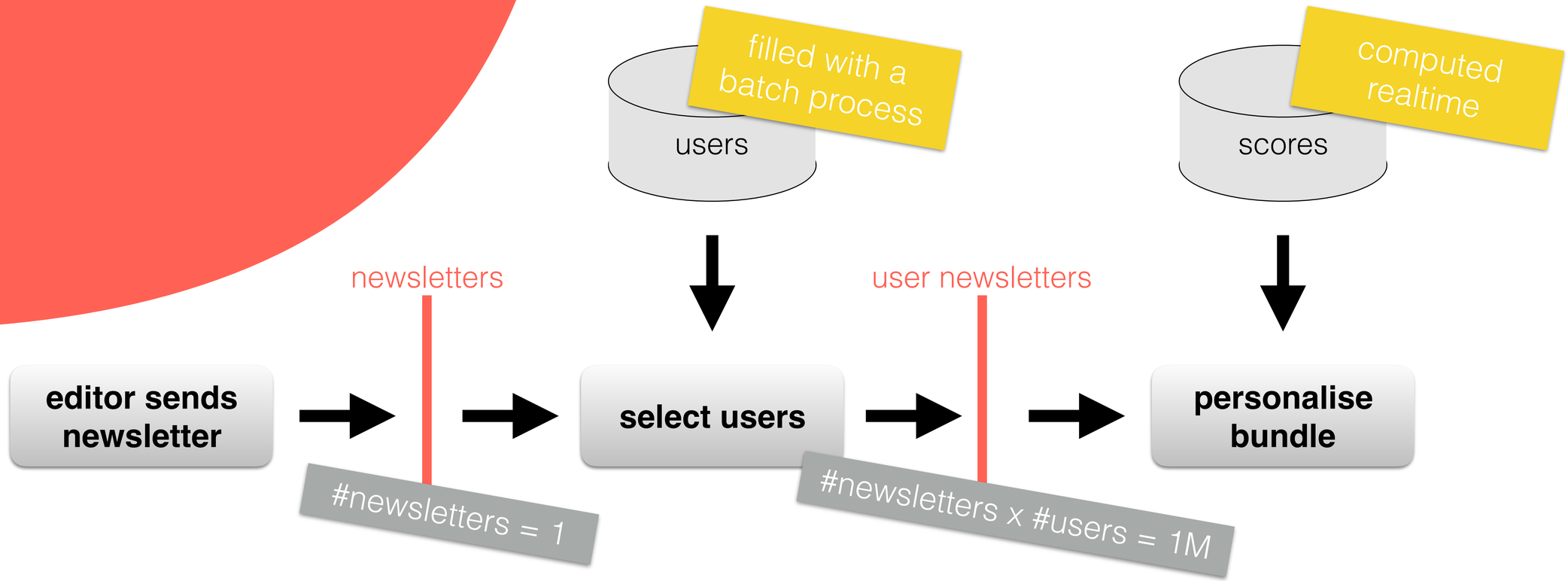
Aboutaleb, Van Aartsen en Van Zanten: 'Het blijft absurd dat wij benoemd zijn'

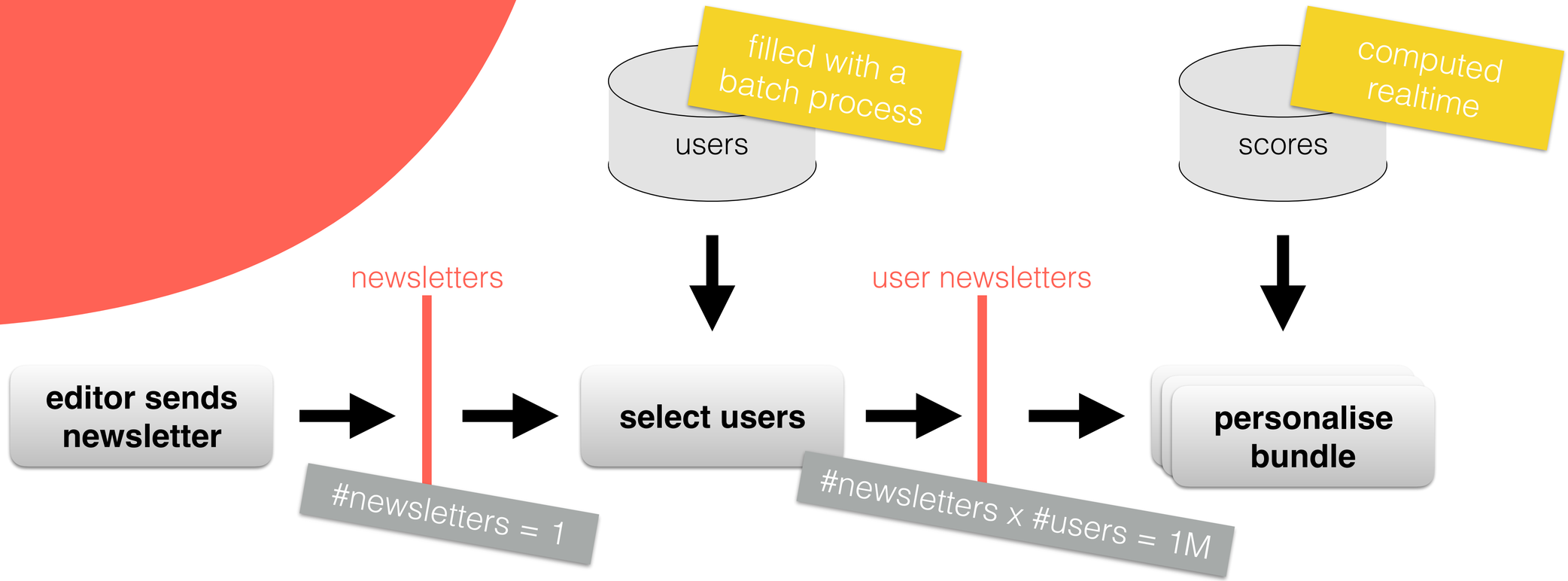
Grote interviews | Politiek

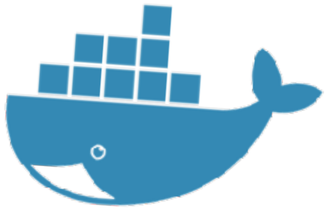
EXPIRES 12/01/2016
RATING: ★★★

Blendle

@annes

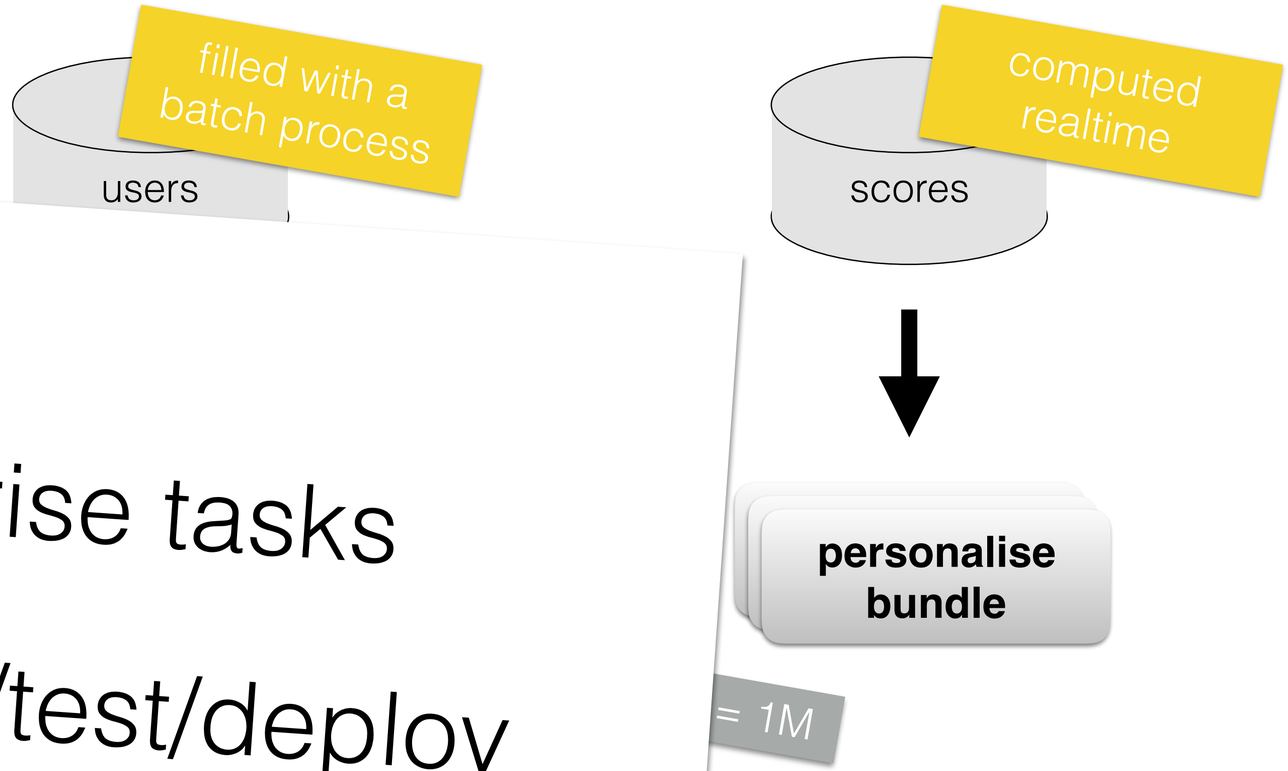

users
filled with a batch process
scores
computed realtime
newsletters
user newsletters
editor sends newsletter
select users
personalise bundle
#newsletters = 1
#newsletters x #users = 1M

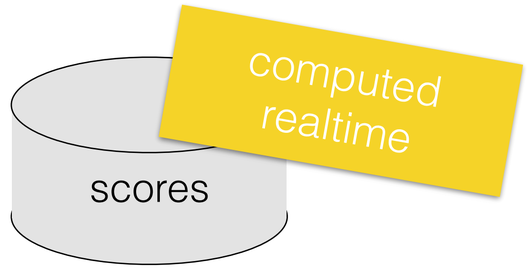

users
filled with a batch process
scores
computed realtime
newsletters
user newsletters
editor sends newsletter
select users
personalise bundle
#newsletters = 1
#newsletters x #users = 1M

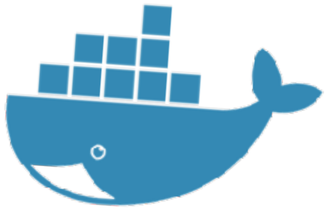


Docker

- containerise tasks
- easy dev/test/deploy





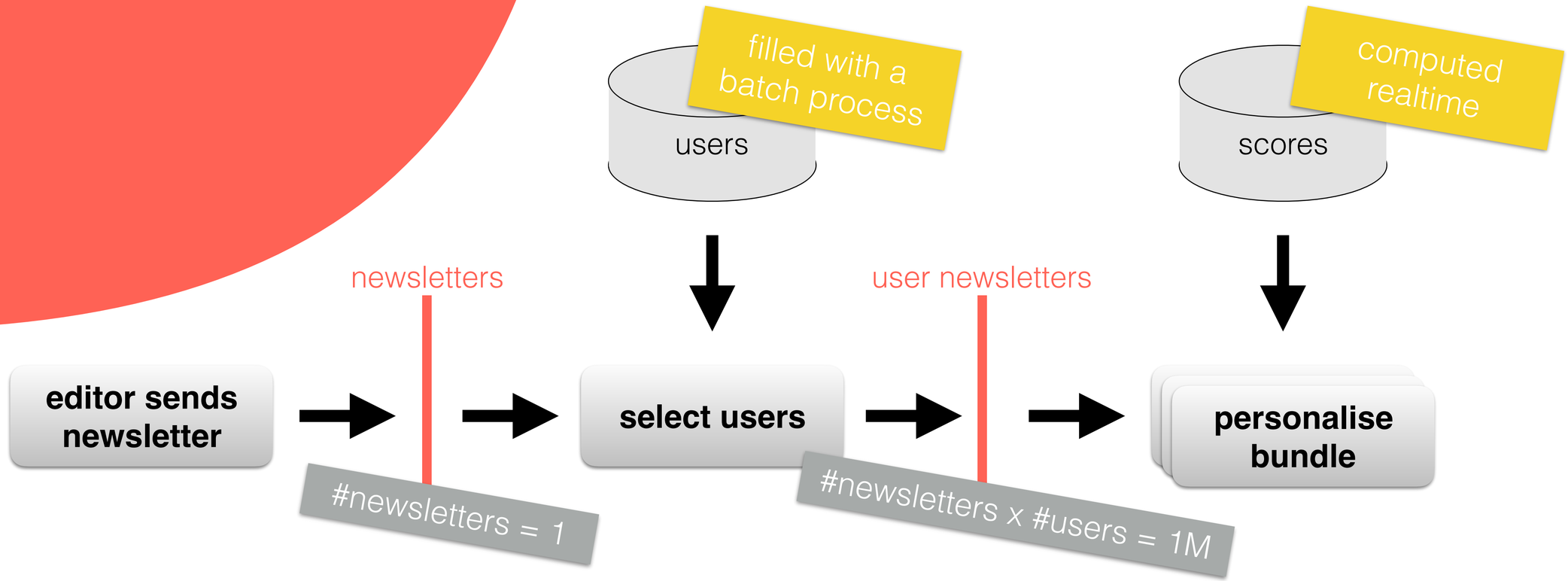


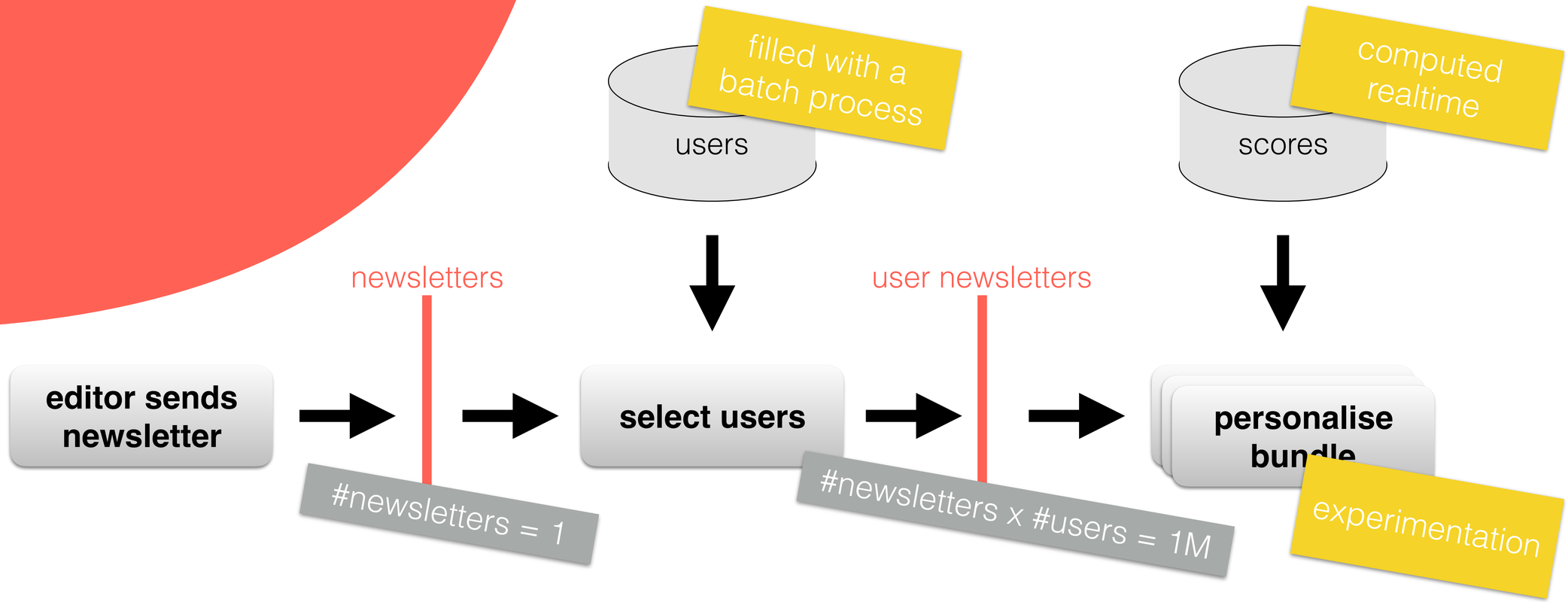
Docker

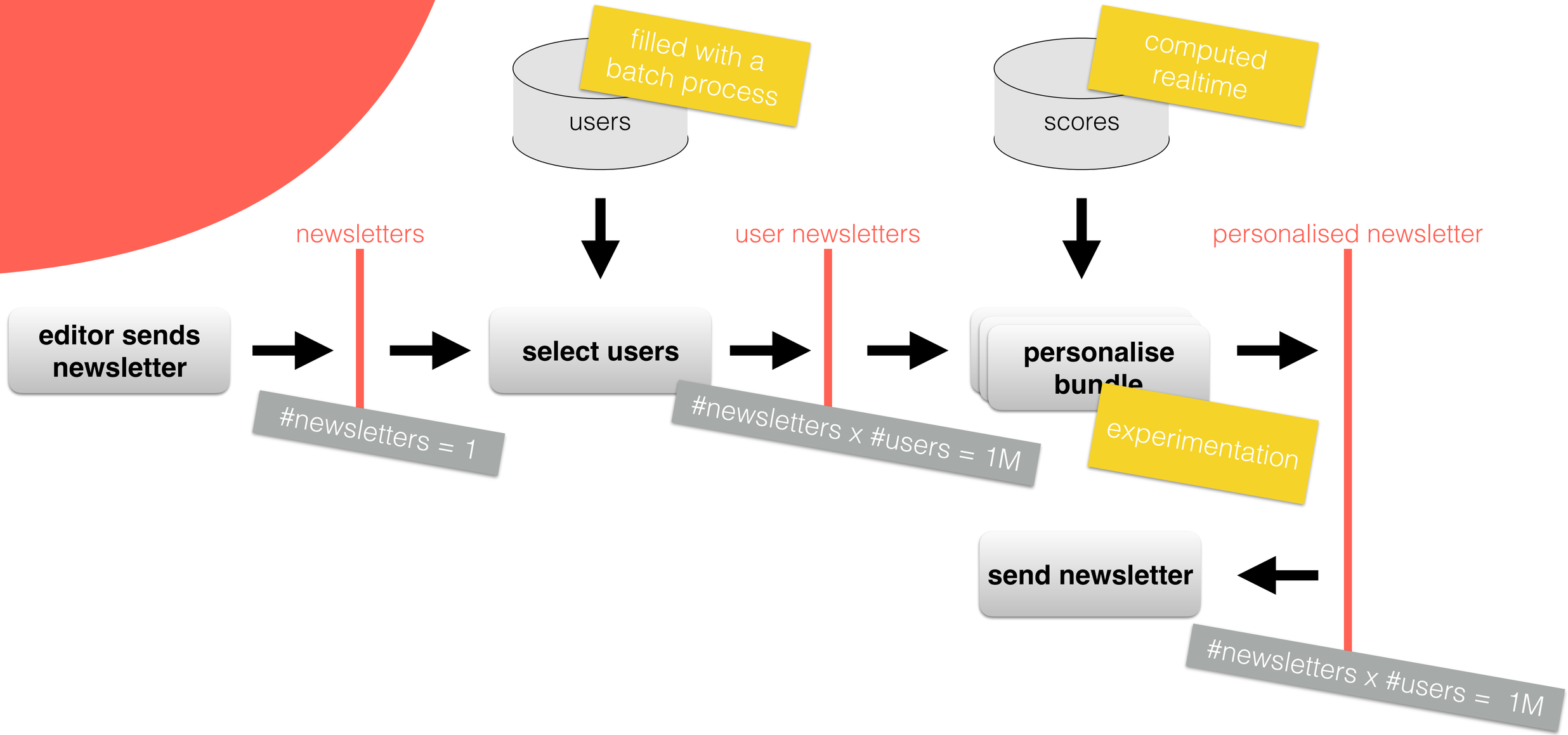
- containerise tasks
- easy dev/test/deploy

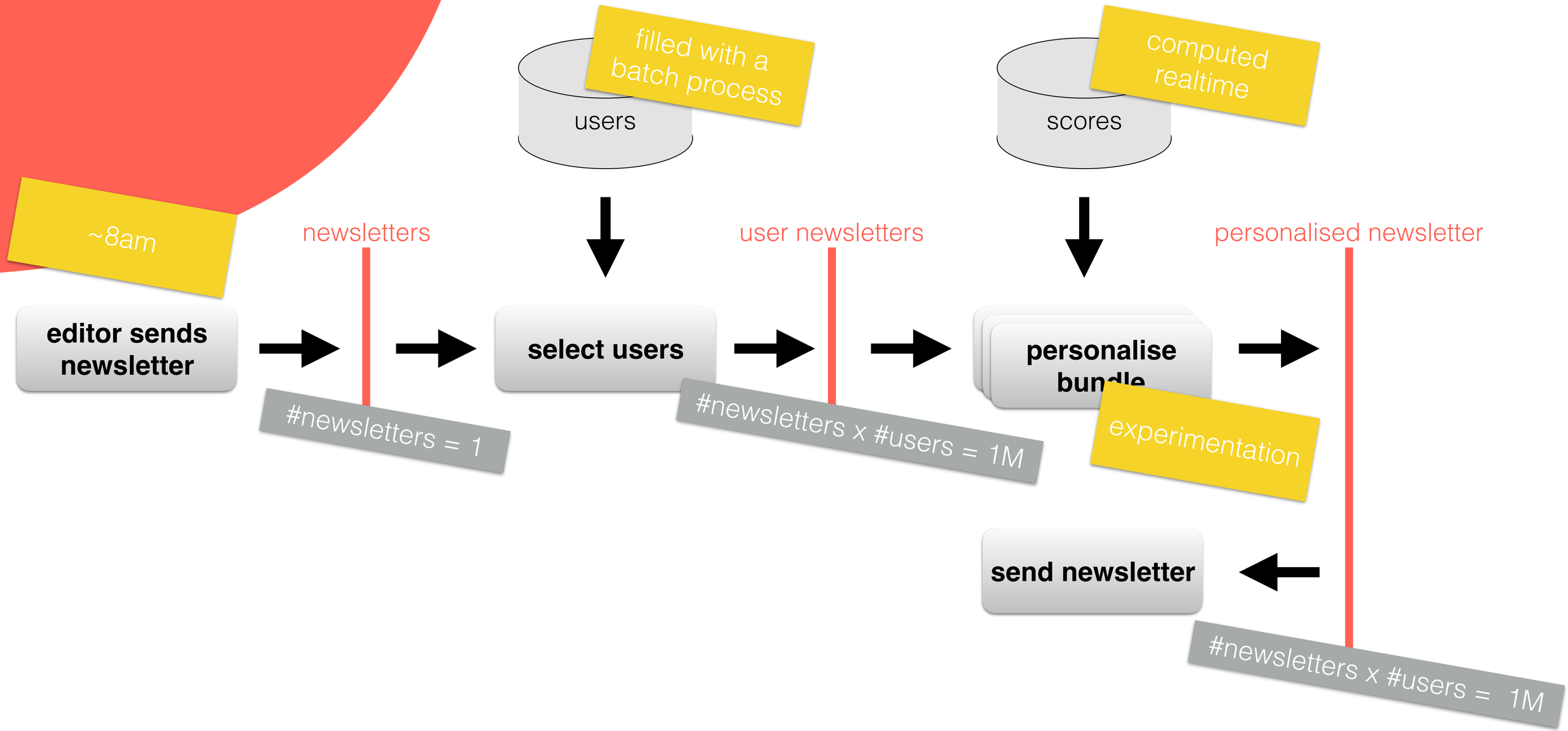
Kubernetes

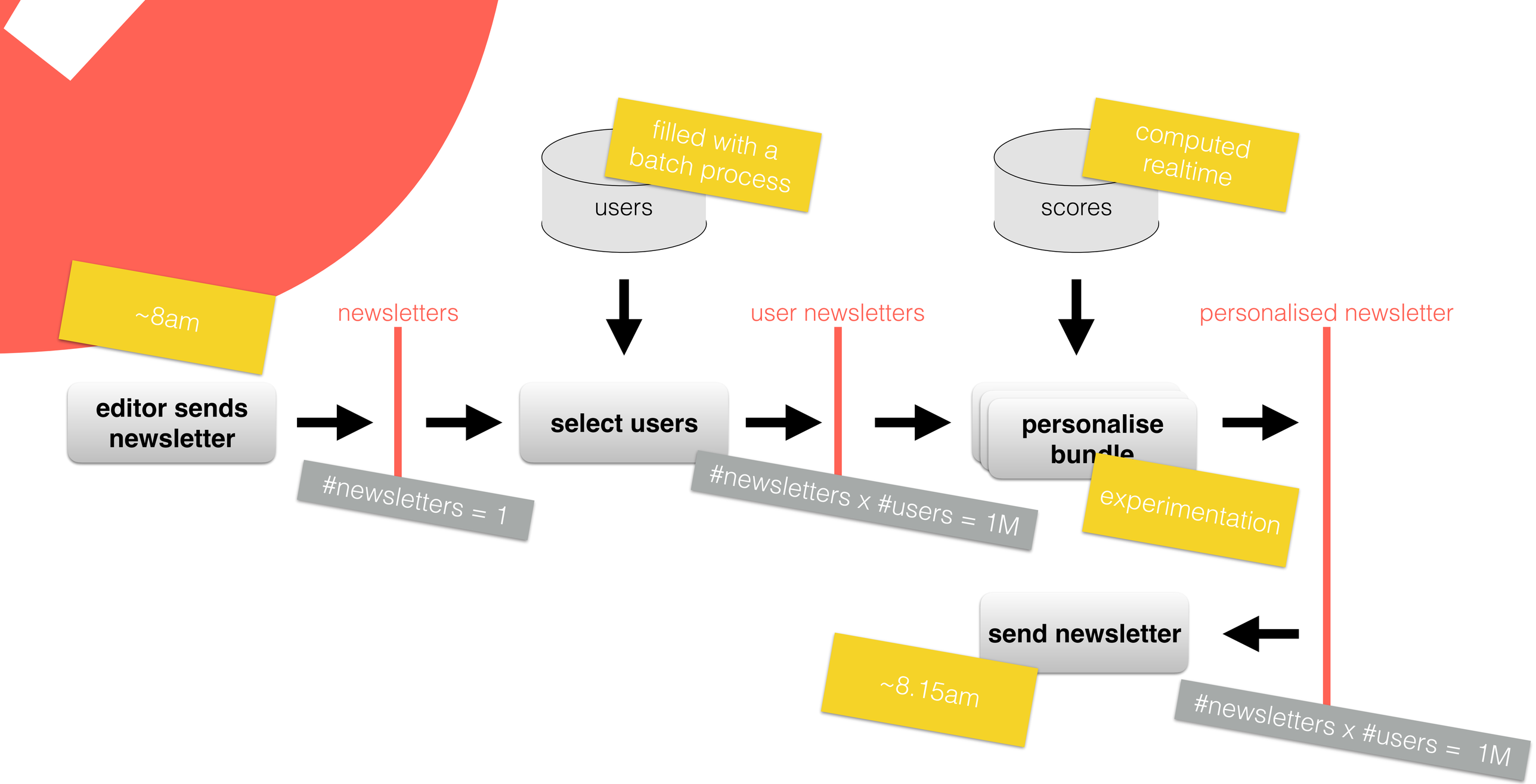
- orchestrate deployment
- autoscaling


users
filled with a batch process
scores
computed realtime
newsletters
user newsletters
editor sends newsletter
select users
personalise bundle
#newsletters = 1
#newsletters x #users = 1M


users
filled with a batch process
scores
computed realtime
newsletters
user newsletters
editor sends newsletter
select users
personalise bundle
#newsletters = 1
#newsletters x #users = 1M
experimentation


users
filled with a batch process
scores
computed realtime
newsletters
user newsletters
personalised newsletter
editor sends newsletter
select users
personalise bundle
send newsletter
#newsletters = 1
#newsletters x #users = 1M
experimentation
#newsletters x #users = 1M


~8am
editor sends newsletter
newsletters
#newsletters = 1
users
filled with a batch process
select users
user newsletters
#newsletters x #users = 1M
scores
computed realtime
personalise bundle
experimentation
personalised newsletter
send newsletter
#newsletters x #users = 1M


users
filled with a batch process
scores
computed realtime
~8am
newsletters
user newsletters
personalised newsletter
editor sends newsletter
select users
personalise bundle
experimentation
#newsletters = 1
#newsletters x #users = 1M
send newsletter
~8.15am
#newsletters x #users = 1M

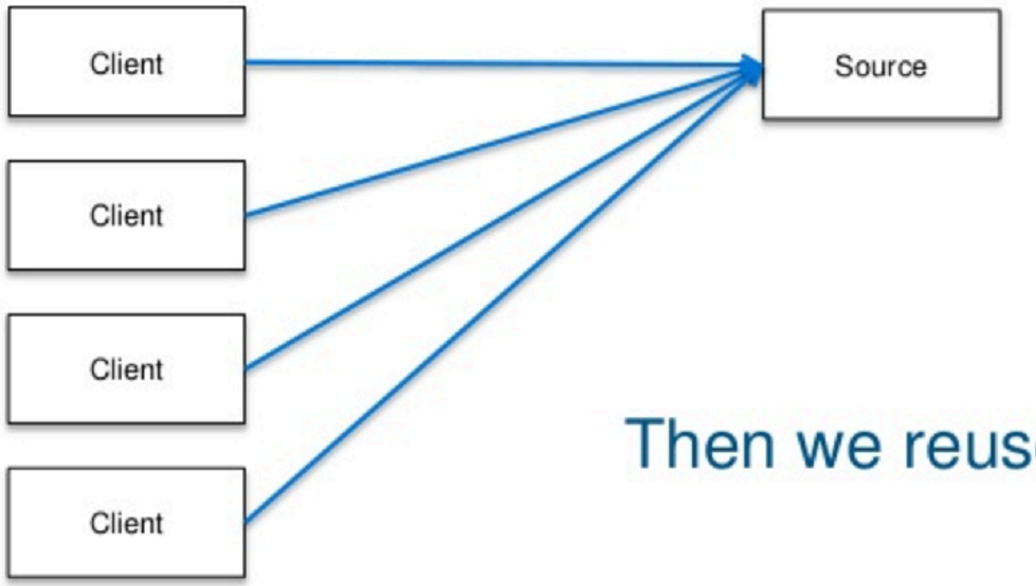
Clip slide

Why Kafka

Data Pipelines Start like this.

© 2015 Cloudera, Inc. All rights reserved.

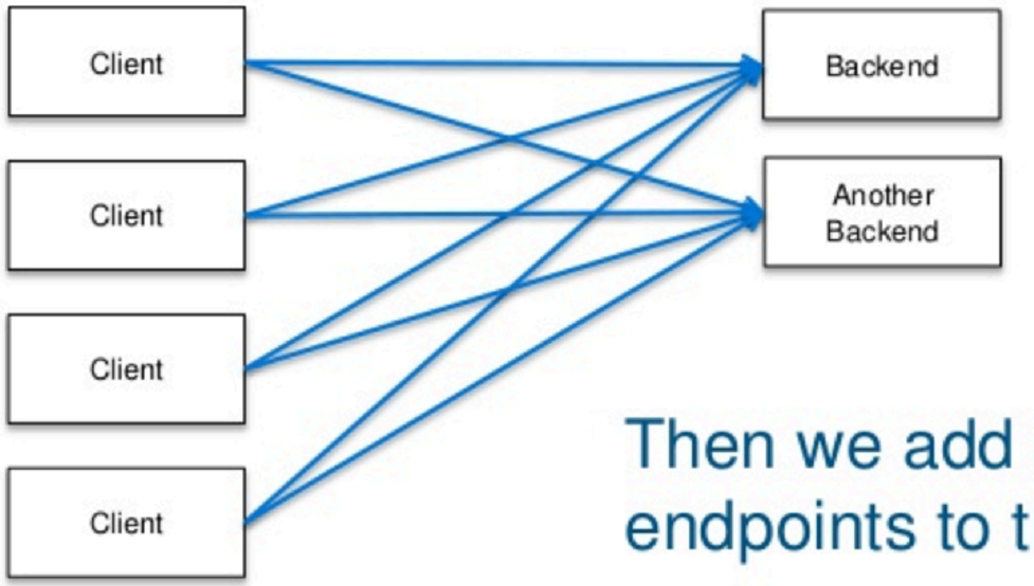
Why Kafka



Then we reuse them

© 2015 Cloudera, Inc. All rights reserved.

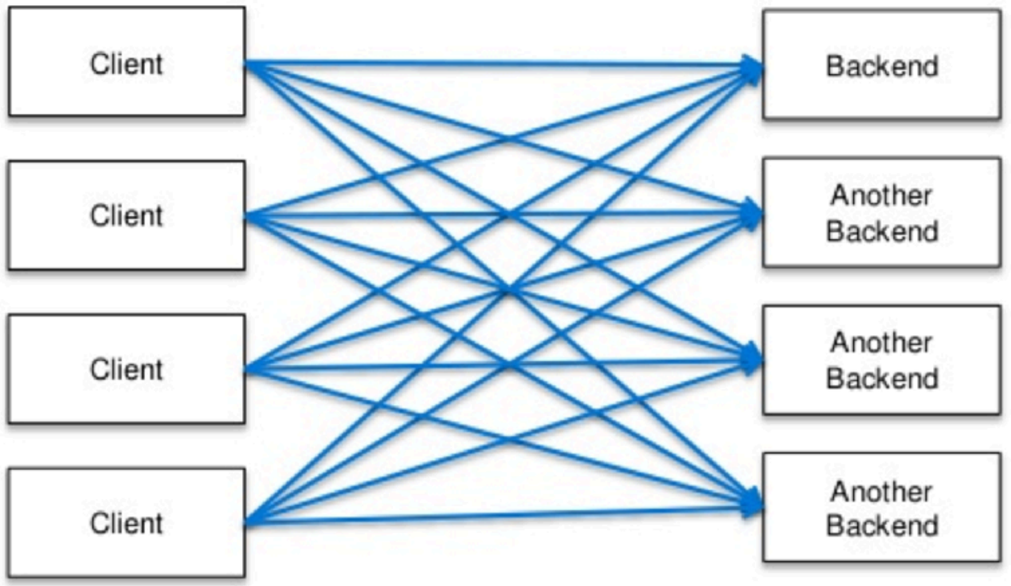
Why Kafka



Then we add additional endpoints to the existing sources

© 2015 Cloudera, Inc. All rights reserved.

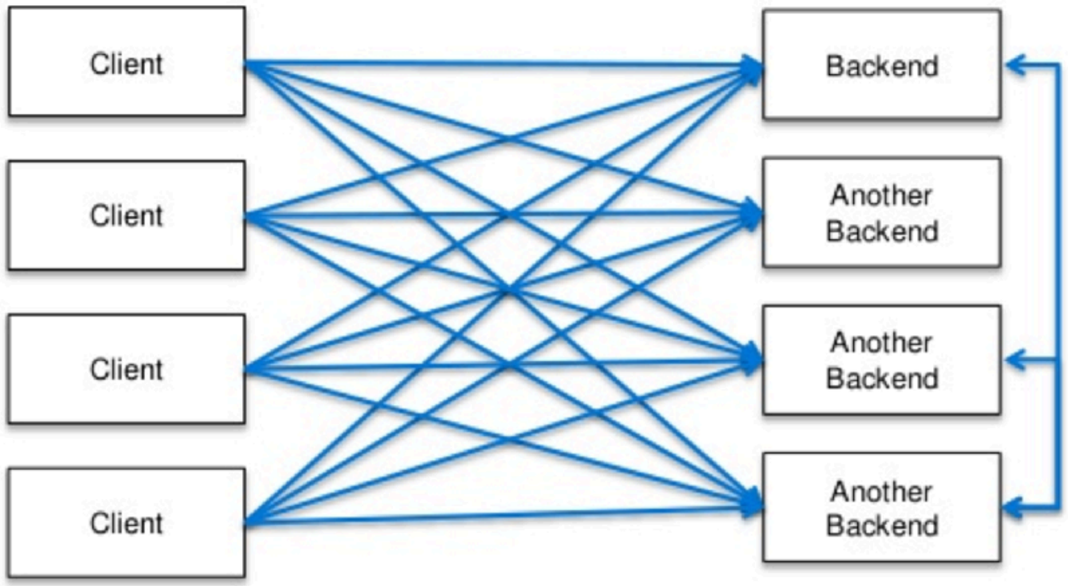
Why Kafka



Then it starts to look like this

© 2015 Cloudera, Inc. All rights reserved.

Why Kafka



With maybe some of this

As distributed systems and services increasingly become part of a modern architecture, this makes for a fragile system

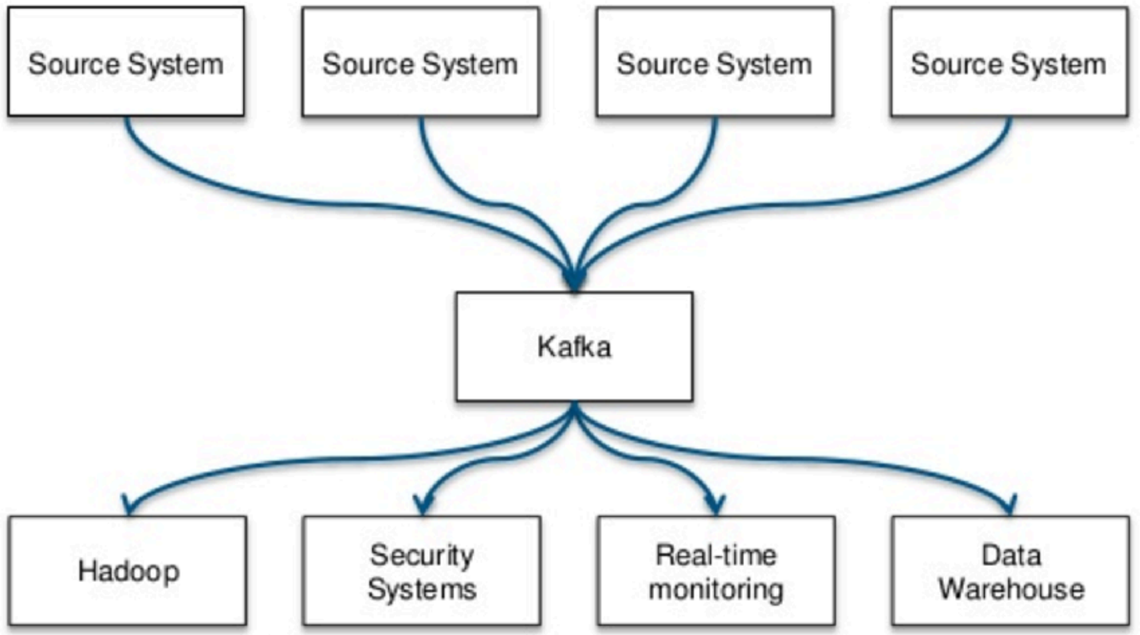
© 2015 Cloudera, Inc. All rights reserved.

Why Kafka

Producers

Brokers

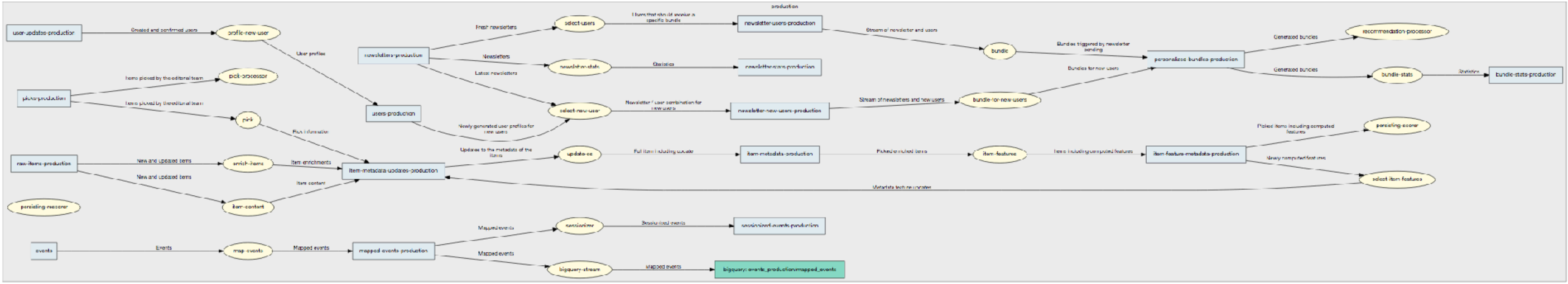
Consumers

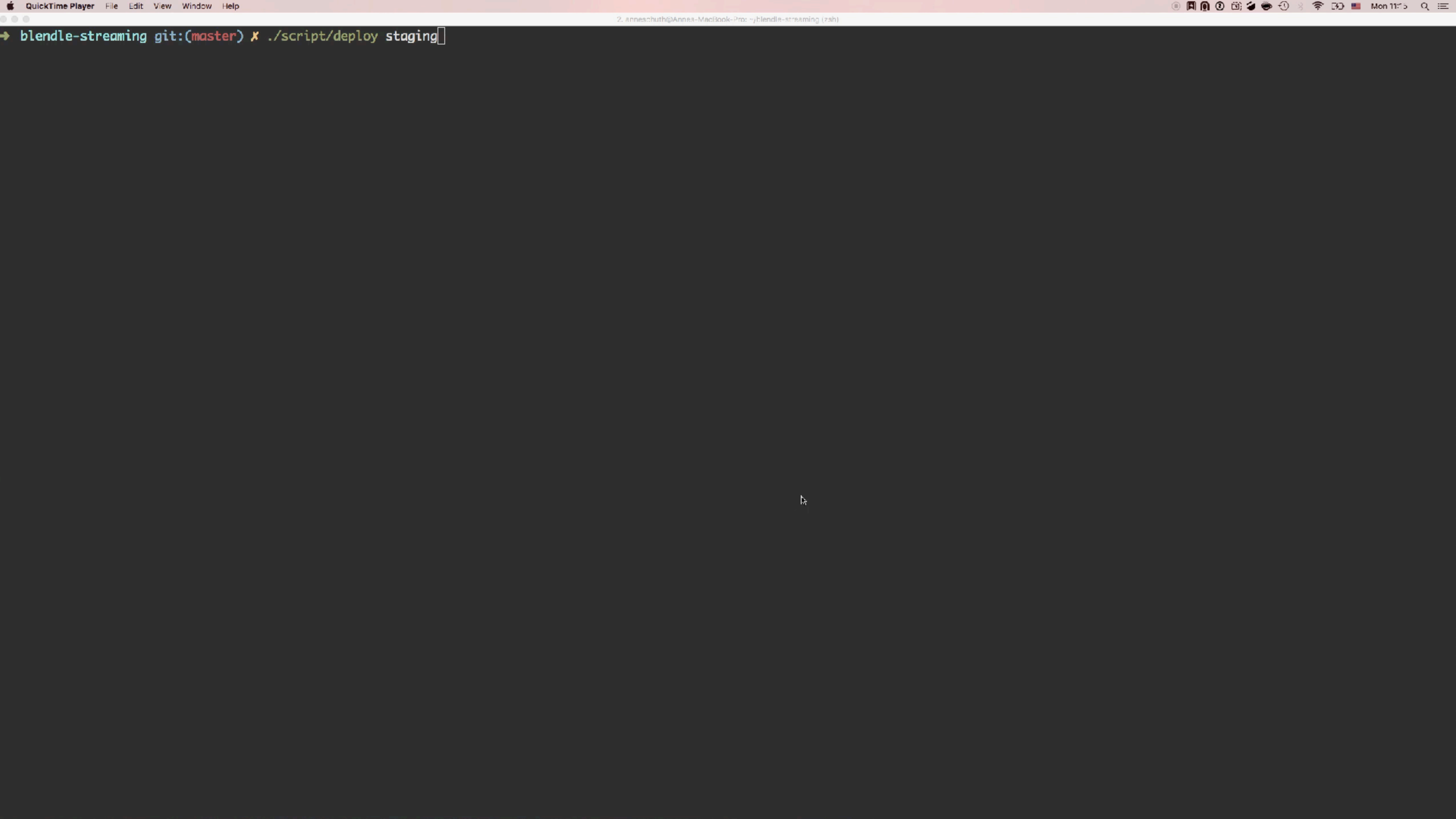


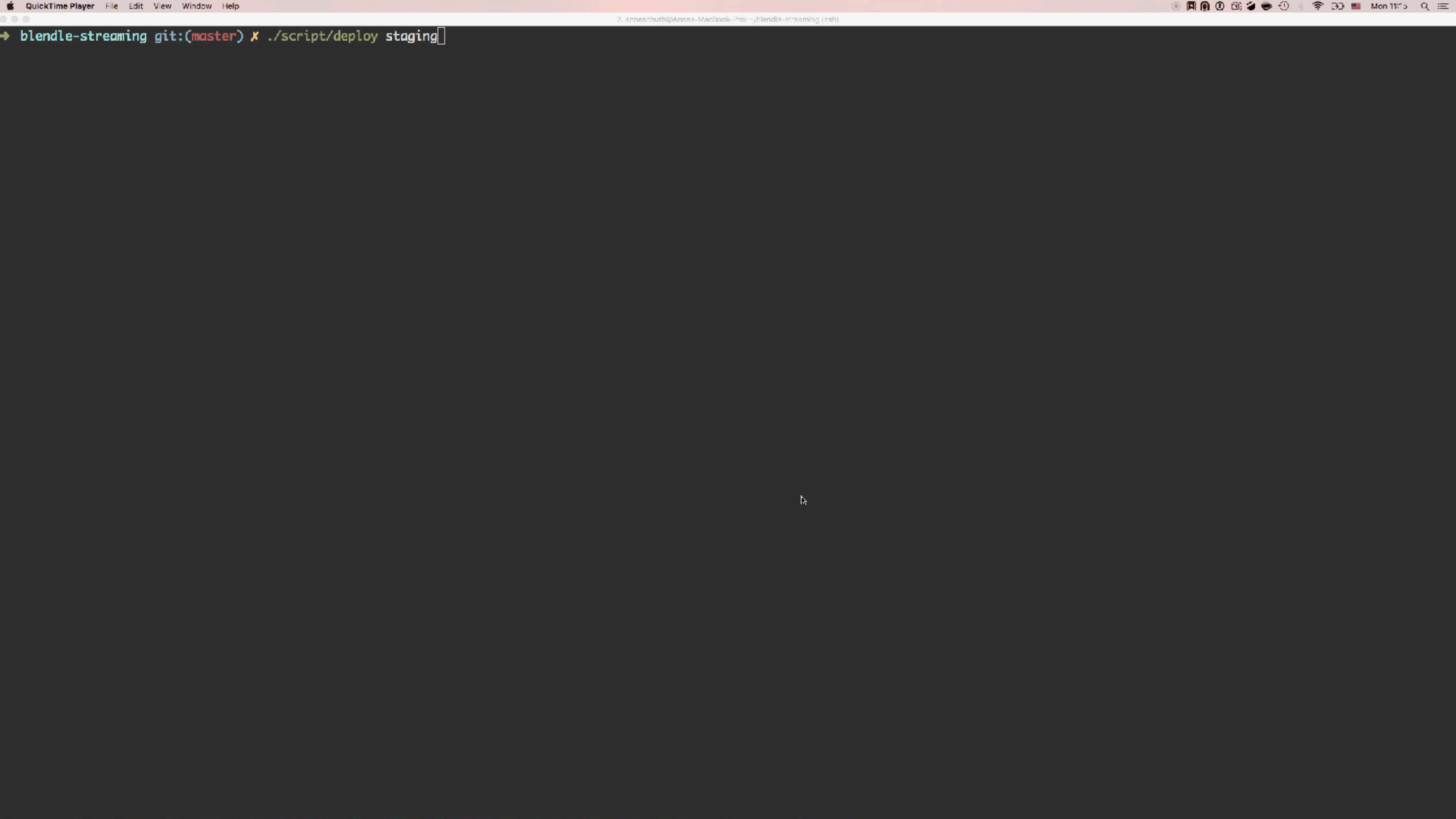
Kafka decouples Data Pipelines

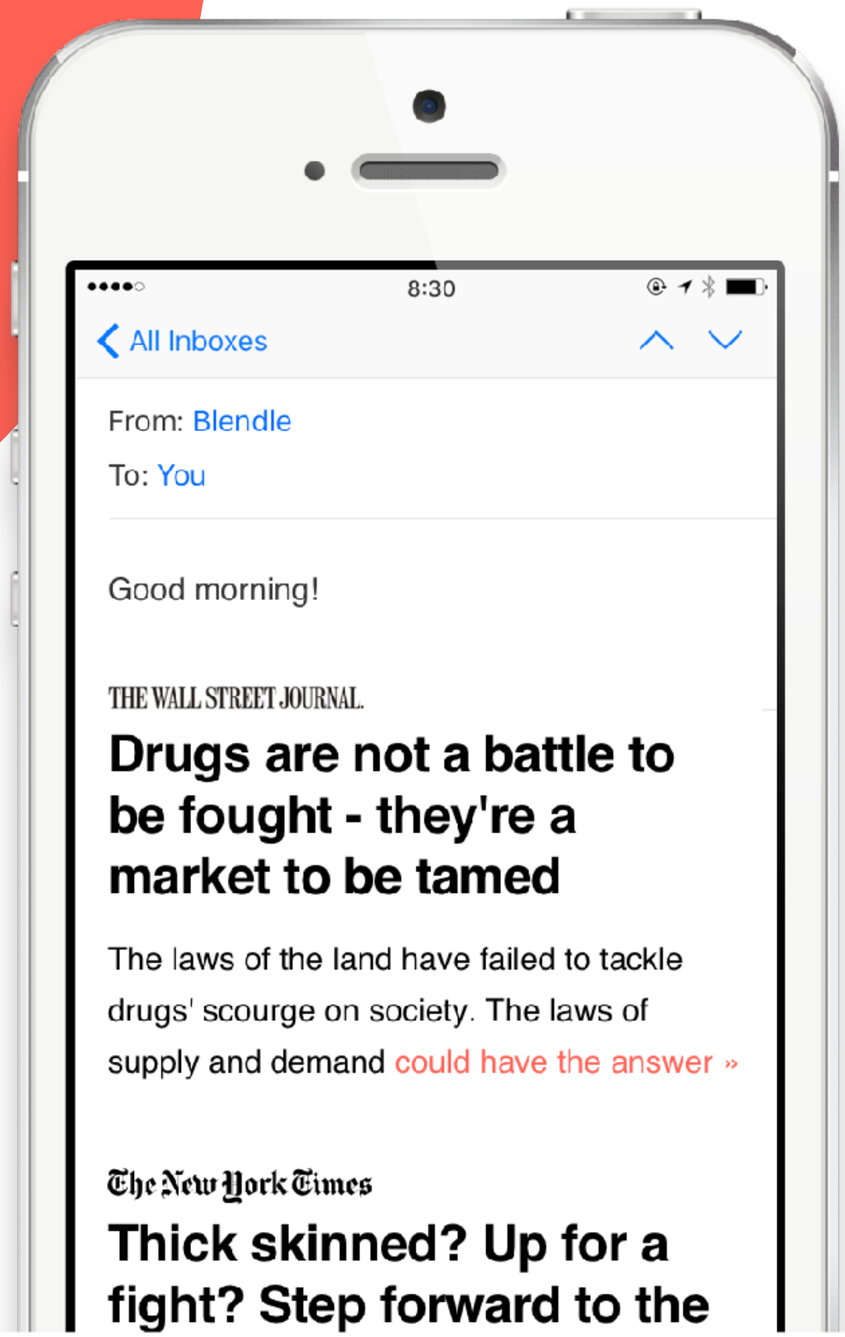
© 2015 Cloudera, Inc. All rights reserved.

Streaming Architecture



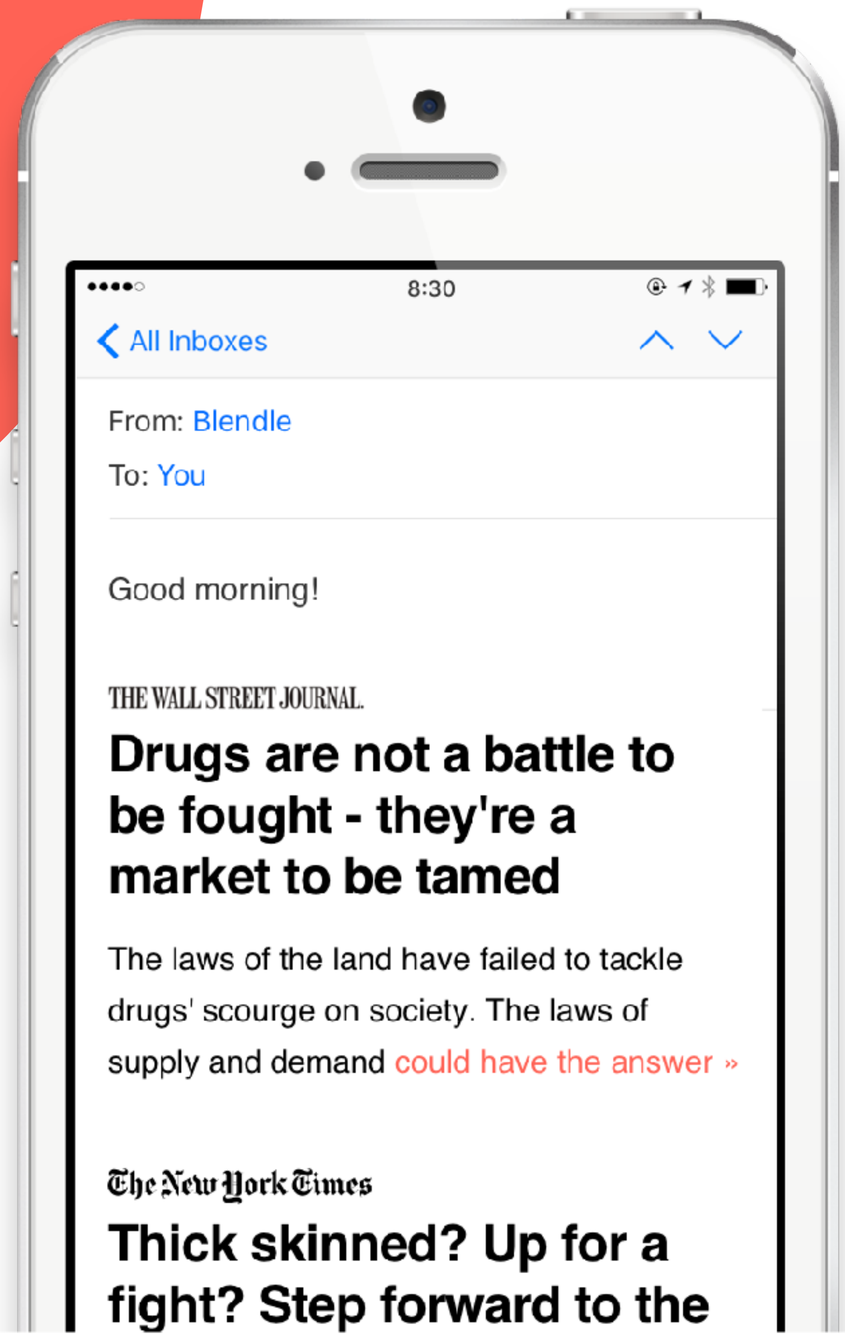

QuickTime Player File Edit View Window Help
2. anneschuth@Annes-MacBook-Pro: ~/blendle-streaming (zsh)
➜ blendle-streaming git:(master) ✗ ./script/deploy staging


QuickTime Player File Edit View Window Help
2. anneschuth@Annes-MacBook-Pro: ~/blendle-streaming (zsh)
➜ blendle-streaming git:(master) ✗ ./script/deploy staging



Cold start problem

- So we enrich our content



Timing problem

- So we precompute as much as possible

Components we use for **Machine Learning**

Components we use for **Machine Learning**

Spark
- processing large data

Components we use for **Machine Learning**

Spark
- processing large data

Components we use for **Machine Learning**

Spark
- processing large data

MLib
- production ML

Components we use for **Machine Learning**

Spark
- processing large data

S3
- raw content

MLib
- production ML

Components we use for **Machine Learning**

Spark
- processing large data

S3
- raw content

MLib
- production ML

Components we use for **Machine Learning**

Database
- data warehouse

Spark
- processing large data

S3
- raw content

MLib
- production ML

Components we use for **Machine Learning**

Luigi
- plumbing

Database
- data warehouse

Spark
- processing large data

S3
- raw content

MLib
- production ML

Components we use for **Machine Learning**

Luigi
- plumbing

Database
- data warehouse

Spark
- processing large data

S3
- raw content

MLib
- production ML

Scikitlearn
- experimental ML

Components we use for **Machine Learning**

Luigi
- plumbing

Database
- data warehouse

Spark
- processing large data

S3
- raw content

MLib
- production ML

Scikitlearn
- experimental ML

Elastic Search
- fast search and access

Components we use for **Machine Learning**

Redis
- fast access

Database
- data warehouse

Luigi
- plumbing

Spark
- processing large data

S3
- raw content

MLib
- production ML

Elastic Search
- fast search and access

Scikitlearn
- experimental ML

Blendle

@anneschuth

Components we use for **Machine Learning**

Redis
- fast access

Database
- data warehouse

Kafka
- streaming

Luigi
- plumbing

Spark
- processing large data

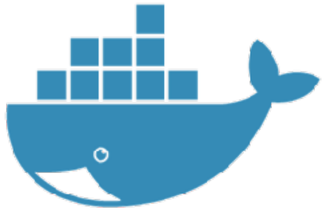
S3
- raw content

MLib
- production ML

Elastic Search
- fast search and access

Scikitlearn
- experimental ML

Components we use for **Machine Learning**



Docker + Kubernetes
- containerisation

Redis
- fast access

Database
- data warehouse

Kafka
- streaming

Luigi
- plumbing

Spark
- processing large data

S3
- raw content

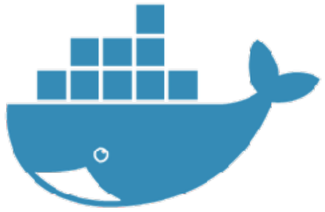
MLib
- production ML

Elastic Search
- fast search and access

Scikitlearn
- experimental ML

Blendle @anneschuth

Components we use for **Machine Learning**



Docker + Kubernetes
- containerisation

Database
- data warehouse

Redis
- fast access

Kafka
- streaming

Luigi
- plumbing

Cucumber
- BDD development

Spark
- processing large data

S3
- raw content

MLib
- production ML

Scikitlearn
- experimental ML

Elastic Search
- fast search and access

Internship Ideas

- Computer Vision

Internship Ideas

- Computer Vision
    - Death Recognition

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku
- Machine Learning

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku
- Machine Learning
  - Editorial Pick Prediction

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku
- Machine Learning
  - Editorial Pick Prediction
  - Blinder™, a Tinder-like app for reading articles

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku
- Machine Learning
  - Editorial Pick Prediction
  - Blinder™, a Tinder-like app for reading articles
  - Explanations of why we recommend

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku
- Machine Learning
  - Editorial Pick Prediction
  - Blinder™, a Tinder-like app for reading articles
  - Explanations of why we recommend
- NLP

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku
- Machine Learning
  - Editorial Pick Prediction
  - Blinder™, a Tinder-like app for reading articles
  - Explanations of why we recommend
- NLP
  - ‘Here’s the Blendle link’ Twitter bot

Internship Ideas

- Computer Vision
  - Death Recognition
  - ‘Share from the newspaper’
  - Find Replacement Images
  - Sudoku
- Machine Learning
  - Editorial Pick Prediction
  - Blinder™, a Tinder-like app for reading articles
  - Explanations of why we recommend
- NLP
  - ‘Here’s the Blendle link’ Twitter bot
  - Reactions on Twitter

Thank You
Blendle
@anneschuth