

# Real-time Recommendation

Anne Schuth  
*data scientist*

» Blendle



# Blendle

## ✓ Mission

Help you discover and support the world's best journalism



# Blendle

## ✓ Mission

Help you discover and support the world's best journalism

## ✓ Micropayment platform

articles, issues, subscriptions

# Blendle

## ✓ Mission

Help you discover and support the world's best journalism

## ✓ Micropayment platform

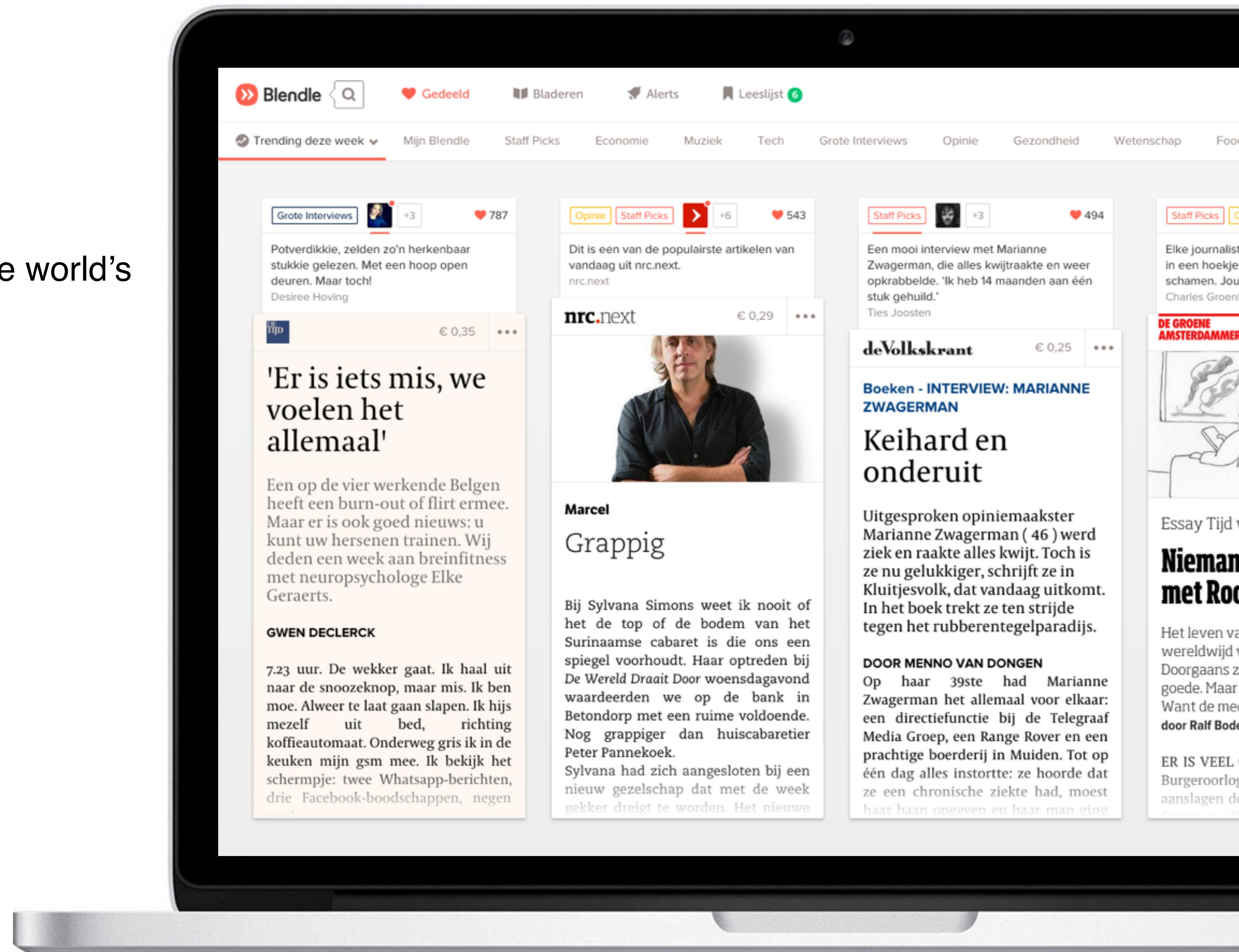
articles, issues, subscriptions

## ✓ International

May 2014 : The Netherlands

Sept 2015 : Germany

March 2016 : United States



# Blendle

## ✓ 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

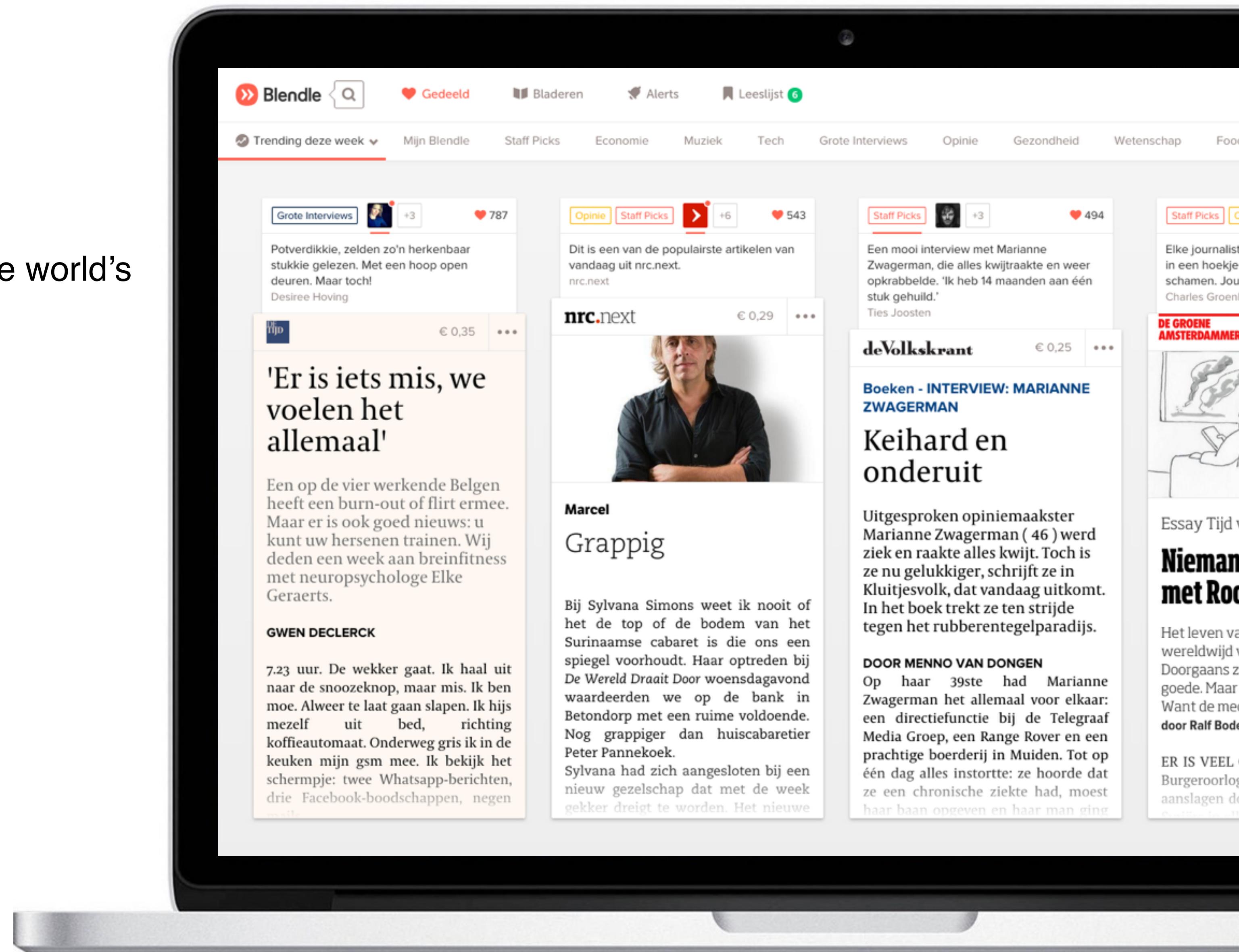
## ✓ Numbers

> 1M users

1 in 5 is paying

70 employees

(15 journalists, 50 developers,  
+ 30 clippers)



In Blendle you can  
browse through all  
quality  
newspapers and  
magazines

The Blendle app interface shows two magazine covers side-by-side. On the left is the front page of NRC Handelsblad, dated Thursday April 7, 2016. The main headline reads "GEENPEEL" and "Simpel nee, complex antwoord". Below it, a sub-headline says "Referendum De kiezer zegt opnieuw nee tegen Europese samenwerking, maar waarom? De politiek heeft geen idee." On the right is the front page of Vrij Nederland, featuring a large portrait of Elon Musk. The main headline reads "ELON MUSK Hoe de man achter Tesla miljardair werd. En moreel leidsman". The Blendle navigation bar at the top includes "Blended", "Gedeeld", "Bladeren", "Alerts", "Leeslijst", and tabs for "Mijn Edities", "Populair", "Nieuw", "Kranten", "Tijdschriften", "Web", "Engels", "Vlaams", and "Duits".

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



@anneschuth

Blendle Gedeeld Bladeren Alerts Leeslijst

Populair Vrij Nederland van gisteren

Vrij Nederland € 0,89 ...

**ELON MUSK**

## PORTRET DE ONDERNEMER ALS MOREEL LEIDER

*Ja, er zit geld in moraal. Elon Musk presenteerde de Tesla Model 3, de coole en betaalbare elektrische auto die de wereld moet veranderen, en hij wil de mensheid redden door Mars te koloniseren. Het heeft hem geen windeieren gelegd, dat redden van de mensheid.*

SANDER PLEIJ | PORTRET DAN WINTERS

Hoeveel tijd kost een vrouw? Zo luidde de vraag van Elon Musk, directeur van Tesla (iPhone onder de elektrische auto's) en SpaceX (rakettenfabriek die de mensheid naar Mars wil emigreren); een rijke vent die het beste met de mensheid voor heeft.

Hij zat met een probleem, zo vertrouwde hij zijn biograaf toe, hij was net voor de tweede keer gescheiden en zocht nu een nieuwe vriendin, maar hij vreesde dat een object van liefde hij

**DUURZAMER Vervoer** De onthulling van Model 3 van de Tesla. Binnen een dag waren er 100.000 reserveringen.

**HOE INSTEED HIJ LEVEN?** Hij zocht steun bij de grote filosofen, maar kwam van een kouwe kernvis thuis.

Tik la Schopenhauer en Nietzsche en het was vreselijk. Niemand zou hem moeten lezen. Ved te deprimerend. Dat waren geen gelukkige mensen.

Eigenlijk anders: Als we streven naar verificatie, kunnen we beter beginnen welke vragen we moeten stellen over het universum. It seems like there's a fundamental god in that.

Het fundamenteel profiel voor de mens. Niet minder. Daarop zullen zijn doelen zijn gebaseerd.

**MUSK VERHAALT EROVER IN DE BIOGRAFIE**

Hij stelt zichzelf de vraag: wat is de voorliefde van de mensheid? En dan valt open: op dat hij er ook een moeite dimensie mee betreft.

Hij zegt: 'Het enige alertheid om naar te streven, is een grotere collectieve verlichting.'

Eigenlijk anders: Als we streven naar verificatie, kunnen we beter beginnen welke vragen we moeten stellen over het universum. It seems like there's a fundamental god in that.

Het fundamenteel profiel voor de mens. Niet minder. Daarop zullen zijn doelen zijn gebaseerd.

**PERMANENTE EXPANSIE**

Verder. Op zijn negen krijgt hij zijn eerste Commodore computer. Op zijn twaalfde vindt hij een prijs met een zelf geschreven computerspel. Op zijn veertienste verlaat hij het geweldadige Zuid-Afrika voorgoed. Met 100 dollar trekt hij via Canada (die moeder is Canadees) naar Amerika en gaat voor de Amerikaanse droom. Go West young man, go West. Voorbij de grenzen van het Westen, voorbij de grenzen.

**RENTIEREN?**

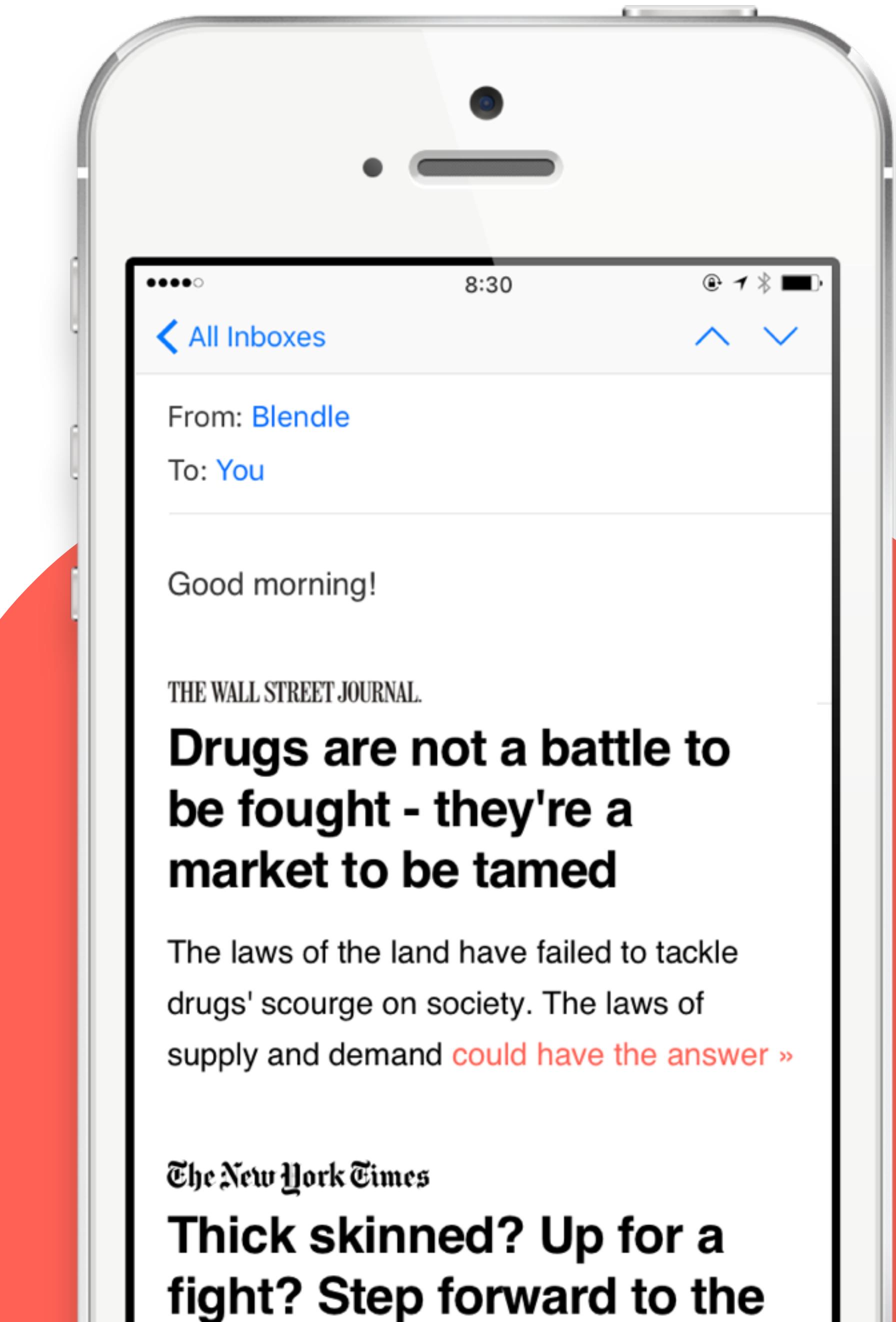
Met zijn broer begint hij socjet Gou des God op in klassieke start-upverhaal maar werken, geen huis kantoor en slak voor de i in 1999 knoopt Compaq zijn lange Musk schrift zijn een bij. Renteren? Hij probeert intern etbank op te richten Pal, dat wordt gekocht door vrienden en te maken. Niks aantrekkelijk lever heel grote fabrieken met alles en iedereen vlak bij elkaar want dat betekent niet wachten, geen tijdelijke, directe communicatie. Binnenkort open hij de Gigafactory, een nieuw batterijenfabrik met de grootste omtrek ter wereld en zal giftig uitstoot.

Allerlijkt onconventioneel: geen reclame maken - met de skeletische reden dat reclame manipuleert en onerijk is. Auto's moeten niet jaarlijks met een nieuw model komen, nee, het huidige model wordt constant beter - als je wilt wachten op een nieuw model kun je voor altijd wachten. Bovendien ondagt de Tesla continu software updates. Iets als automatisch van rijbaan verwisselen is een nachtelijke download. Over vijf jaar, denkt Musk, zijn die zover dat de Tesla's zichzelf kunnen rijden.

Ook onconventioneel: raketten maken en lanceren vanuit het afgelegen atol Kwajalein in de Grote Oceaan. Alle klanen hij ook en ondernemt hij zelf. De technische handboeken over ruimtevaart leest hij zelf. Dat de werkzaamheden is, mag dat mogelijk zijn, maar hij blijft nooit met hij een dagdag.

**EXCENTRISCHE CELEBRITY** Musk met zijn tweede echtgenote, filmster

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



# Data at Blendle

- Articles
  - > 3.5M in total
  - > 6K new every day

# Data at Blendle

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

# Data at Blendle

- Articles
  - > 3.5M in total
  - > 6K new every day
- Users
  - > 1M users
  - ~ 1 in 5 converts to paying user
- Events
  - > 800M in total
  - > 2M new every day

# Data at Ble

- Articles

**> 3.5M in total**

**> 6K new every day**

- Users

**> 1M users**

**~ 1 in 5 converts to paying**

- Events

**> 800M in total**

**> 2M new every day**

The screenshot shows the Blendle website interface. At the top, there's a navigation bar with links like 'Gedeeld', 'Bladeren', 'Alerts', and 'Leeslijst'. Below the navigation, there are tabs for 'Trending vanavond', 'Mijn Blendle' (which is active), 'Staff Picks', 'Opinie', and 'Wat lees je graag?'. The main content area displays a news feed with articles from 'nrc.next' and 'Knack'. The 'nrc.next' article is titled 'Bluesheld Seasick Steve loog over' and has a price of € 0,29. The 'Knack' article is titled 'DOCUMENT Hoe ik. Intisar.' and has a price of € 0,59. To the right, there are sections for 'Staff Picks' and 'Opinie'. A sidebar on the left shows event logs, and a developer tools panel at the bottom shows an XHR request for a timeline view.

```
events
Name Path
events events.blendle.com/v1
events events.blendle.com/v1

x Headers Preview Response Timing
0: {type: "URL Change", session_id: "69fe8e3d-e99e-4188-b0f2-531a440d14ff", origin: "blendle.com",...}
1: {type: "View Timeline", session_id: "69fe8e3d-e99e-4188-b0f2-531a440d14ff", origin: "blendle.com",...}
  client_version: "9c1e16e"
  origin: "blendle.com"
  payload: {timeline: "following", country: "NL", locale: "nl_NL", current_uri: "/following",...}
    AA_EXPERIMENT: "AA_VARIANT_1"
    DeeplinkSignupFB: "DeeplinkSignupFBControl"
    ShorterOnboardingFlow: "OnboardingSkipChannelsAbout"
    ab_tests_ids: "AA_EXPERIMENT, ShorterOnboardingFlow, DeeplinkSignupFB"
    client_timestamp: "2016-06-09T21:34:18+02:00"
    country: "NL"
    current_uri: "/following"
    locale: "nl_NL"
    previous_uri: "/i/vrij-nederland/de-boodschap-is-ik-ben-wit-dus-ik-ben-superieur/bnl-vn-20160608-28991/refund"
    session: "69fe8e3d-e99e-4188-b0f2-531a440d14ff"
    timeline: "following"
    tracking_uid: "a8bdf675-2813-4519-b125-431c044752d2"
    session_id: "69fe8e3d-e99e-4188-b0f2-531a440d14ff"
    type: "View Timeline"
```

# 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)



## Movies and series

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



## Articles

- **Single** consumption

# 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

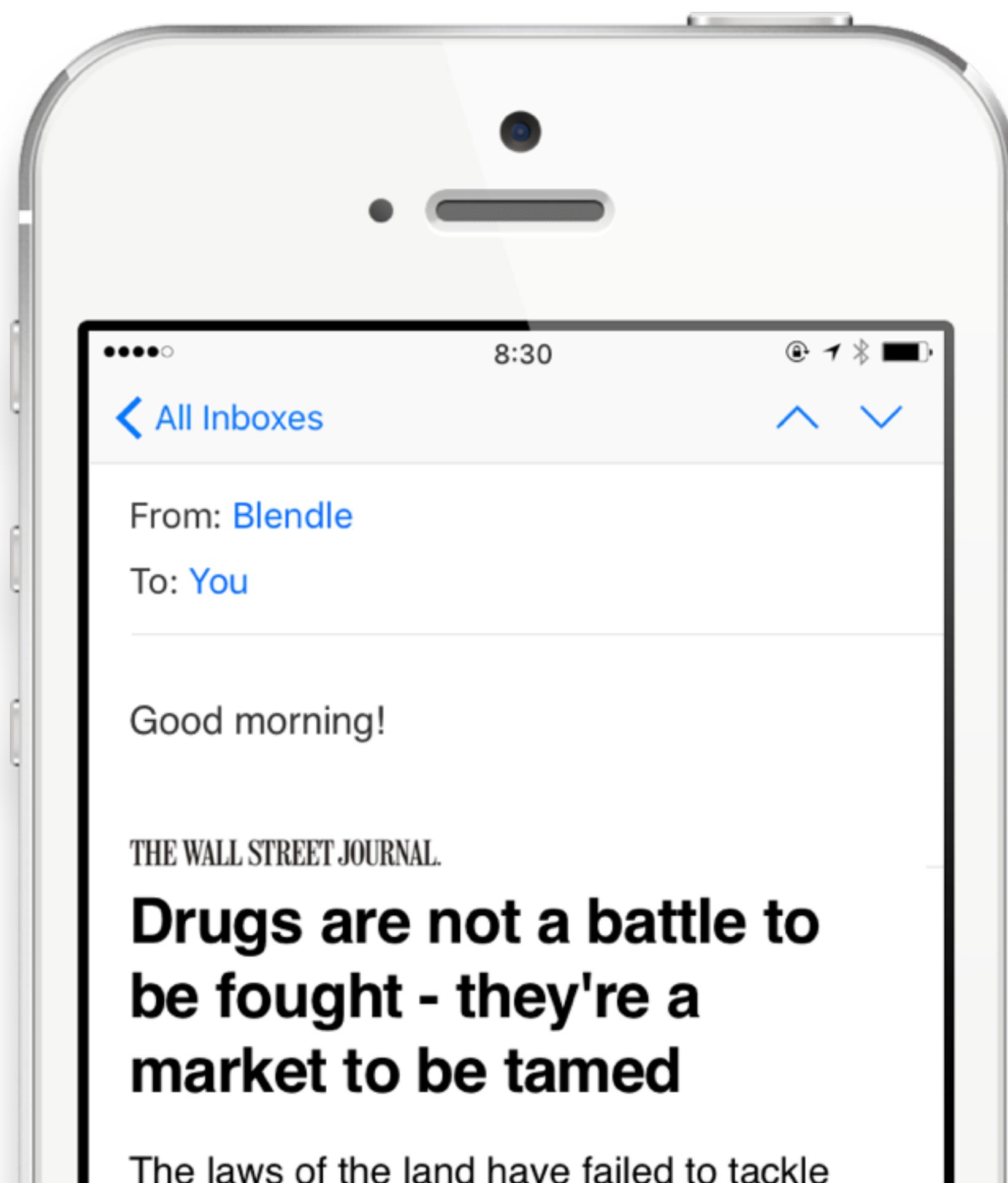
- **Single** consumption
- Less niche than music
- 75.000.000 users
- **Long** shelf life
- Feedback: explicit ratings



## 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

# Rankings at Blendle



# Rankings at Blendle

The image shows a smartphone on the left displaying a news article from Blendle, specifically the 'All Inbox' section. On the right is a screenshot of a web browser window titled 'Blindle Widgets for Developers' at the URL <https://widgets.blindle.nl/developers/>. The browser window is divided into two main sections: 'Settings' on the left and 'Preview' on the right.

**Settings:**

- Widget type:  Widget with popular articles
- Widget with articles by search query
- Widget with articles shared by user
- Widget with singular publication
- Widget with quote from singular article
- Widget with button to article

Amount of articles:

From publications:  All publications

Line template: `<strong>{{item:provider}}</strong> {{item:title}}`

Use the following variables in the line template:

- `{{item:title}}` Article title
- `{{item:snippet}}` Article snippet
- `{{item:provider}}` Article provider
- `{{item:index}}` Ascending number per article, starting at 1
- `{{item:date}}` Publication date
- `{{item:price}}` Article price
- `{{item:image}}` Article image

**Preview:**

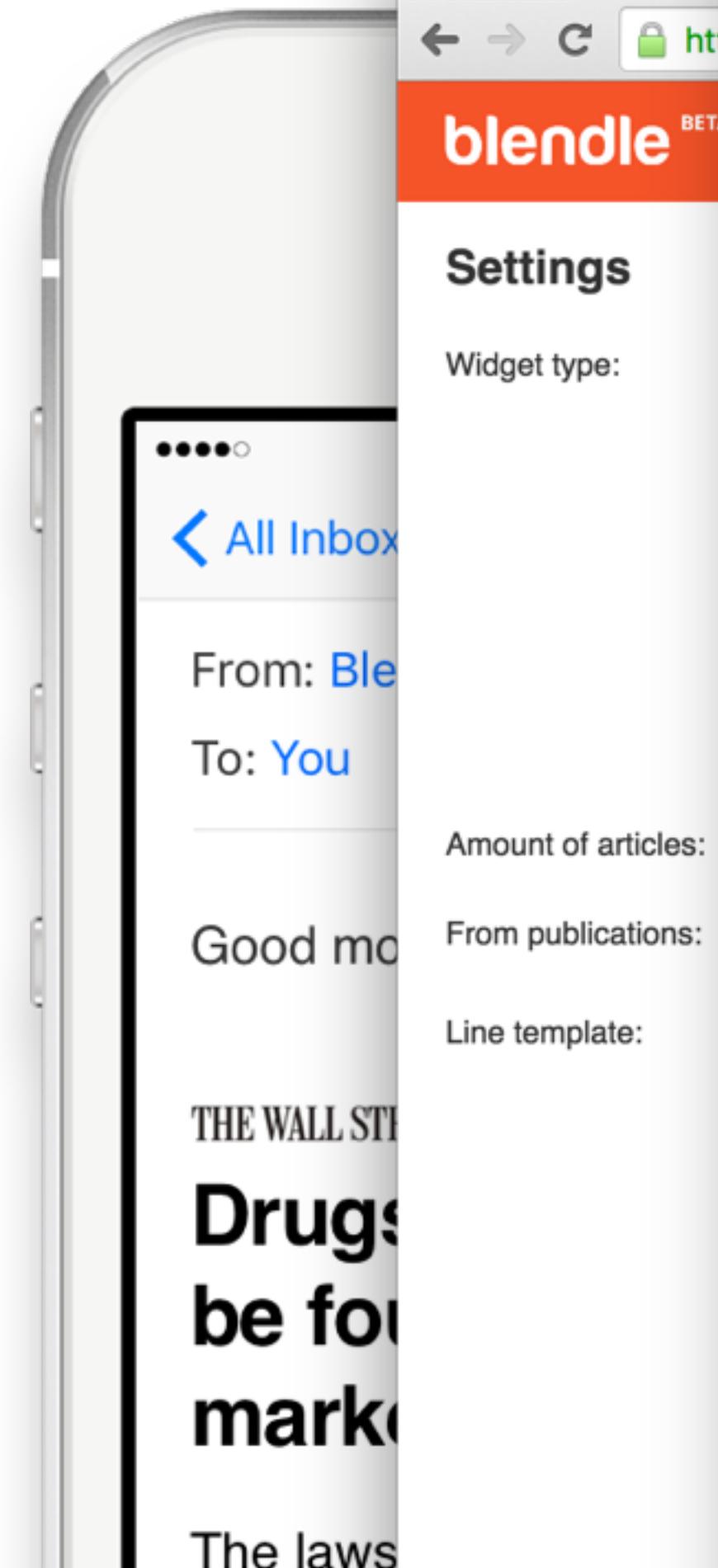
### Populair op Blindle

1. [NRC Handelsblad Edelachtbare](#)
2. [NRC Handelsblad Berouw](#)
3. [Humo 'Ellende en tegenslag achter je kunnen laten en doorgaan: dát is het geheim'](#)
4. [Trouw De vrouw van een pedoseksueel](#)
5. [Vrij Nederland 'De boodschap is: ik ben wit, dus ik ben superieur'](#)

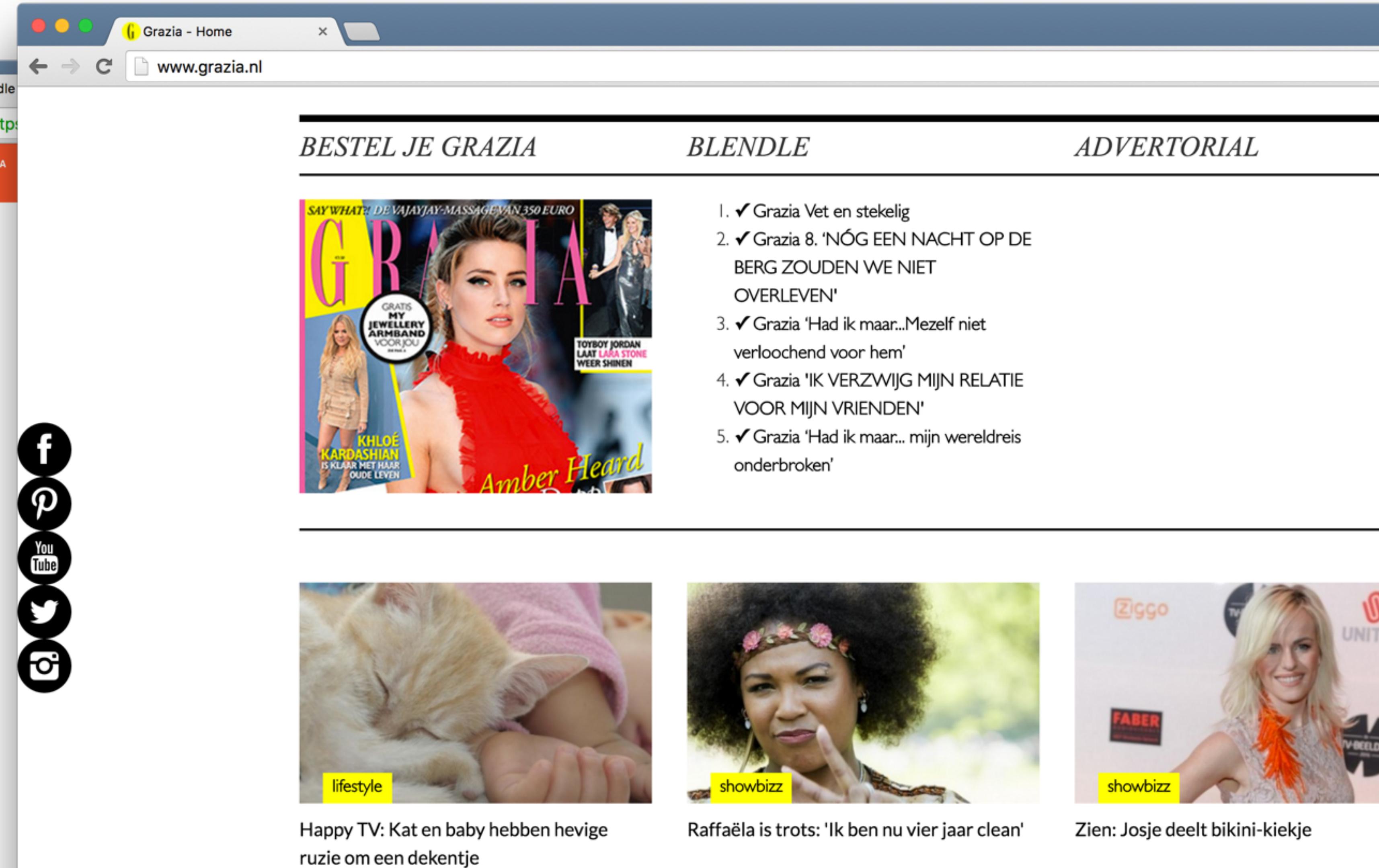
`<div class="blindle-widget" data-widget-type="popular" data-widget-item-template="`

Copy and paste the code shown above into the desired place of the HTML of your website.

# Rankings at Blendle



The Blendle mobile application interface is shown. It features a sidebar with navigation options like 'All Inbox', 'From: Ble', 'To: You', 'Good mor...', 'THE WALL ST...', 'Drugs be fo...', and 'market'. Below the sidebar are sections for 'Amount of articles:', 'From publications:', and 'Line template:'. A vertical column of social sharing icons for Facebook, Pinterest, YouTube, and Twitter is also present.



A screenshot of the Blendle website showing the Grazia magazine cover. The cover features a woman in a red dress and includes headlines such as 'SAY WHAT? DE VAJAJAY-MASSAGE VAN 350 EURO', 'GRATIS MY JEWELLERY ARMBAND VOOR JOU', 'TOYBOY JORDAN LAAT LARA STONE WEER SHINEN', 'KHAOÉ KARDASHIAN IS KLAAR MET HAAR OUDLEVEN', and 'Amber Heard'. The website layout includes sections for 'BESTEL JE GRAZIA', 'BLENDLE', and 'ADVERTORIAL'. The 'BLENDLE' section lists five items from the magazine, each with a checkmark and a short description. Below the magazine cover are three smaller news snippets: 'lifestyle' (Happy TV: Kat en baby hebben hevige ruzie om een dekentje), 'showbizz' (Raffaëla is trots: 'Ik ben nu vier jaar clean'), and 'showbizz' (Zien: Josje deelt bikini-kiekje).

**BESTEL JE GRAZIA**

**BLENDLE**

**ADVERTORIAL**

- ✓ Grazia Vet en stekelig
- ✓ Grazia 8. 'NÓG EEN NACHT OP DE BERG ZOUDEN WE NIET OVERLEVEN'
- ✓ Grazia 'Had ik maar...Mezelf niet verloochend voor hem'
- ✓ Grazia 'IK VERZWIJG MIJN RELATIE VOOR MIJN VRIENDEN'
- ✓ Grazia 'Had ik maar... mijn wereldreis onderbroken'

**lifestyle**

Happy TV: Kat en baby hebben hevige ruzie om een dekentje

**showbizz**

Raffaëla is trots: 'Ik ben nu vier jaar clean'

**showbizz**

Zien: Josje deelt bikini-kiekje

# Rankings at Blendle

The image displays a collage of three devices showcasing Blendle's ecosystem:

- Mobile App (iPhone):** Shows the Blendle app interface with a sidebar for "All Inbox" and various sharing options (Facebook, Pinterest, YouTube, Twitter, Instagram) for a news item from "THE WALL STREET JOURNAL" about "Drugs be found market".
- Web Browser Extension (Grazia):** Shows the Blendle extension installed in a browser window, with a sidebar for "All Inbox" and sharing options.
- Web Browser (Blende.com):** Shows the Blendle website with a navigation bar including "Trending vanavond", "Mijn Blendle" (highlighted), "Staff Picks", "Opinie", and "Wat lees je graag?". The main content area features news items from "nrc.next" and "Knack".

**nrc.next** (€ 0,29)

Bluesheld Seasick  
Steve loog over  
leeftijd en  
levensverhaal

**BIOGRAFIE** Blueslegende  
Seasick Steve heeft hele delen  
van zijn tumultueuze  
levensverhaal verzonnen, zo  
blijkt uit een nieuwe biografie

Door onze medewerker **Jan Vollaard**

Seasick Steve verzon grote delen van zijn ruige levensverhaal. Een nieuwe biografie over de Amerikaanse bluesmuzikant bevat steekhoudend bewijs dat de zanger en gitarist bij zijn internationale doorbraak heeft gelogen over zijn leeftijd, zijn echte

**Knack** (€ 0,59)

DOCUMENT

**Hoe ik, Intisar, radicaliseerde**

Hoe wordt een vrolijke Antwerpse tienerdochter een soldaat van de jihad? En wat doet haar uiteindelijk toch van gedachte veranderen? Intisar Umm Mansur vertelde haar verhaal aan jihadexpert Montasser AlDe'emeh. Een voorpublicatie

Wow. Na deze brief van een 23-jarig verkrachtingsslachtoffer ben je even stil.  
Ties Joosten

NRC  
HANDELSBLAD (€ 0,39)

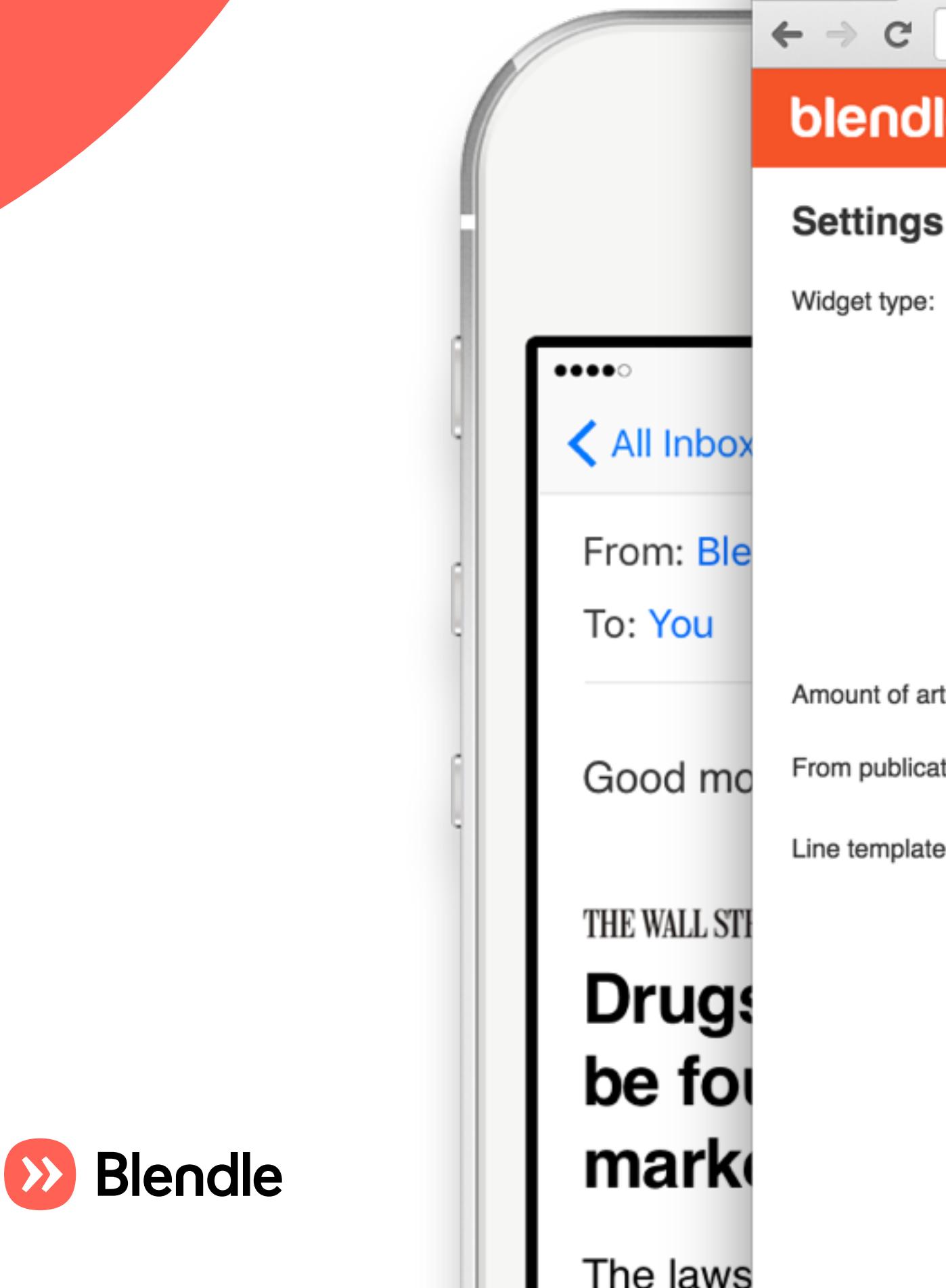
Edelachtbare,  
**Aanranding**  
Het 23-jarige slachtoffer van een verkrachting op Stanford University in de VS schreef een brief aan de rechter, over wat haar is overkomen. De dader, een veelbelovend zwemmer van 20, kreeg zes maanden cel.

# Rankings at Blendle

The image is a collage of several screenshots from the Blendle platform, illustrating its various features and user interface elements.

- Top Left:** A screenshot of a mobile device displaying the Blendle app's inbox screen. It shows a list of items with titles like "All Inbox", "From: Ble", "To: You", and "Good mo". Below the list are social sharing icons for Facebook, Pinterest, YouTube, Twitter, and Instagram.
- Middle Left:** A screenshot of the Blendle website's "Settings" page. It includes sections for "Widget type:", "Amount of articles:", "From publications:", and "Line template:". To the right of these settings are two columns of news cards. The first column features "nrc.next" and "Bluesher". The second column features "LINDA. meiden" (with a red heart icon showing 143 likes) and "Steve lo".
- Bottom Left:** A screenshot of a news article from "Drugs be fo". The headline reads "Drugs be fo market". Below the headline is a snippet: "The laws".
- Top Right:** A screenshot of a web browser window titled "Blendle - Blader online doo". The URL is <https://blendle.com/trending/now>. The page displays a "Trending vanavond" section with a red underline. Other menu items include "Mijn Blendle", "Staff Picks", "Opinie", and "Wat lees je graag?". Below the menu, there are several news cards. One card for "LINDA. meiden" has a red heart icon showing 143 likes. Another card for "NRC HANDELSBLAD" has a red heart icon showing 472 likes. A third card for "Wow. Na deze brief van een 23-jarig verkrachtingsslachtoffer ben je even stil." by Ties Joosten is also visible.
- Middle Right:** A screenshot of a news article from "Edelachtbare, Aanranding". The headline reads "Edelachtbare, Aanranding". Below the headline is a snippet: "Het 23-jarige slachtoffer van een verkrachting op Stanford University in de VS schreef een brief aan de rechter, over wat haar is overkomen. De dader, een veelbelovend zwemmer van 20, kreeg zes maanden cel. Met uw goedvinden zou ik me met deze verklaring graag rechtstreeks tot".
- Bottom Right:** A screenshot of a news article from "GLORIA WEKKER I OV GELIJKHEID". The headline reads "INTERVIEW 'DE BOODS IK BEN WI BEN SUPE". Below the headline is a snippet: "400 jaar kolonie na in hoe je na ander kijkt, zegt hoogleraar genetica Gloria Wekker. 'White Innocence' waarom 'ras' in reactie opropt. altijd, maar steeds vinden het normaal zijn. Geef mij maar correctheid.'

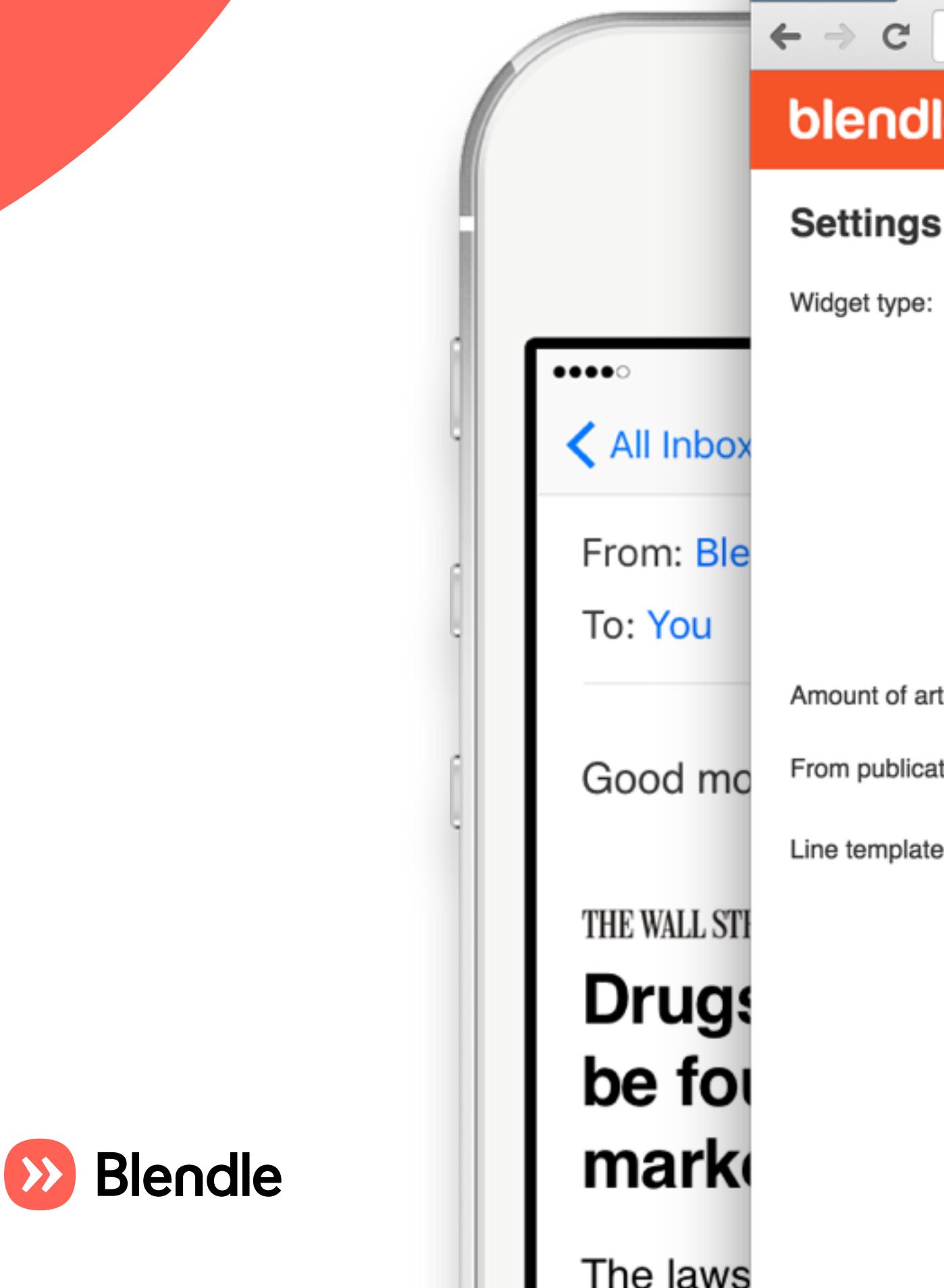
# Rankings at Blendle



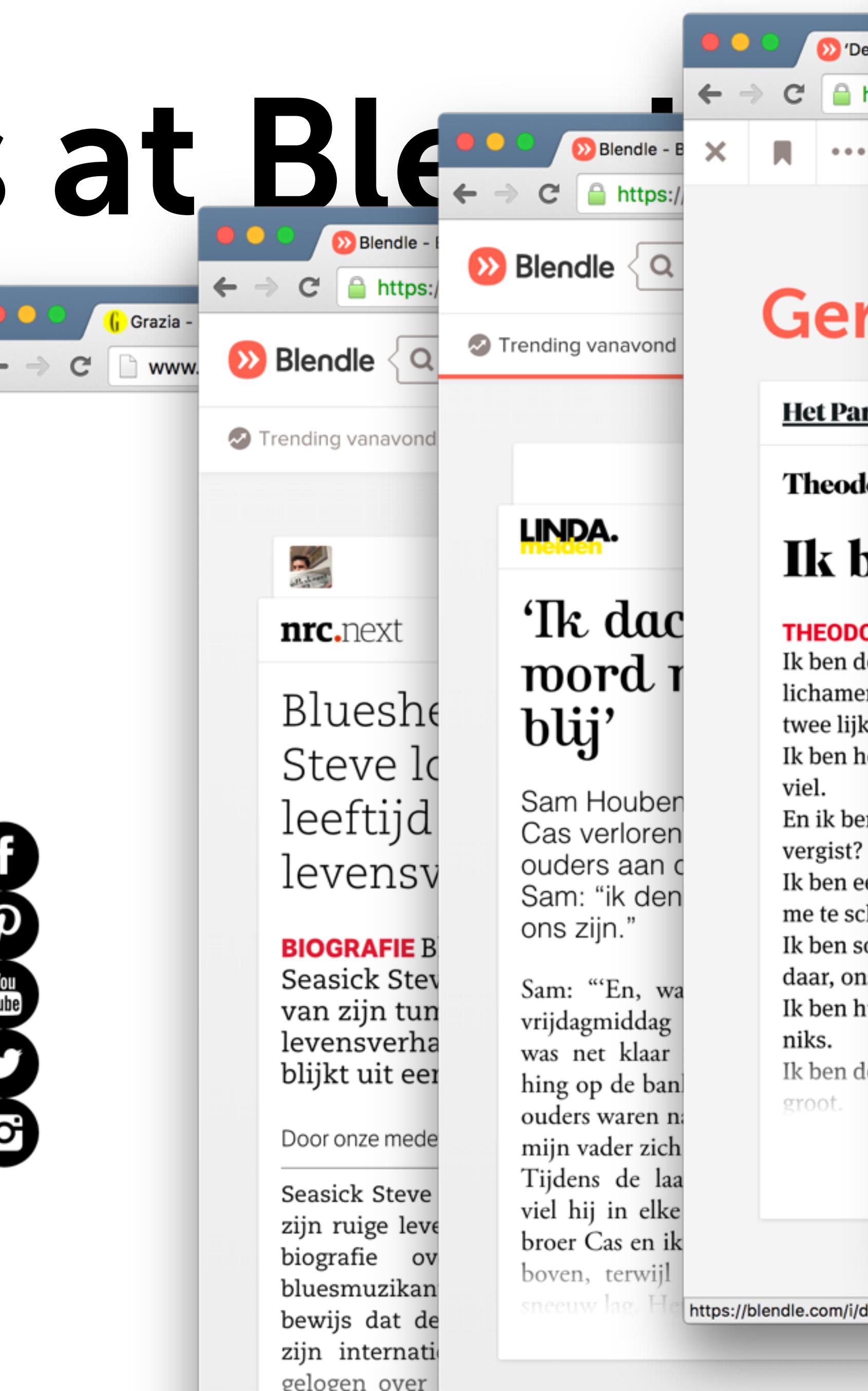
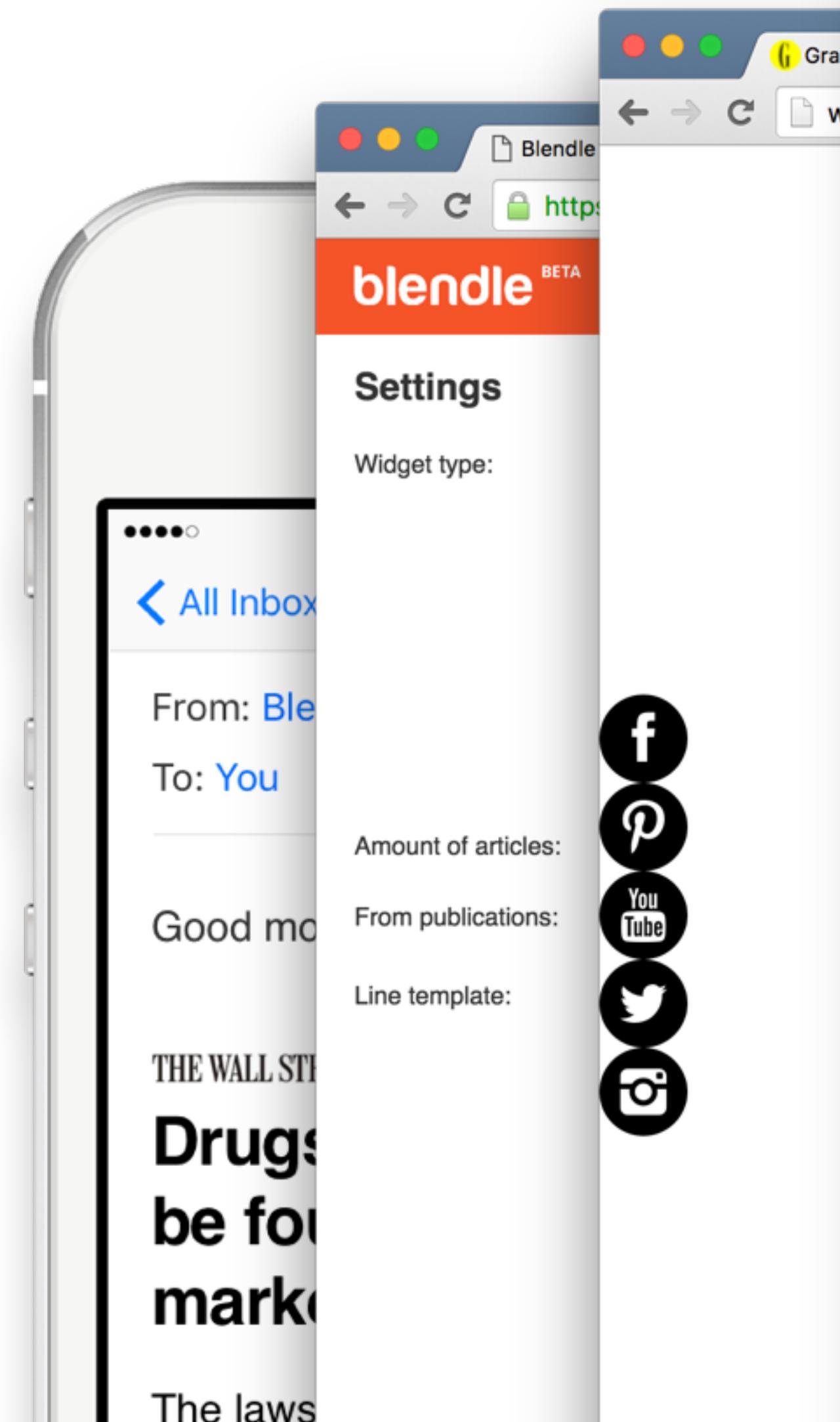
The image shows a desktop environment with multiple windows. In the foreground, a web browser has several tabs open. The active tab is for Blendle, showing a 'Trending vanavond' section. Other tabs include 'Grazia', 'www.', and another Blendle tab. To the right of the browser, there's a news article from 'nrc.next' by 'LINDA. meiden' with the title 'Ik dac mord n blij'. Below it is a 'BIOGRAFIE' section about Seasick Steve.

The image shows a desktop environment with a single browser window. The page displays a news article from 'De Telegraaf' with the title 'Ik ben wie ik ben'. Above this, there's a section titled 'Gerelateerde artikelen' with links to articles from 'Het Parool' and 'De Telegraaf', both featuring 'Theodor Holman'.

# Rankings at Blendle



Blendle



Blendle

https://blendle.com/search/data

Gedeeld Bladeren Alerts

Artikelen

Zoeken in alle kranten en tijdschriften.

data vandaag

Wegens succes: extra data tocht over de Mark

BN DeStem

**data**

Alle perioden

Nederlands

Toon resultaten

Voeg toe als alert

**Tip:**  
Gebruik dubbele aanhalingstekens " om woorden te combineren.  
Gebruik - om een woord uit te sluiten.

[https://blendle.com/i/de-standaard/doping-ik-ben-gelukkig-ik-ben-clean/bnl-destandaard-20150829-57034224\\_4da6\\_11e5\\_a21a](https://blendle.com/i/de-standaard/doping-ik-ben-gelukkig-ik-ben-clean/bnl-destandaard-20150829-57034224_4da6_11e5_a21a)

deze verklaring graag rechtstreeks tot

TEKST GRETA RIE

**Gere**

**Het Parool**

**Theodor H**

**Ik ben**

**THEODOR H**

Ik ben de boel lichamen zag twee lijken o. Ik ben het lic veel. En ik ben reb vergist? Ik ben een ra me te scham. Ik ben soldaa daar, ons tre. Ik ben hun vniks. Ik ben de pre groot.

Een dagtocht per schip Brabant, inclusief uitgewarmer snacks. Binnen twee weken zijn vaartochten - met ruim uitverkocht. Omdat er zoveel animo in samenwerking met meerdere extra vaartuiken kunnen de leden van onder andere op woensdag 18 juni het rui

# Rankings at Blendle

# Rankings at Blendle

- Should be personalised

# Rankings at Blendle

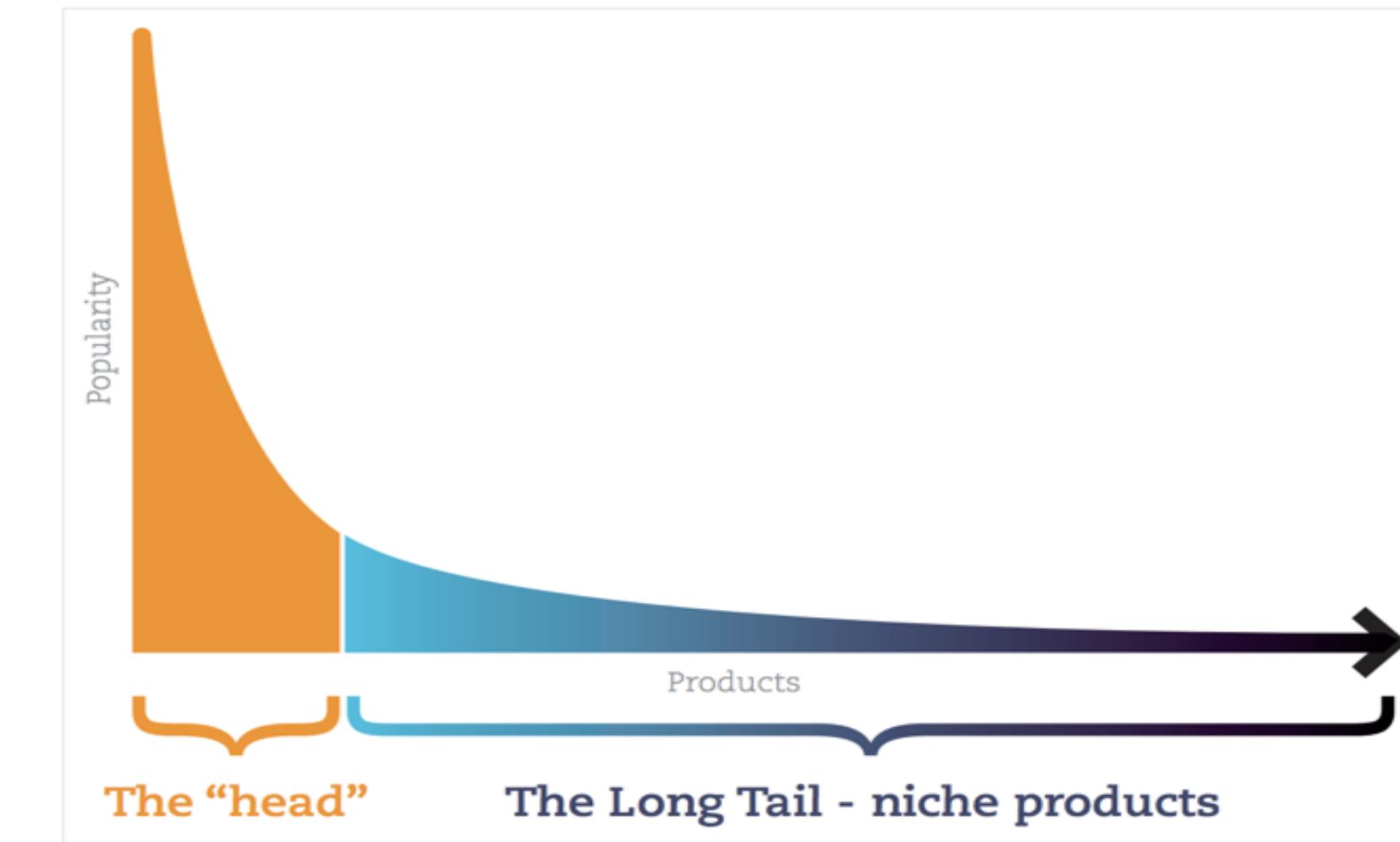
- Should be personalised
  - Lots of content

# Rankings at Blendle

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

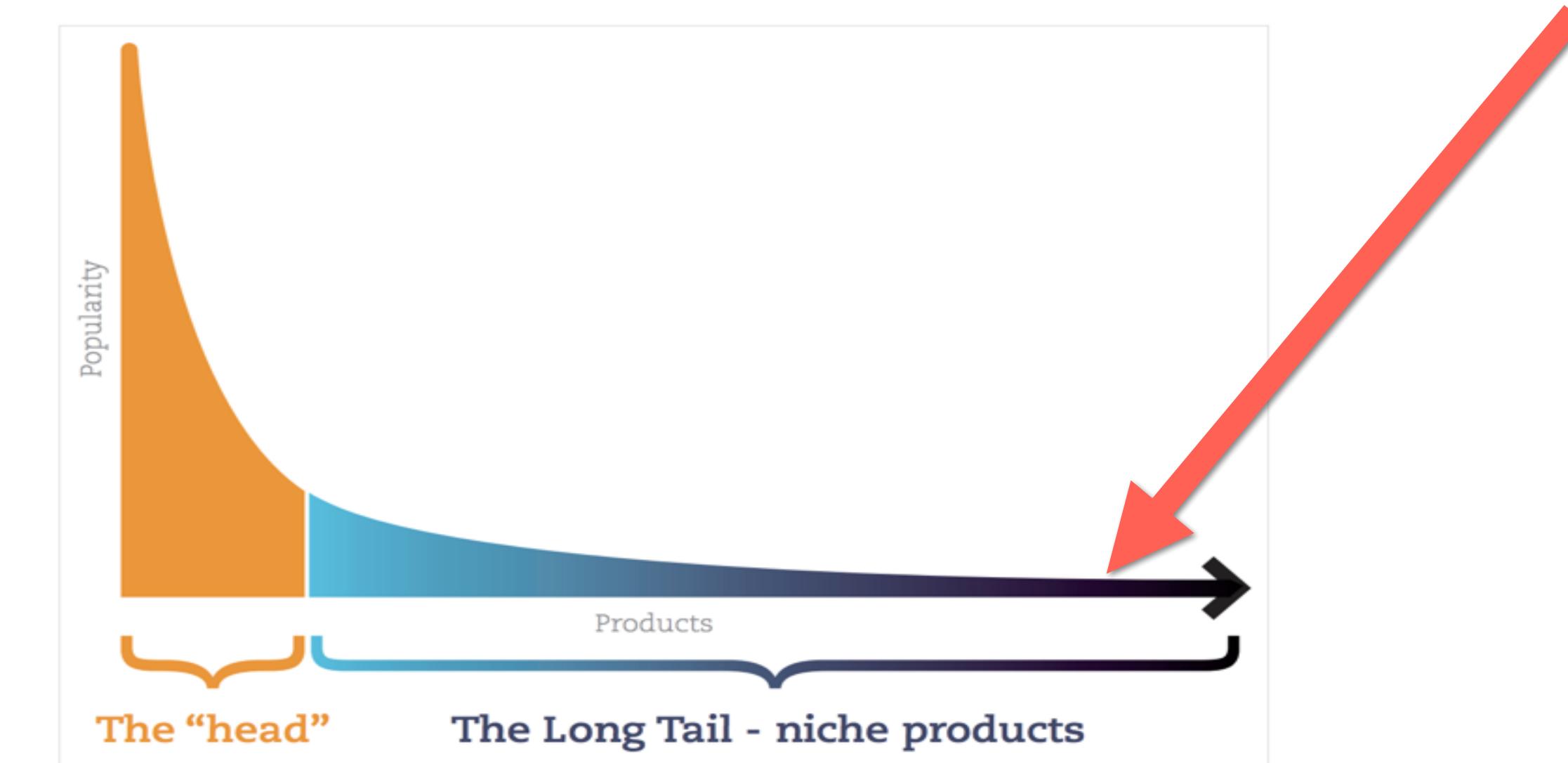
# Rankings at Blendle

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



# Rankings at Blendle

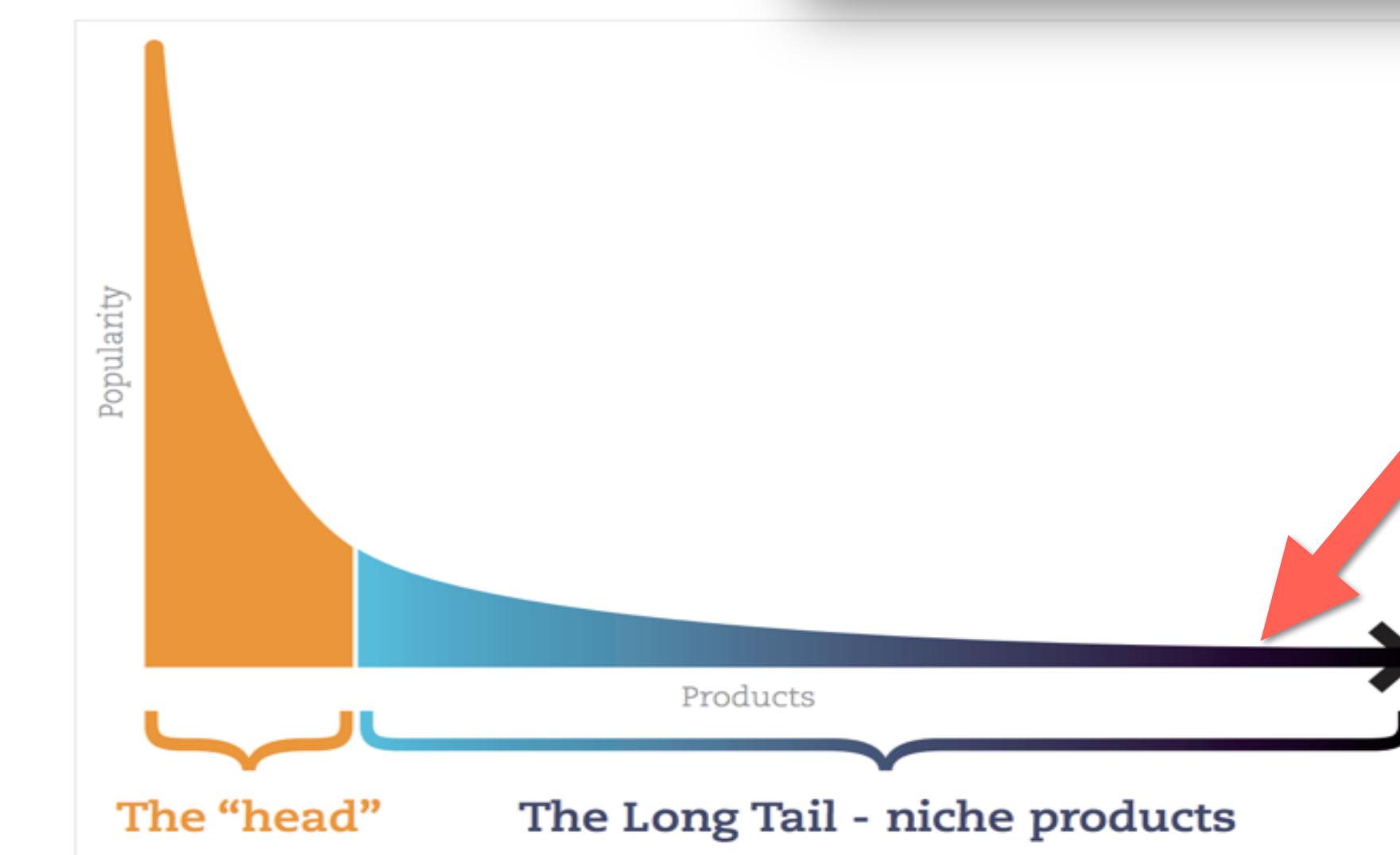
- Should be personalised
  - Lots of content
  - Lots of different tastes



# Rankings at Blendle

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

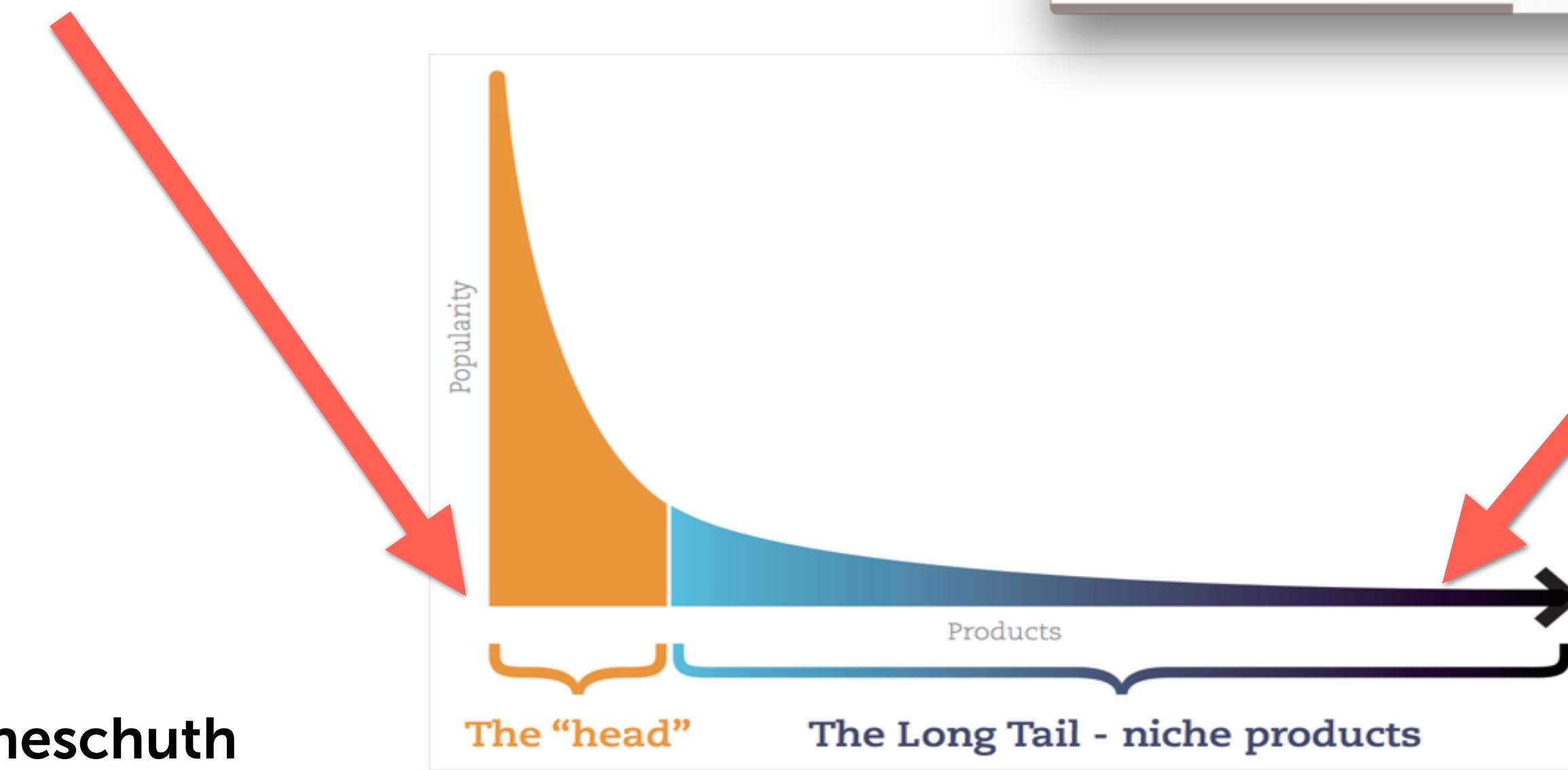
The screenshot shows a web browser window with the URL <https://blendle.com/i/leeuwarder-courant/colofon/bnl-ic-20160603-8G7SPMQ>. The page title is "COLOFON". It contains information about the editorial team: Eindredactie Gerda Douma, Wiebe van der Hout, Roel Snijder (chef), Theo Zandstra, Janna Zuiderveld; Vormgeving Erik Hilverda; Art director Alie Veenhuizen; Telefoon cultuurredactie 058 - 2845395; and E-mail cultuur@lc.nl. There is also a "Goed stuk" button with a heart icon. To the right, there's a sidebar with social sharing links (Facebook, Twitter, LinkedIn) and a "Delen op de interwebs" section. Below that is a "Mail dit artikel naar een vriend" section with an input field for the recipient's email address. At the bottom right, there's a "Lees editie" button.



# Rankings at Blendle

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

The screenshot shows a web browser window with the URL <https://blendle.com/i/leeuwarder-courant/colofon/bnl-ic-20160603-8G7SPMQ>. The page title is "COLOFON". It contains information about the editorial team: Eindredactie Gerda Douma, Wiebe van der Hout, Roel Snijder (chef), Theo Zandstra, Janna Zuiderveld; Vormgeving Erik Hilverda; Art director Alie Veenhuizen; Telefoon cultuurredactie 058 - 2845395; and E-mail cultuur@lc.nl. There is also a link to "Goed stuk". To the right, there's a sidebar with social sharing buttons for Facebook, Twitter, and LinkedIn, and sections for "Delen op de interwebs" and "Mail dit artikel naar een vriend". A small thumbnail of the newspaper is visible.



LINDA.  
woensdag 20 januari 2016 - 5 minuten

# Lustige dames

Tv-vrouw Geraldine Kemper (25) en model Kim Feenstra (30) rollebollen graag met vrouwen, maar willen de mannen niet missen. "Als het maar geen softie is."

Tekst Eva Loesberg, fotografie Marc De Groot, styling Jos Van Heel, visagie Irena Ruben

**VROUWEN**

*Kim* "Ik val op mannen én vrouwen. Daar ben ik kort geleden pas eerlijk voor uitgekomen. Op mijn achttiende heb ik een lesbische relatie gehad."

*Geraldine* "Ik heb nog nooit een relatie gehad met een vrouw, maar ik speel wel met vrouwen."

**RODDEL**

*Kim* "Sinds ik in *Spuiten en Slikken* heb laten merken dat ik Geraldine heel aantrekkelijk vind, gaat het gerucht dat er iets tussen ons is gebeurd."

*Geraldine* "Maar daar gaan we verder niets over zeggen. Sommige dingen blijven privé."

*Kim* "Die blijven in de slaapkamer."

*Geraldine* "We doen mee aan deze LINDA.-shoot omdat we fotograaf Marc de Groot geweldig vinden. Met hem weet je

*Kim* "Ja, misschien was die relatie ook meer spelen. Het was sex. Het was ontdekken. Het duurde drie maanden. Vrouwen zijn zo zacht. Fijn om te voelen."

*Geraldine* "Ik vind sex met een vrouw ook fijn. Ik word aangetrokken door een persoon – of het nou een man of een vrouw is kan me niet schelen."

*Kim* "Het hippie-idee staat me aan: love everyone; geen hokjes, vrijheid."

**MODE**

*Geraldine* "We k  
zou ik dat nooit  
genoeg voel ik m  
*Kim* "Of iets sexy  
draagt. Bij jou is  
*Geraldine* "Jij zag  
van Christian Lo  
*Kim* "Ik vond he  
*Geraldine* "Norm  
ruitje. Simpel. C  
Be zweet. In een l  
maxjes. En dan h  
*Kim* "Ik heb een  
Lieslaarzen vind

dat het classy wo  
oplait, vinden w

COLOFON - Leeuwarder Courant

EEL  
afgelopen vrijdag - één minuut

# COLOFON

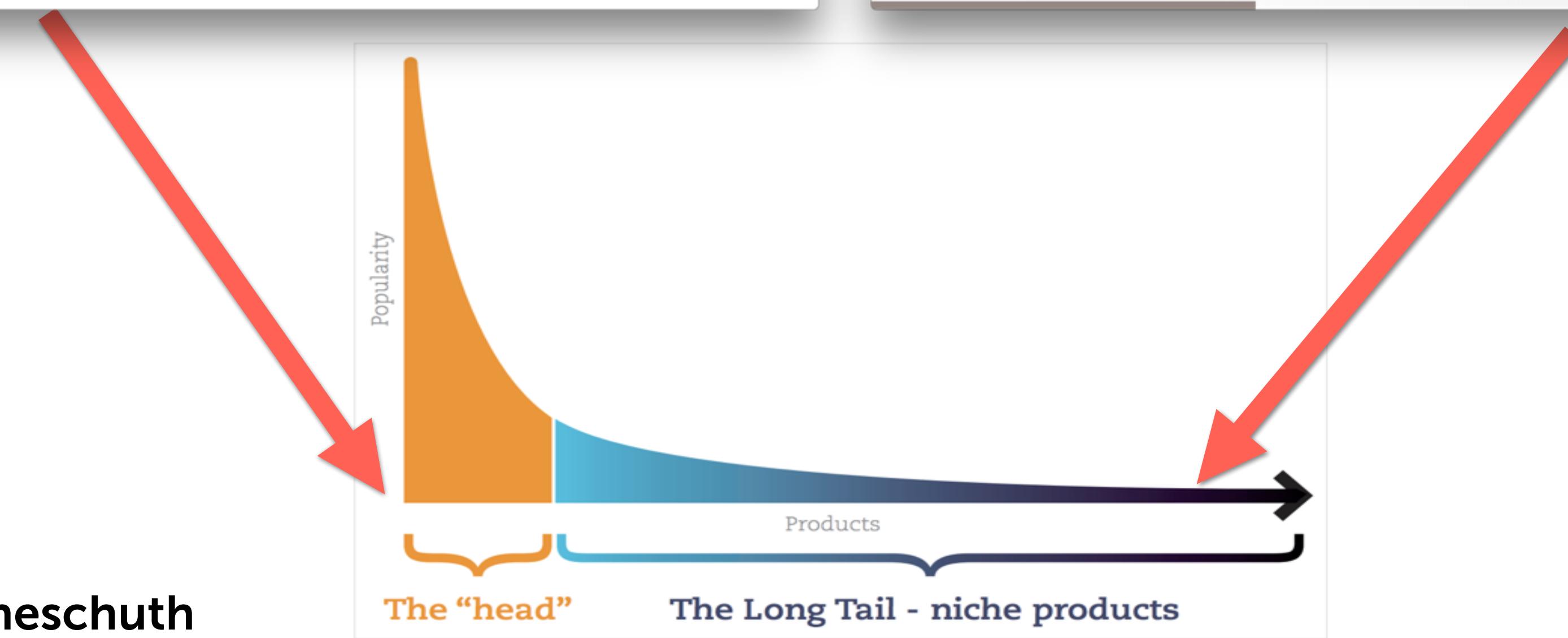
Eindredactie Gerda Douma, Wiebe van der Hout, Roel Snijder (chef), Theo Zandstra, Janna Zuiderveld  
Vormgeving Erik Hilverda  
Art director Alie Veenhuizen  
Telefoon cultuurredactie 058 - 2845395  
E-mail cultuur@lc.nl

Dit artikel komt uit **Leeuwarder Courant** van afgelopen vrijdag.

Meer lezen uit Leeuwarder Courant?  
Hele editie  
Eenmalig € 1,78 Lees editie

Goed stuk

Dit stuk las je voor € 0,01 via Blendle. Viel het tegen? [Vraag je geld terug](#).



# 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

The screenshot shows a web browser window with the Blendle website. The URL in the address bar is <https://blendle.com/i/vrij-nederland/de-boodschap-is-ik-ben-wit-dus-ik-ben-superieur/bnl-vn-20160608-28991>. The page title is 'Blendle - Blader online doe'. The main content is an interview with a person from 'Vrij Nederland' (Free Netherlands). The text discusses the history of racism in the Netherlands, mentioning figures like Simon Schama, Russell Shorto, and Jonathan Franzen. It also touches on the climate debate and political correctness. A red button on the right says 'Goed stuk' (Good piece).

iseerd worden, het begint al bij de grote wordingsgeschiedenis van Nederland. Simon Schama, Russell Shorto en Jonathan Franzen in hun boeken staan hoe het werkelijk hadden gaan over wat de Gouden Eeuw oerderde, de strijd tegen het water, de eedsstrijd tegen Spanje, maar het had de slavenhandel spelen geen rol. Het over Nederland moet veranderen en dan we in het onderwijs iets doen aan vaardigheden: hoe gaan we met elkaar.

*ence' pleit u voor meer etiquette in onze samenleving?*

terugkwam uit de Verenigde Staten, viel onbeschaamdheid op waarmee hier opmerkingen worden gemaakt. Met groot respect voor iemand hier van alles over anderen en dan niet. Het gevoel is: wij kunnen dit doen, niet racistisch. Ik schuif Amerika zeker

niet naar voren als het voorbeeld dat we moeten volgen, maar beledigende opmerkingen worden daar onmiddellijk afgestraft.'

*Als u het klimaat toen al erg vond in Nederland, moet u nu van uw stoel vallen. Nogal wat mensen zeggen: weg met de politieke correctheid.*

'Geef mij maar meer politieke correctheid. Ik vind het ongehoord wat er allemaal kan worden gezegd en dat is alleen maar erger geworden. Racisme was er altijd, maar steeds meer mensen vinden het normaal om grof te zijn.'

*Hoe is het eigenlijk om dit soort opmerkingen aan te moeten horen sinds u een kind was?*

Peinzend: 'Weet je, als ik kan verklaren waarom dit soort dingen wordt gezegd, zoals ik in dit boek heb gedaan, werkt dat verzachtend. Het was heel pijnlijk om het te schrijven, soms heb ik aan mijn bureau zitten huilen, als ik iets over mezelf vertelde of als ik een connectie zag die ik nog niet eerder had gezien. Ik zat er

Dit stuk las je voor €

# 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

The screenshot shows a web browser window with the title 'Blendle - Blader online doe'. The URL in the address bar is <https://blendle.com/i/vrij-nederland/de-boodschap-is-ik-ben-wit-dus-ik-ben-superieur/bnl-vn-20160608-28991>. The page content is in Dutch and discusses the climate in the Netherlands. A context menu is open over some text, showing options: 'Toevoegen aan Pocket', 'Niet goed? Geld terug (€ 0,89)', 'Printen', and 'Uitgave weergeven'. The text on the page includes:

niet naar voren als het voorbeeld dat we moeten volgen, maar beledigende opmerkingen worden daar onmiddellijk afgestraft.'

*Als u het klimaat toen al erg vond in Nederland, moet u nu van uw stoel vallen. Nogal wat mensen zeggen: weg met de politieke correctheid.*

'Geef mij maar meer politieke correctheid. Ik vind het ongehoord wat er allemaal kan worden gezegd en dat is alleen maar erger geworden. Racisme was er altijd, maar steeds meer mensen vinden het normaal om grof te zijn.'

*Hoe is het eigenlijk om dit soort opmerkingen aan te moeten horen sinds u een kind was?*

Peinzend: 'Weet je, als ik kan verklaren waarom dit soort dingen wordt gezegd, zoals ik in dit boek heb gedaan, werkt dat verzachtend. Het was heel pijnlijk om het te schrijven, soms heb ik aan mijn bureau zitten huilen, als ik iets over mezelf vertelde of als ik een connectie zag die ik nog niet eerder had gezien. Ik zat er

Dit stuk las je voor € 0,89

# Rankings a

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

The screenshot shows a web browser window with a news article from 'Vrij Nederland' titled "'De boodschap is: ik ben wit' dus ik ben superieur". The article discusses the history of racism in the Netherlands. A modal window is overlaid on the page, asking "Waarom wil je je € 0,89 terug?" (Why do you want your € 0,89 back?). The modal lists several reasons with radio buttons, and the option "Het artikel was niet wat ik ervan verwachtte" (The article was not what I expected) is selected. Below the list are two buttons: "Ik wil mijn € 0,89 terug" (I want my € 0,89 back) in red, and "Laat maar zitten" (Leave it) in grey.

'De boodschap is: ik ben wit' dus ik ben superieur

Vrij Nederland  
gisteren · 16 minuten

niet naar voren als het voorbeeld dat we moeten volgen, maar beledigende opmerkingen worden door... ook niet op te... ik moest.'

Goed stuk

Waarom wil je je € 0,89 terug?

We willen graag weten waarom je je geld terug wilt, zodat we dit aan de auteur en uitgever kunnen vertellen.

Ik klikte per ongeluk op het artikel

Ik vind de prijs te hoog

Het artikel was te kort

Het artikel was te lang

Het artikel was niet wat ik ervan verwachtte

Het artikel was niet goed leesbaar

Anders, namelijk...

Ik wil mijn € 0,89 terug

Laat maar zitten

# 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

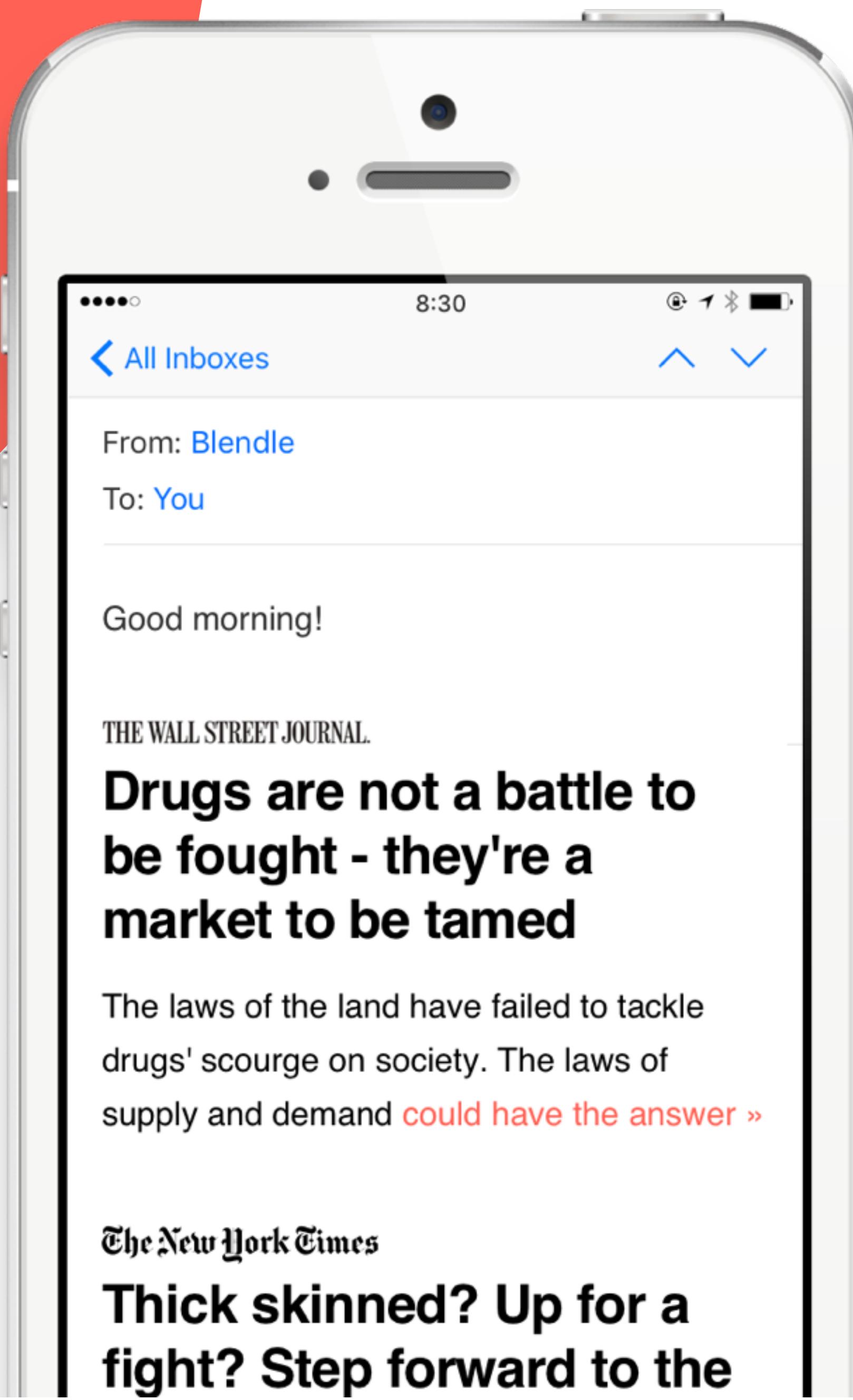
# Rankings at Blendle

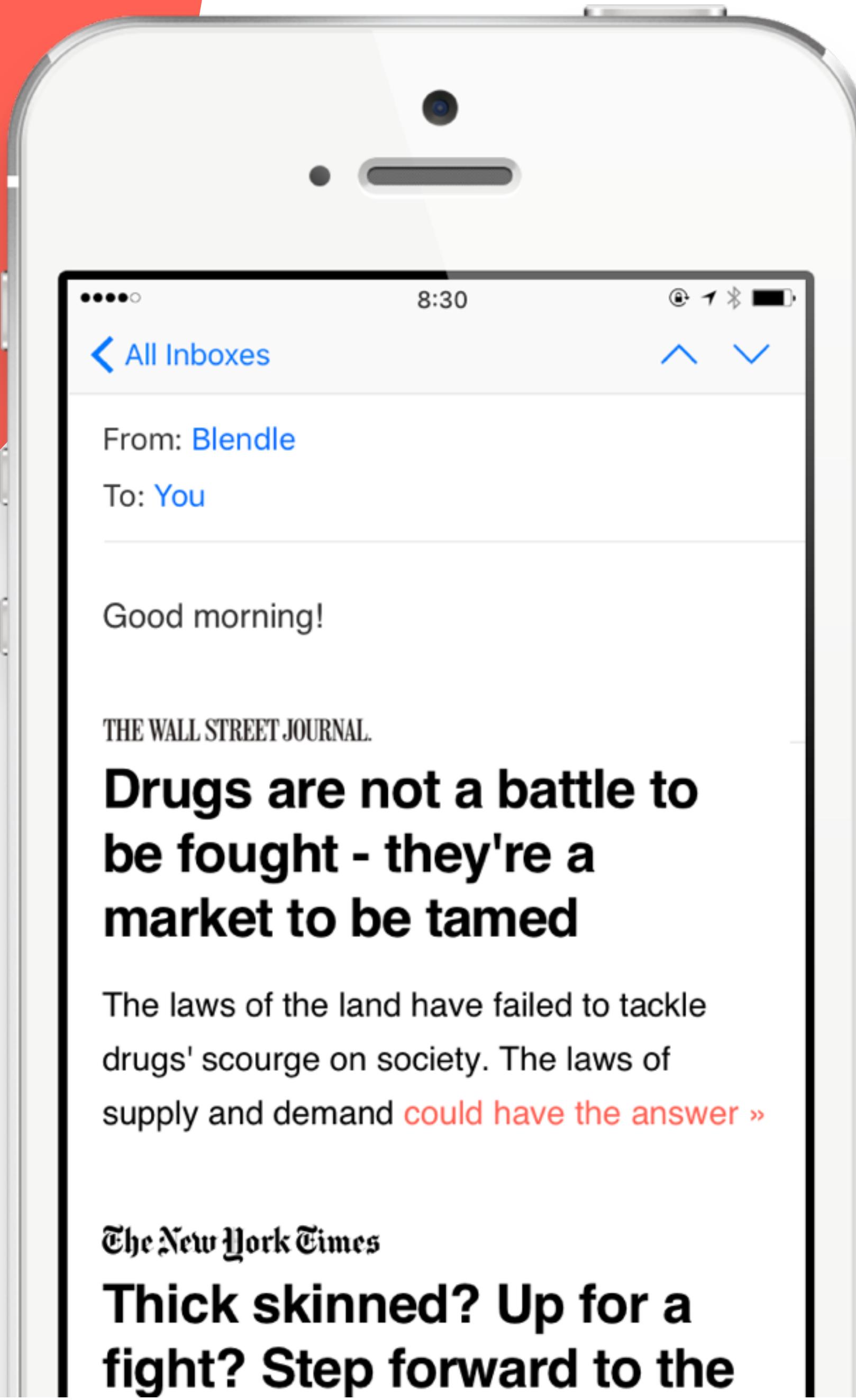
- Should be personalised
  - Lots of content
  - Lots of different tastes
- Can be personalised
  - We know what our users read
  - What they didn't like
  - When they read

# 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
  - ...

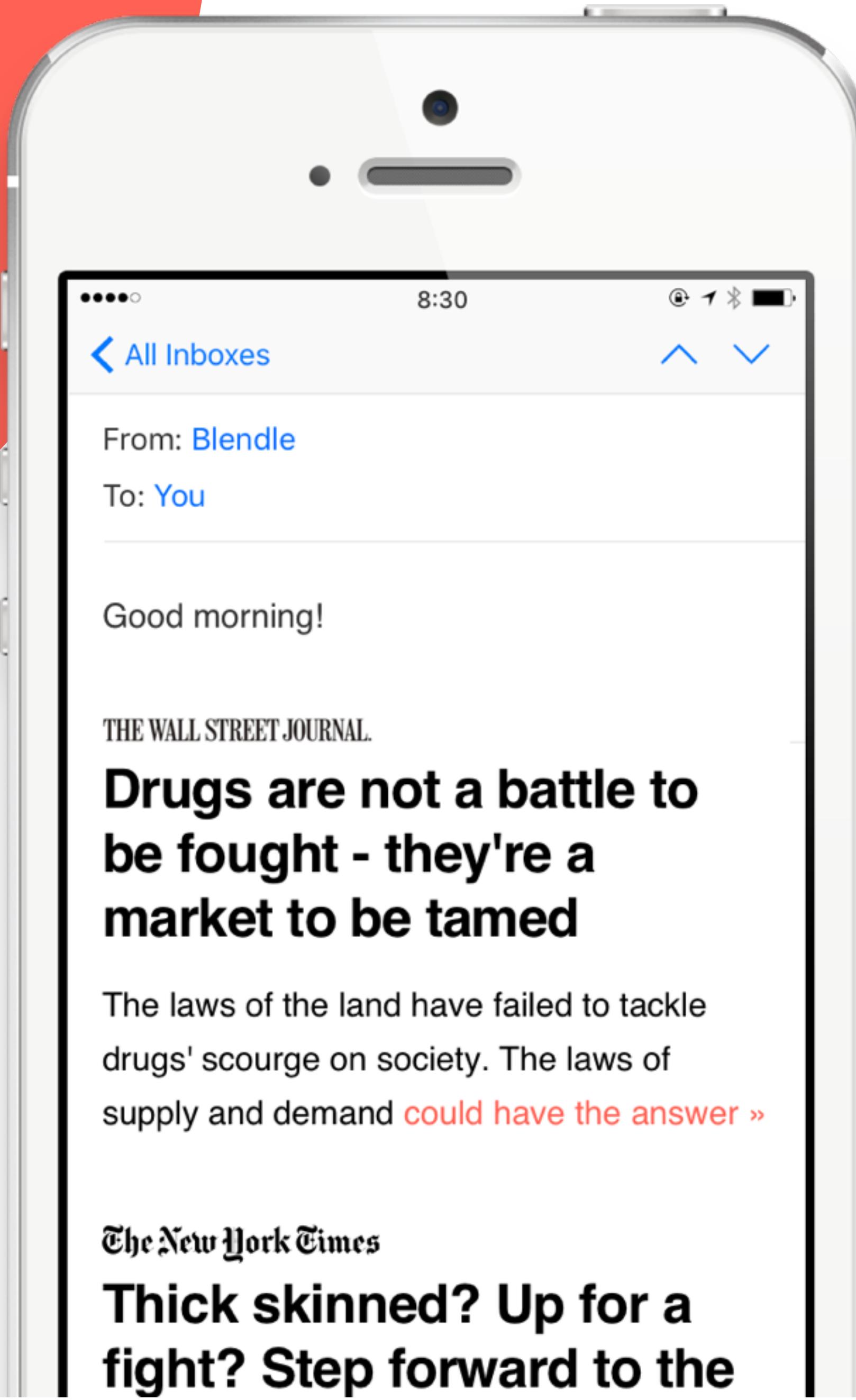
# Cold start problem





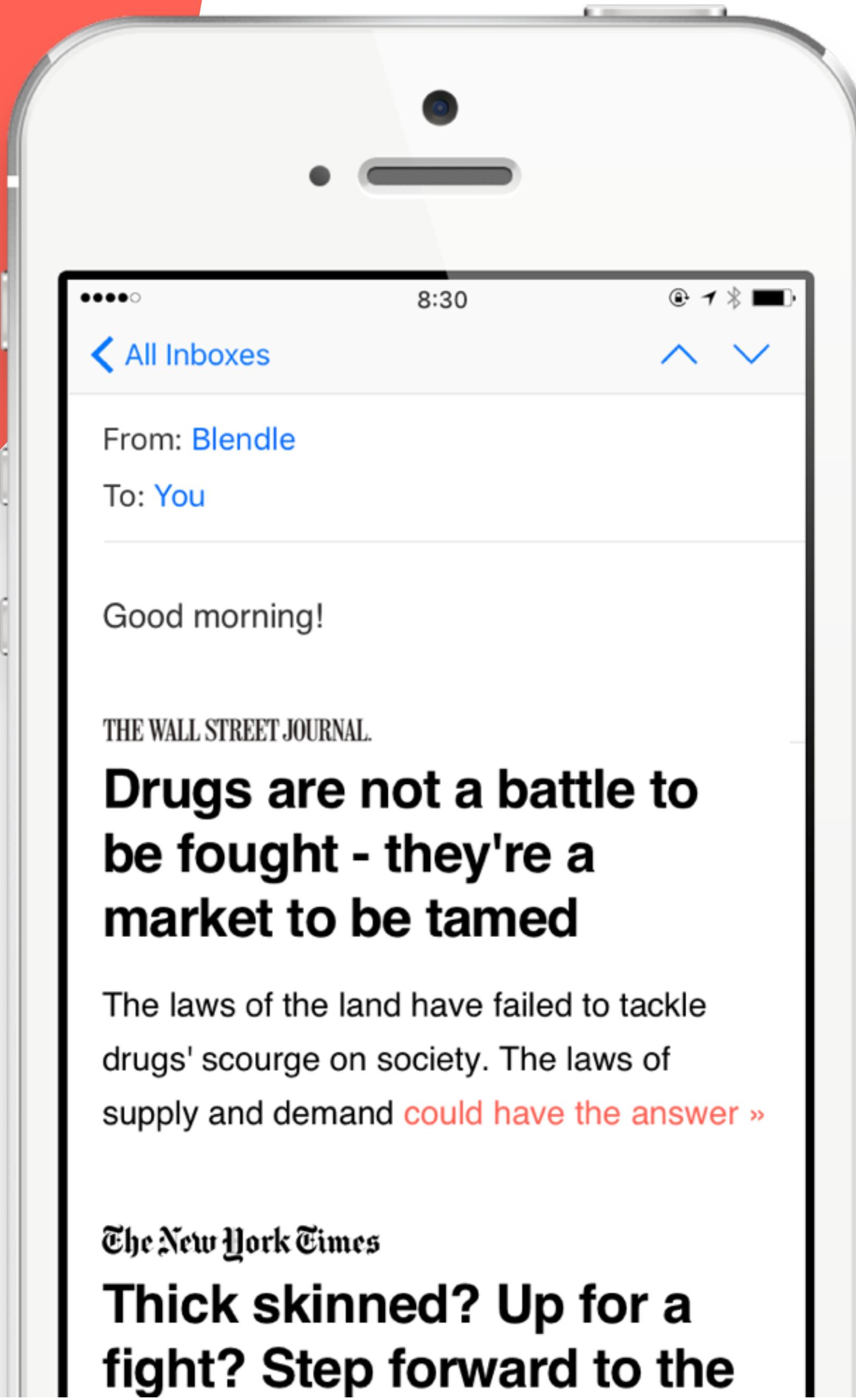
# Cold start problem

- 6K new articles every night



# Cold start problem

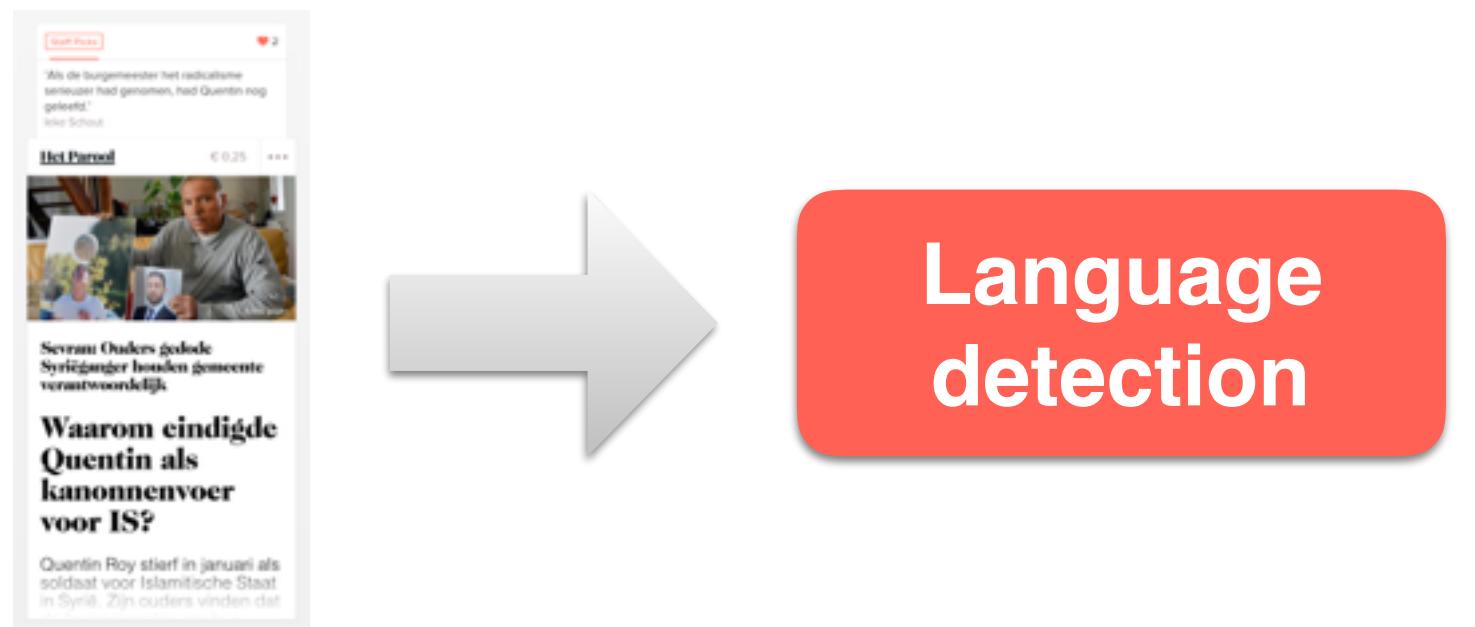
- 6K new articles every night
- Our newsletter is an important traffic driver



# 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

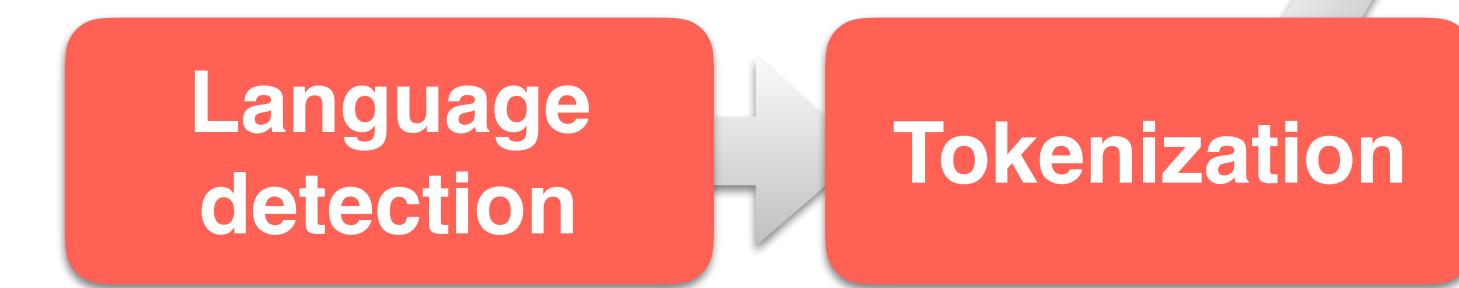
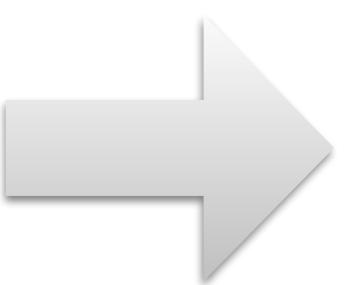
# Article enrichment pipeline



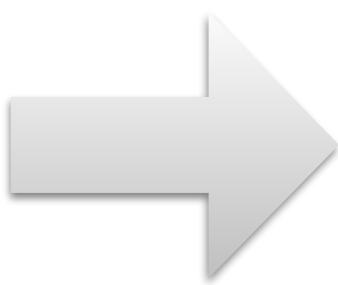
# Article enrichment pipeline



# Article enrichment pipeline



# Article enrichment pipeline



Language  
detection

Tokenization

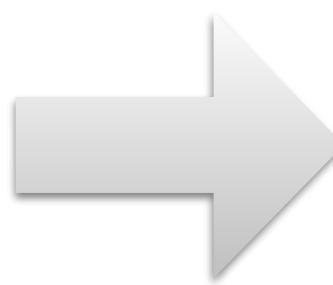
Stylometry

Sentiment  
analysis

Length, word  
variation, vocabulary  
richness, ...

Polarity scores  
(negative, positive)

# Article enrichment pipeline



Language  
detection

Tokenization

PoS-tagging

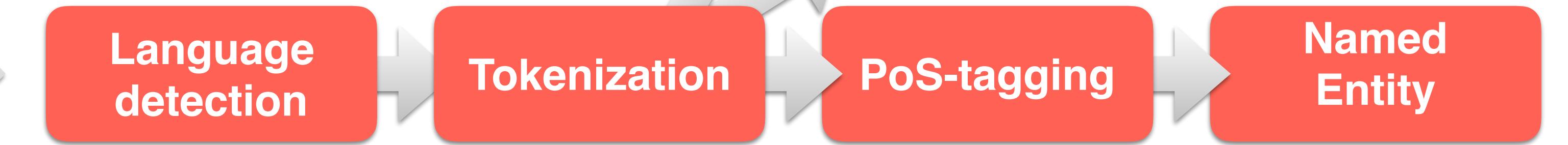
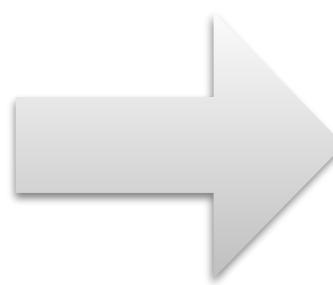
Stylometry

Sentiment  
analysis

Length, word  
variation, vocabulary  
richness, ...

Polarity scores  
(negative, positive)

# Article enrichment pipeline

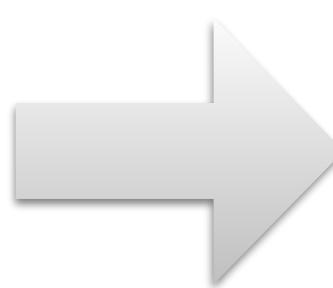


Length, word variation, vocabulary richness, ...

Polarity scores (negative, positive)

Locations, persons, organizations

# Article enrichment pipeline



Language detection

Tokenization

PoS-tagging

Author extraction

Named Entity

Stylometry

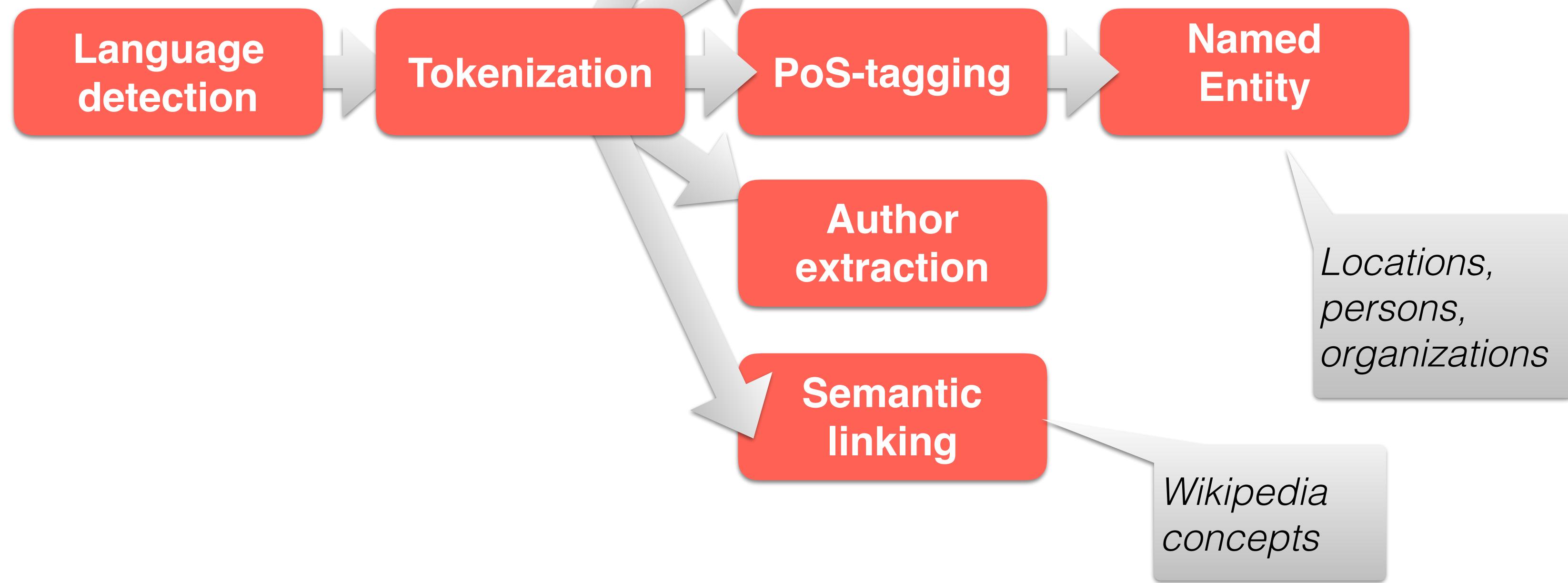
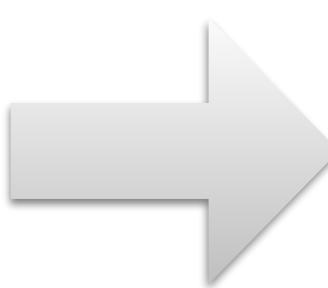
Sentiment analysis

Length, word variation, vocabulary richness, ...

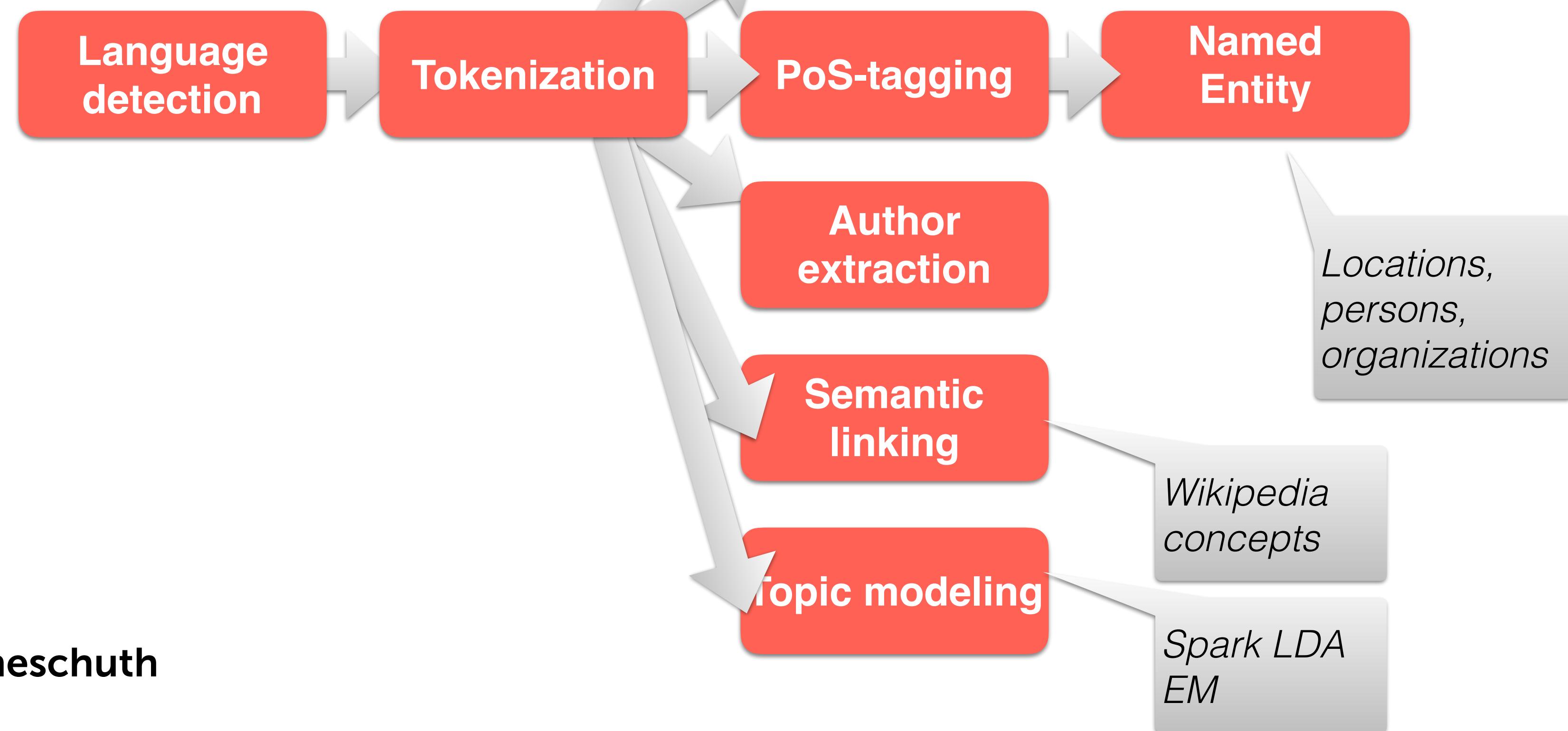
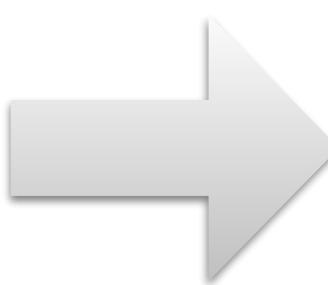
Polarity scores (negative, positive)

Locations, persons, organizations

# Article enrichment pipeline



# Article enrichment pipeline



# Article enrichment pipeline

## 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 vegetari\u00e9r", "type": "h1"}, {"content": "TOM KELLERHUIS", "type": "byline"},  
{"content": "Ik ken een flink aantal restaurateurs dat gruw bij het woord vegetari\u00e9r. 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 z\u00f3nder vis of vlees op de kaart stonden.", "type": "p"}, {"content": "Een goede  
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 \u00e9\u00e4n 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 vegetari\u00e9r 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 temp\u00e9 voor de avondmaaltijd.", "type": "p"}, {"content": "Ik kreeg laatst het receptenboek De dronken  
vegetari\u00e9r van Floor Overgoor, Hanny Roskamp en Yannick Slagter toegezonden, dat gehakt maakt van dat oude clich\u00e9.  
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 vegetari\u00e9r een betere ambassadeur is voor een diervriendelijke wereld dan een  
vegetari\u00e9r 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 vegetari\u00e9r 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 fonuestel, 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  
knoflooksauw 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	0.3666666666666664	1	<a href="http://dbpedia.org/resource/Supper">http://dbpedia.org/resource/Supper</a>
http://nl.wikipedia.org/wiki/Tic	Tic	tic	1.0416666666666667	1	<a href="http://dbpedia.org/resource/Tic">http://dbpedia.org/resource/Tic</a>
http://nl.wikipedia.org/wiki/Platvoeten	Platvoeten	platvoeten	0.47058823529411764	1	<a href="http://dbpedia.org/resource/Flat_feet">http://dbpedia.org/resource/Flat_feet</a>
http://nl.wikipedia.org/wiki/Groente	Groente	groente	0.41304347826086957	2	<a href="http://dbpedia.org/resource/Vegetable">http://dbpedia.org/resource/Vegetable</a>
http://nl.wikipedia.org/wiki/Sojaboon	Sojaboon	sojabonen	0.6206896551724138	1	<a href="http://dbpedia.org/resource/Soybean">http://dbpedia.org/resource/Soybean</a>
http://nl.wikipedia.org/wiki/Vegetarisme	Vegetarisme	vegetari�	2.6956521739130435	6	<a href="http://dbpedia.org/resource/Vegetarianism">http://dbpedia.org/resource/Vegetarianism</a>
http://nl.wikipedia.org/wiki/Pinot%20noir	Pinot noir	pinot noir	0.8827586206896552	1	<a href="http://dbpedia.org/resource/Pinot_noir">http://dbpedia.org/resource/Pinot_noir</a>
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	<a href="http://dbpedia.org/resource/Zucchini">http://dbpedia.org/resource/Zucchini</a>
http://nl.wikipedia.org/wiki/Tofoe	Tofoe	tofoe	0.8275862068965517	3	<a href="http://dbpedia.org/resource/Tofu">http://dbpedia.org/resource/Tofu</a>
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	- bouillon	0.5752212389380531	1	<a href="http://dbpedia.org/resource/Broth">http://dbpedia.org/resource/Broth</a>
http://nl.wikipedia.org/wiki/Bouillon%20%28voedsel%29	Bouillon (voedsel)	bouillon	0.5752212389380531	2	<a href="http://dbpedia.org/resource/Broth">http://dbpedia.org/resource/Broth</a>
http://nl.wikipedia.org/wiki/Aubergine%20%28plant%29	Aubergine (plant)	aubergine	0.8666666666666667	1	<a href="http://dbpedia.org/resource/Eggplant">http://dbpedia.org/resource/Eggplant</a>
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	4.986111111111111

**And** these named entities

entity	tag	count
YANNICK SLAGTER	PER	1
WIJNFOONDUE	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	4.986111111111111

**And** these named entities

entity	tag	count
YANNICK SLAGTER	PER	1
WIJNFOONDUE	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

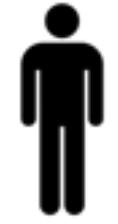
**Then** the user "anne" should have the following profile:

```
"""\n{\n    \"count\": 2,\n    \"enrichments\": {\n        \"language\": [\n            {\n                \"count\": 1,\n                \"name\": \"nl\"\n            },\n            {\n                \"count\": 1,\n                \"name\": \"en\"\n            }\n        ],\n        \"author\": [\n            {\n                \"count\": 1,\n                \"name\": \"TOM KELLERHUIS\"\n            },\n            {\n                \"count\": 1,\n                \"name\": \"PETER LANDERS\"\n            }\n        ]\n    },\n    \"stylometry\": [\n        {\n            \"count\": 2,\n            \"mean\": 2177.5,\n            \"stdev\": 515.4808434849931,\n            \"name\": \"article_length\"\n        },\n        {\n            \"count\": 2,\n            \"mean\": 5.165410504361017,\n            \"stdev\": 0.2535676336592832,\n            \"name\": \"avg_word_length\"\n        }\n    ]\n},\n    \"last_100_purchased_items\": [\"item1\", \"item2\"],\n    \"tracking_uid\": \"anne\"\n}....
```

# Learning to rank: preference learning



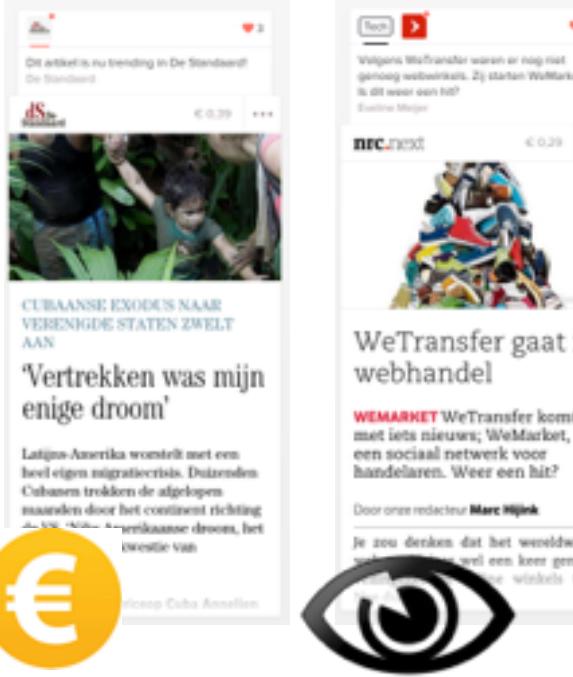
# Learning to rank: preference learning



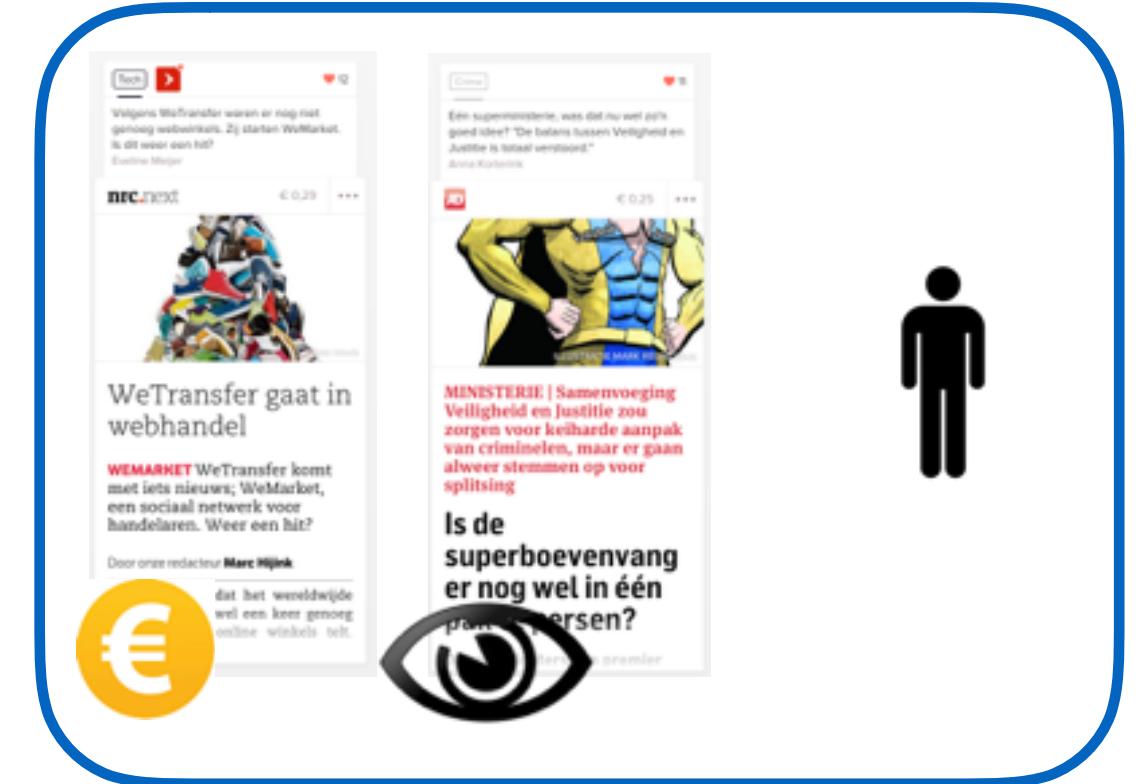
# Learning to rank: preference learning



# Learning to rank: preference learning



Enrich

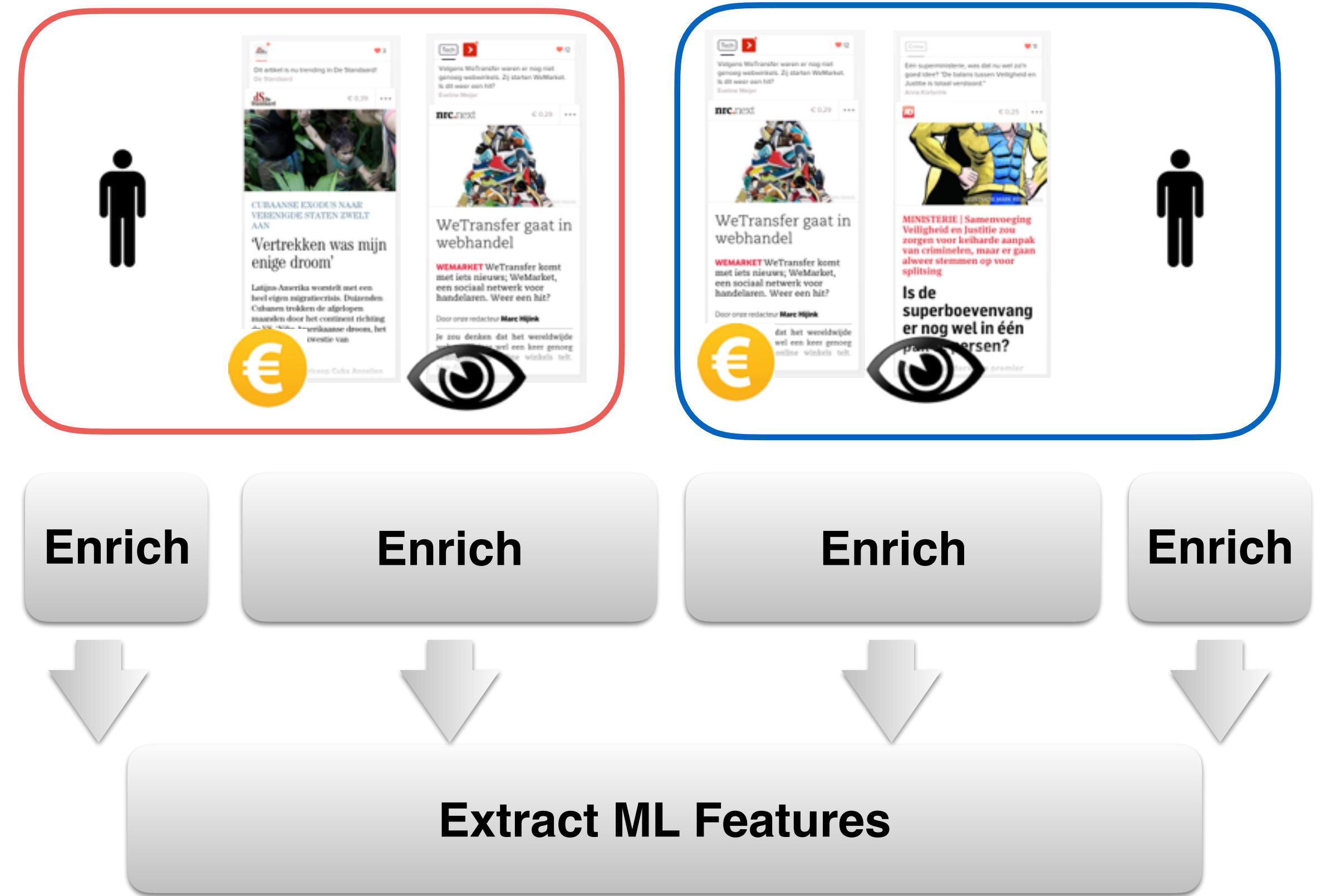


Enrich

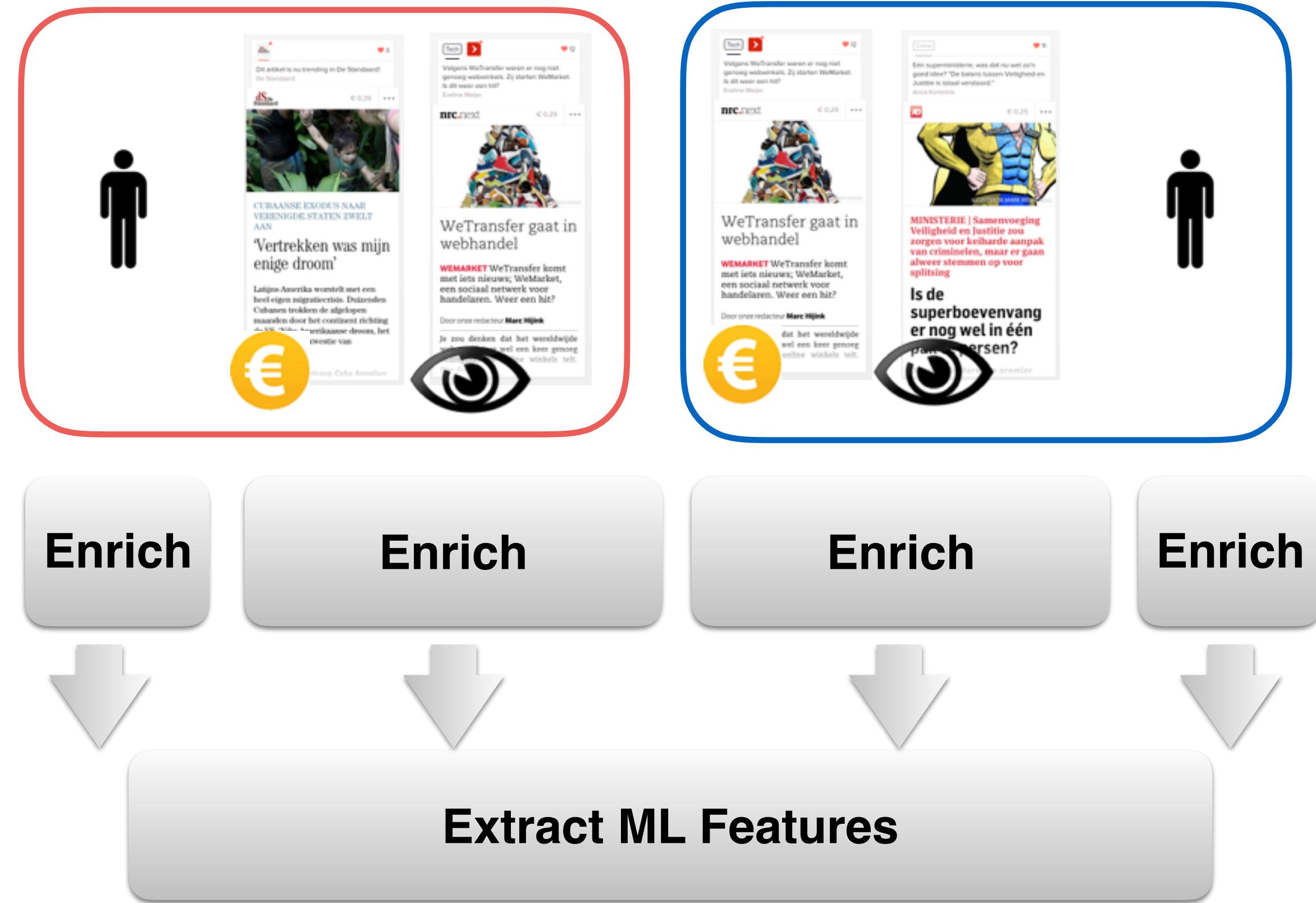


Enrich

# Learning to rank: preference learning



# Learning to rank: preference learning

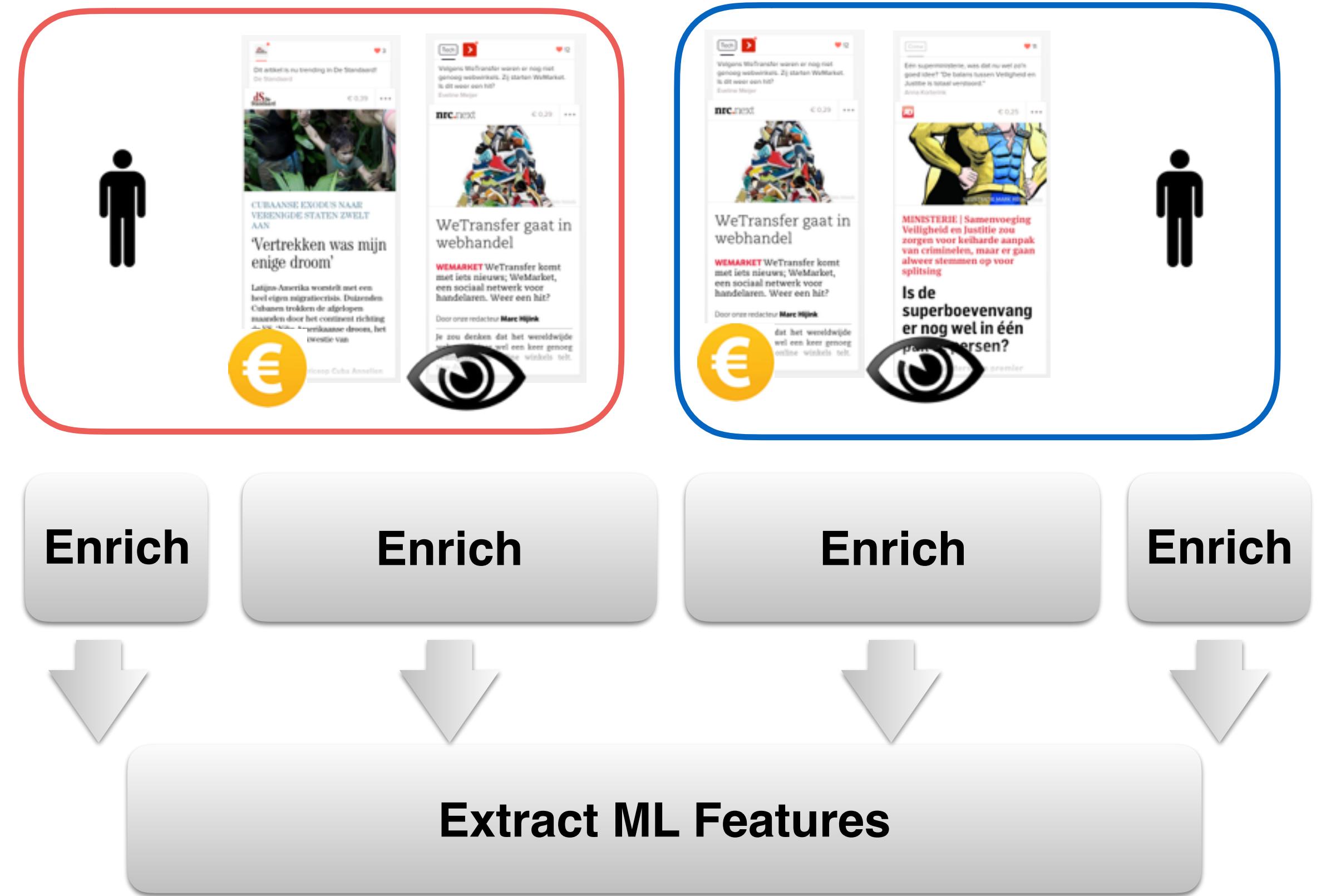


When I compute features

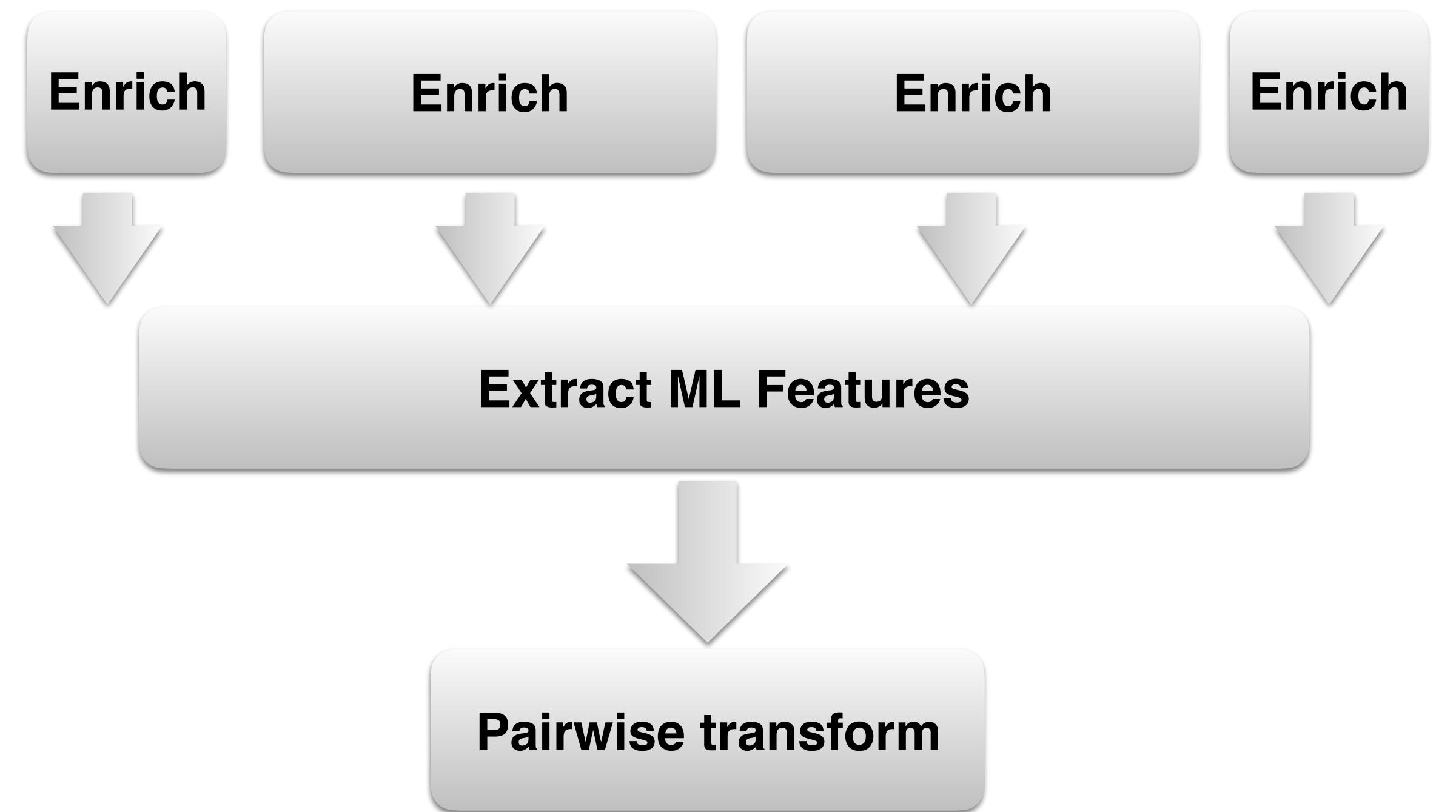
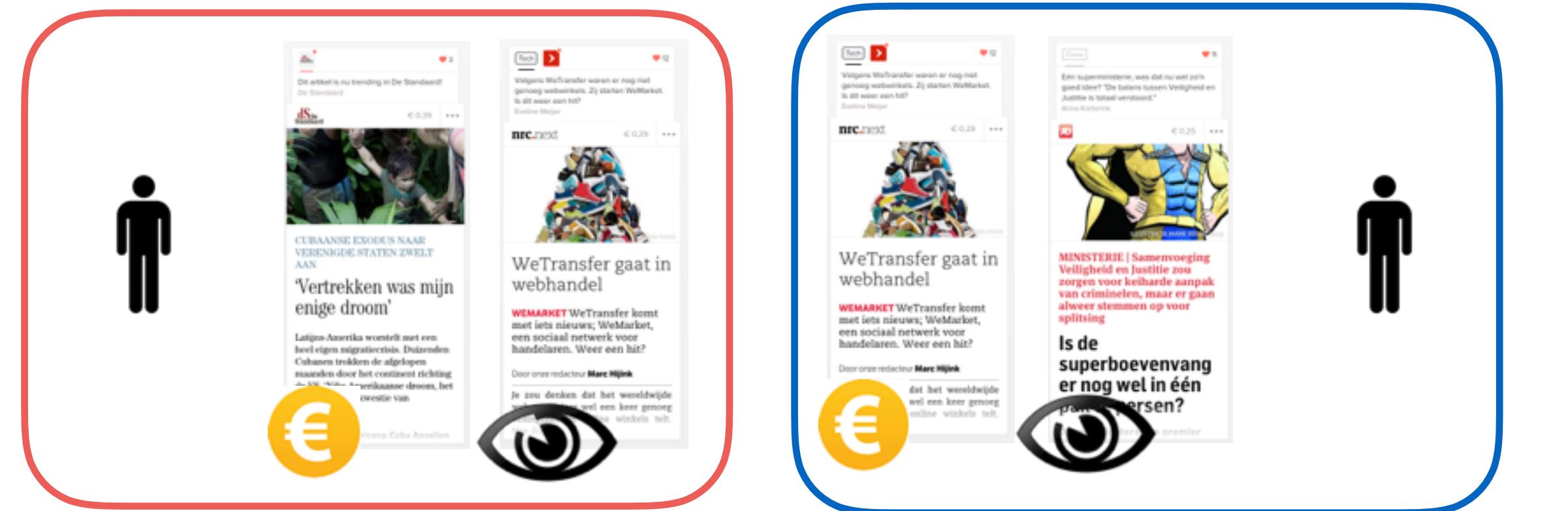
Then I should have the following features and values:

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	0.0977777777777777
user_item_named_entity_overlap	0.04081632653061
user_item_length_pdf	0.00316108841344

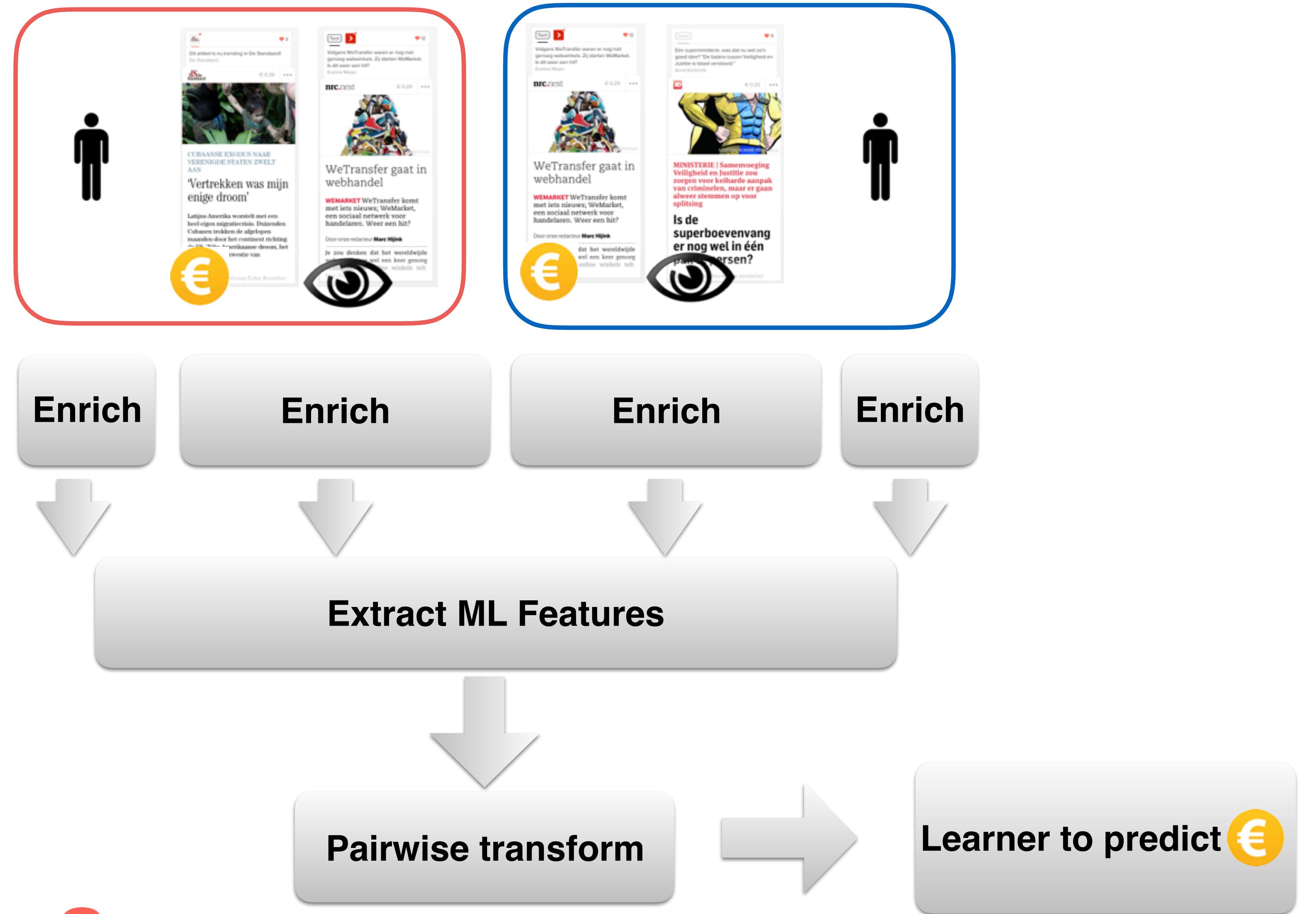
# Learning to rank: preference learning



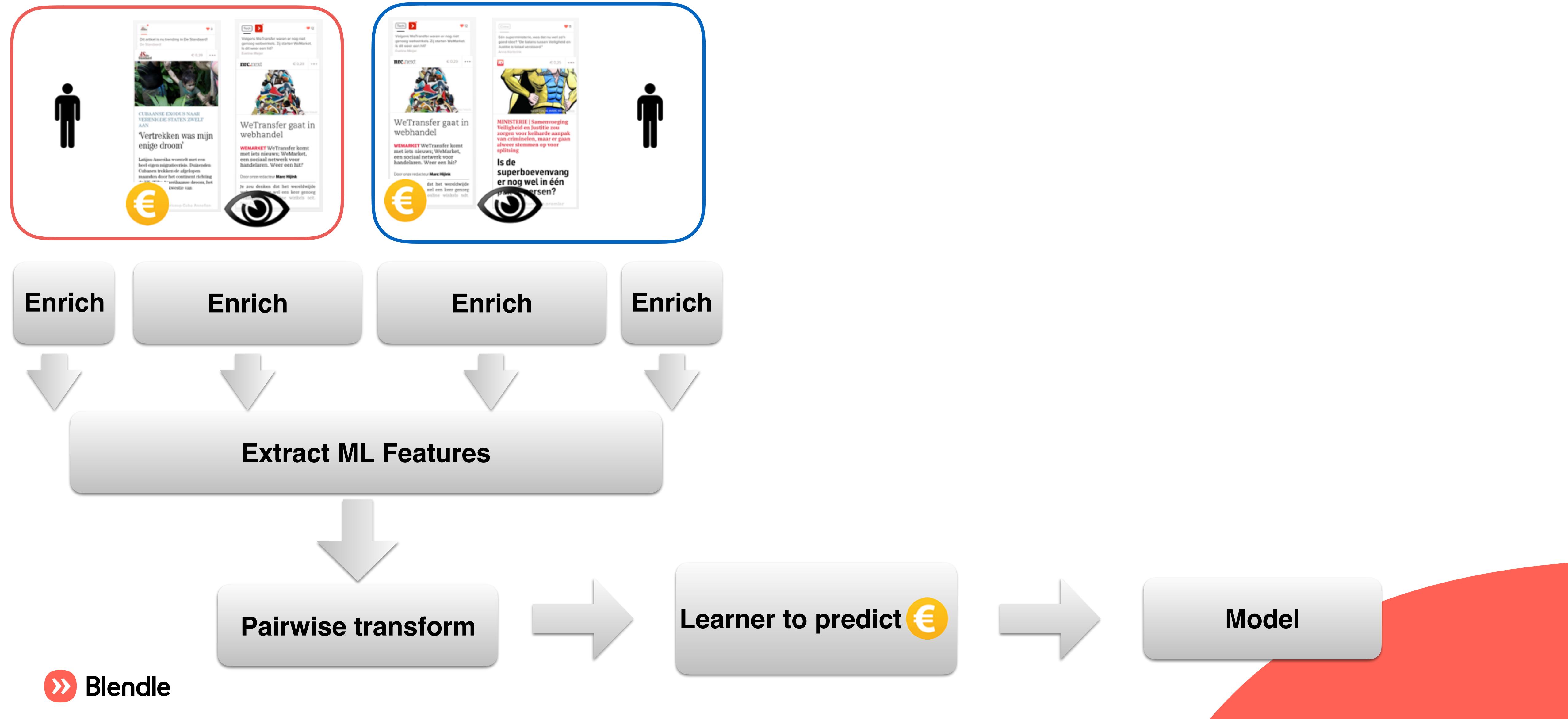
# Learning to rank: preference learning



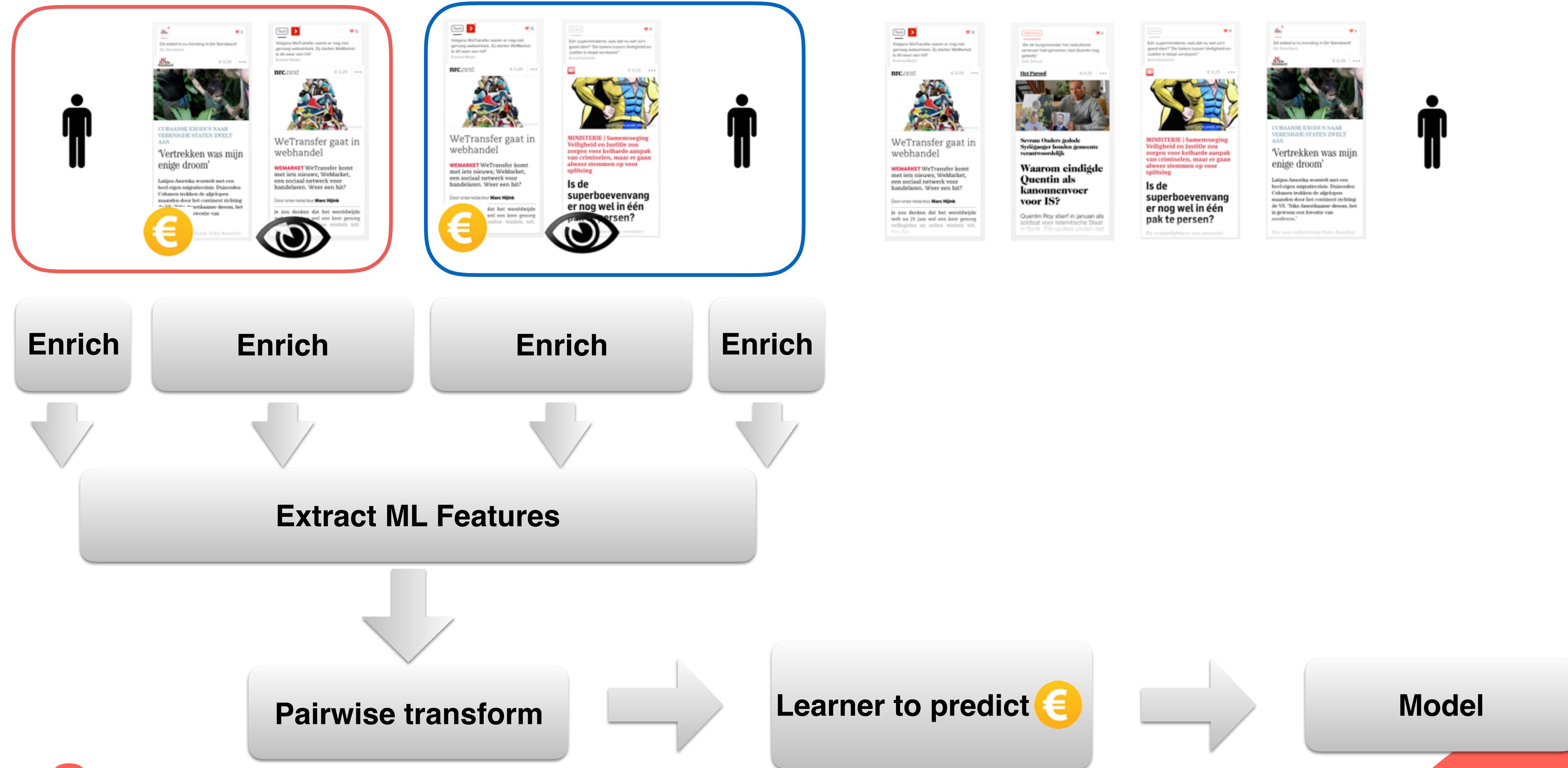
# Learning to rank: preference learning



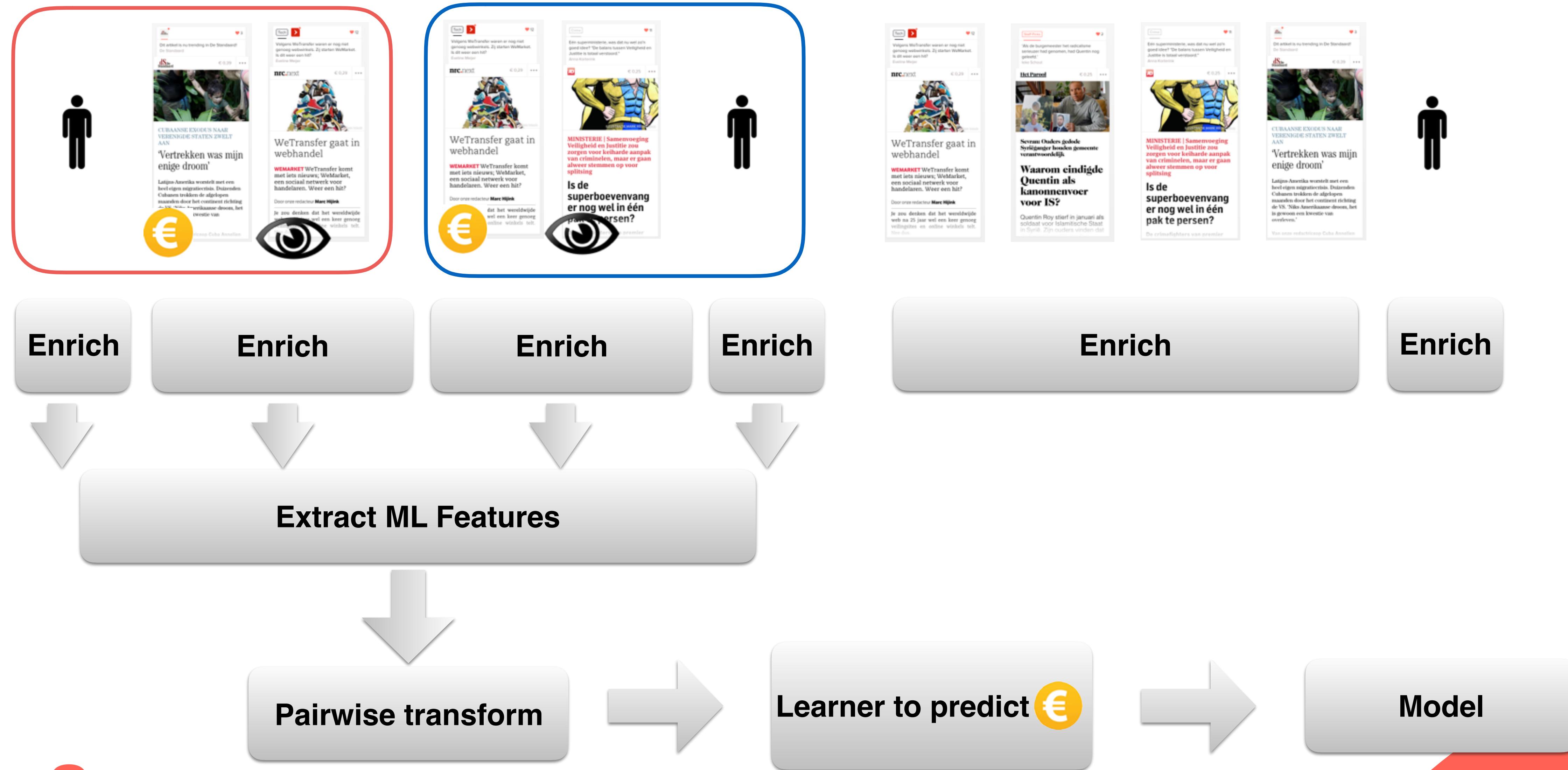
# Learning to rank: preference learning



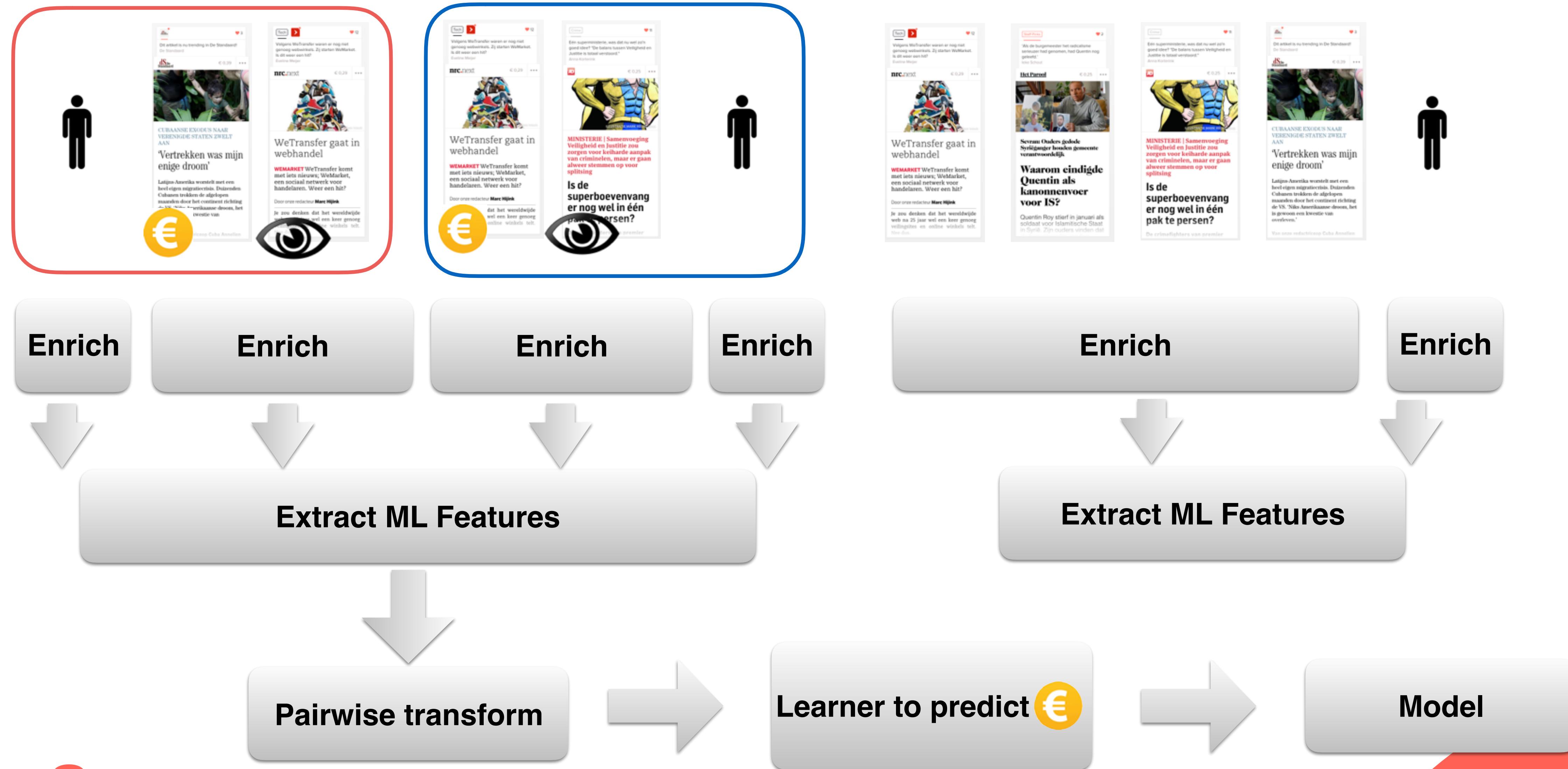
# Learning to rank: preference learning



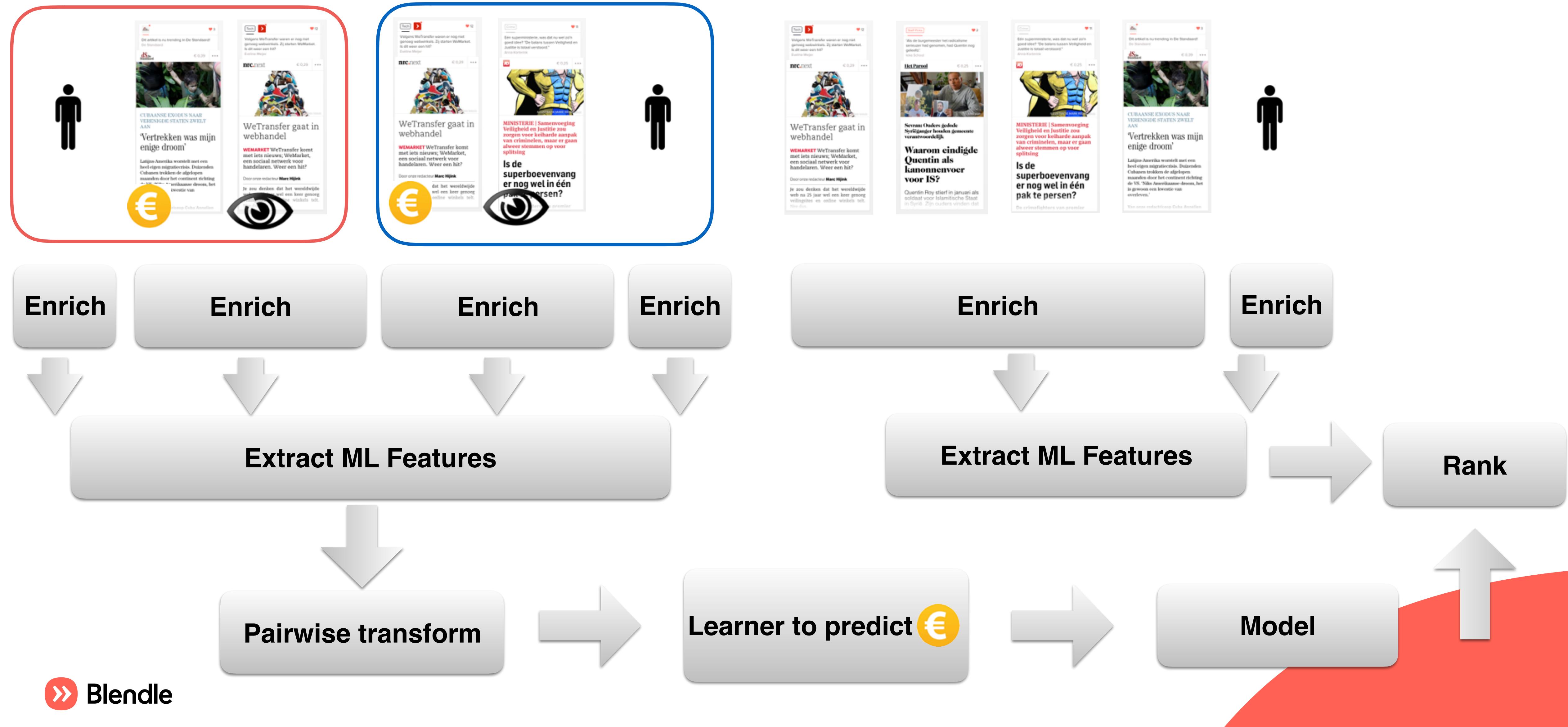
# Learning to rank: preference learning



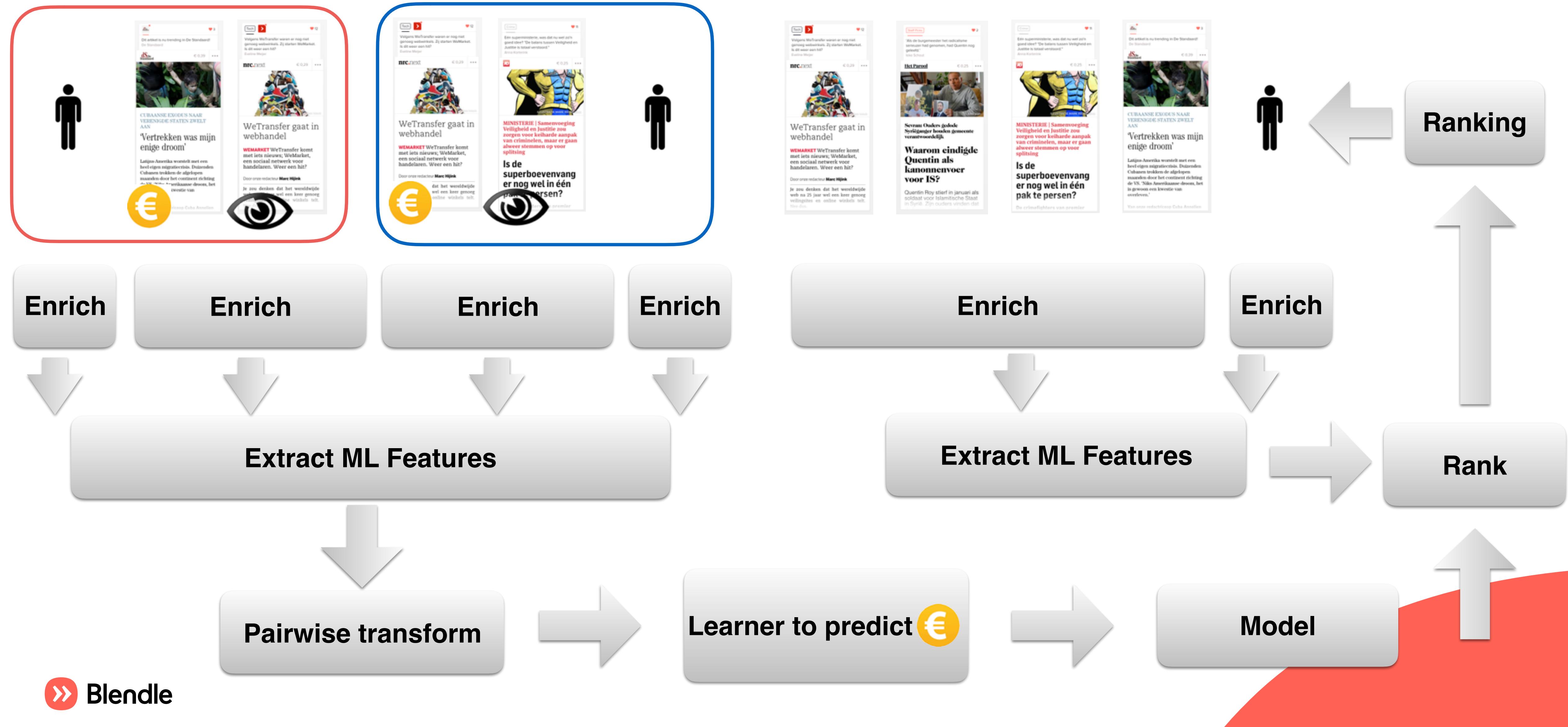
# Learning to rank: preference learning

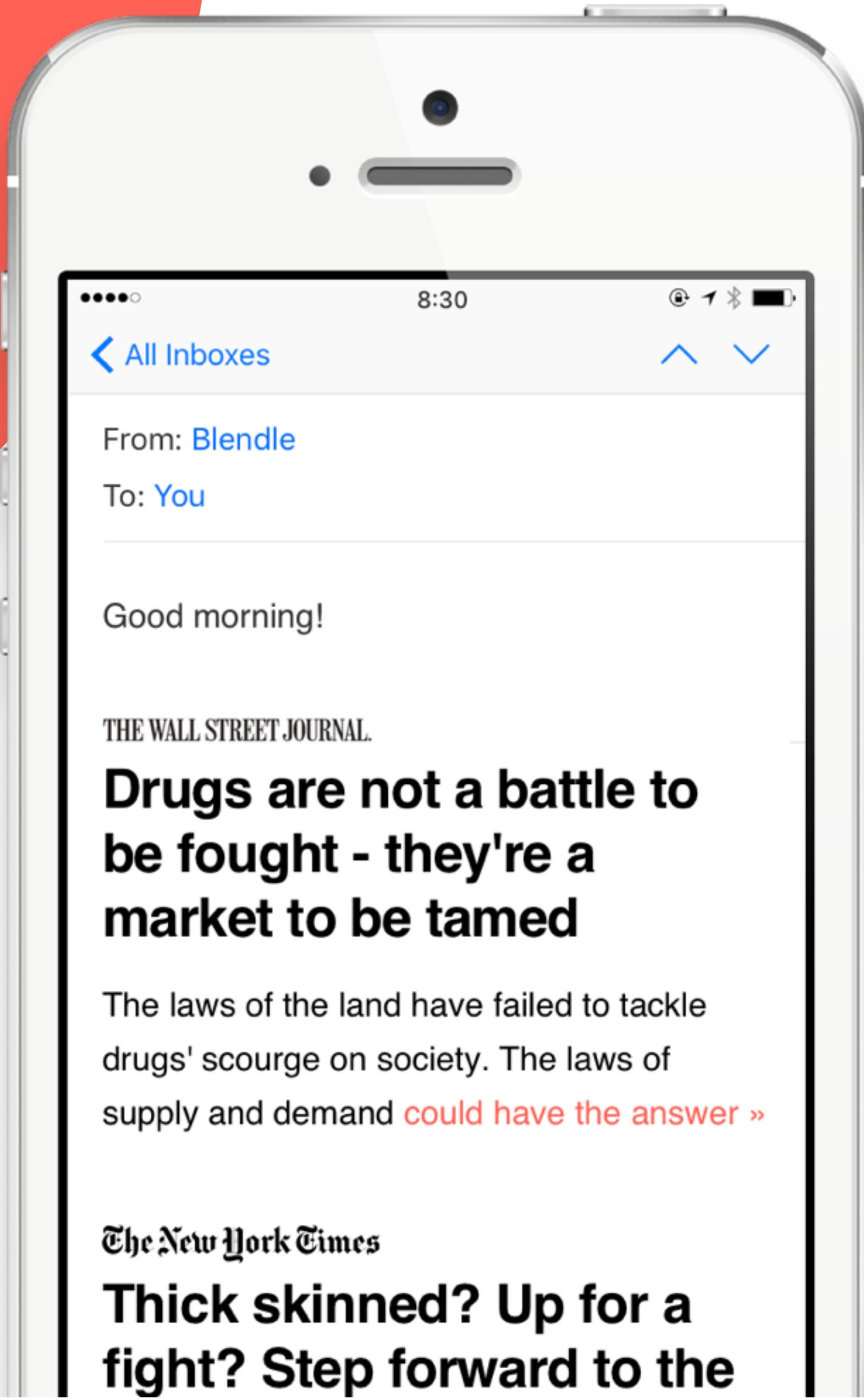


# Learning to rank: preference learning

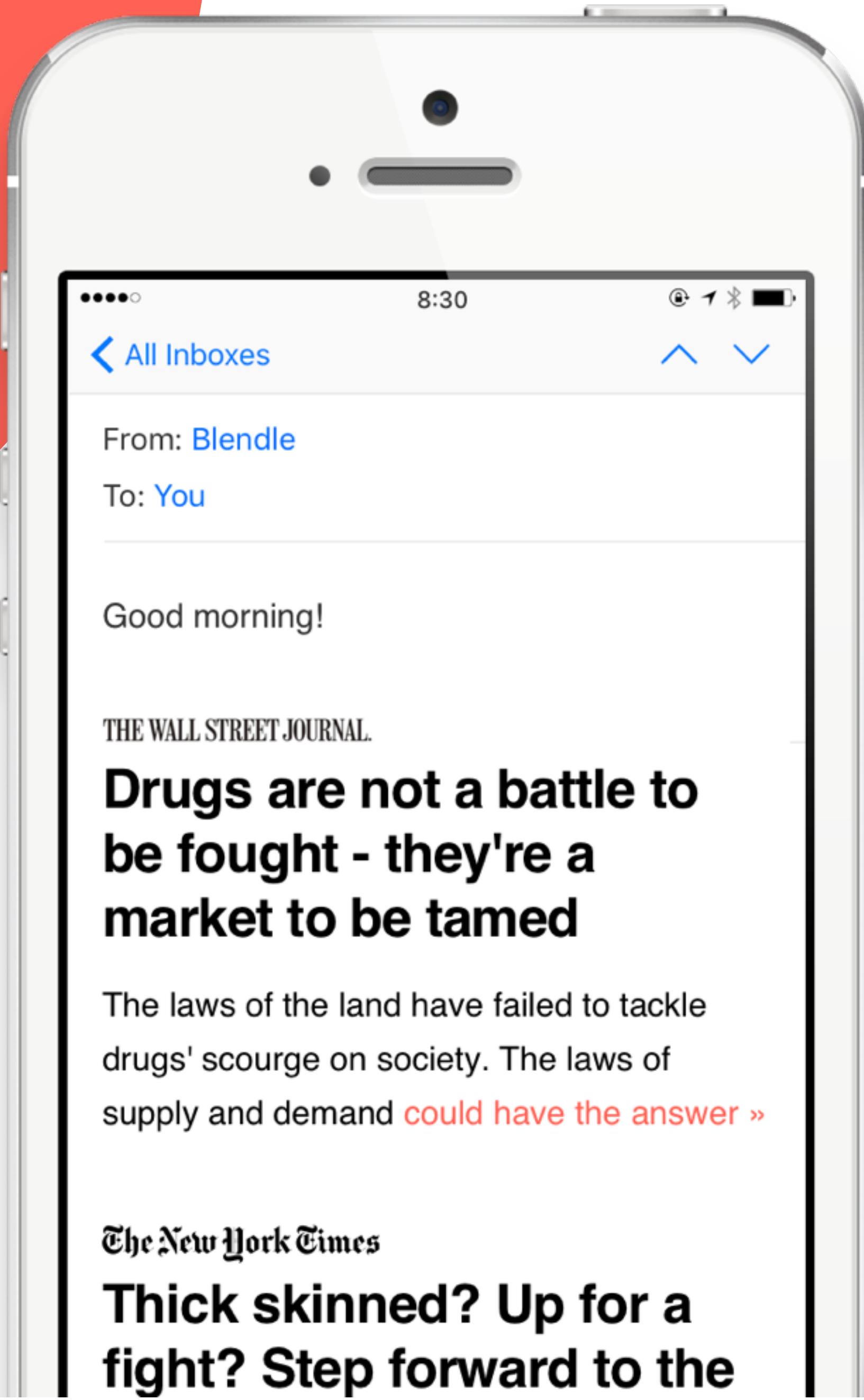


# Learning to rank: preference learning



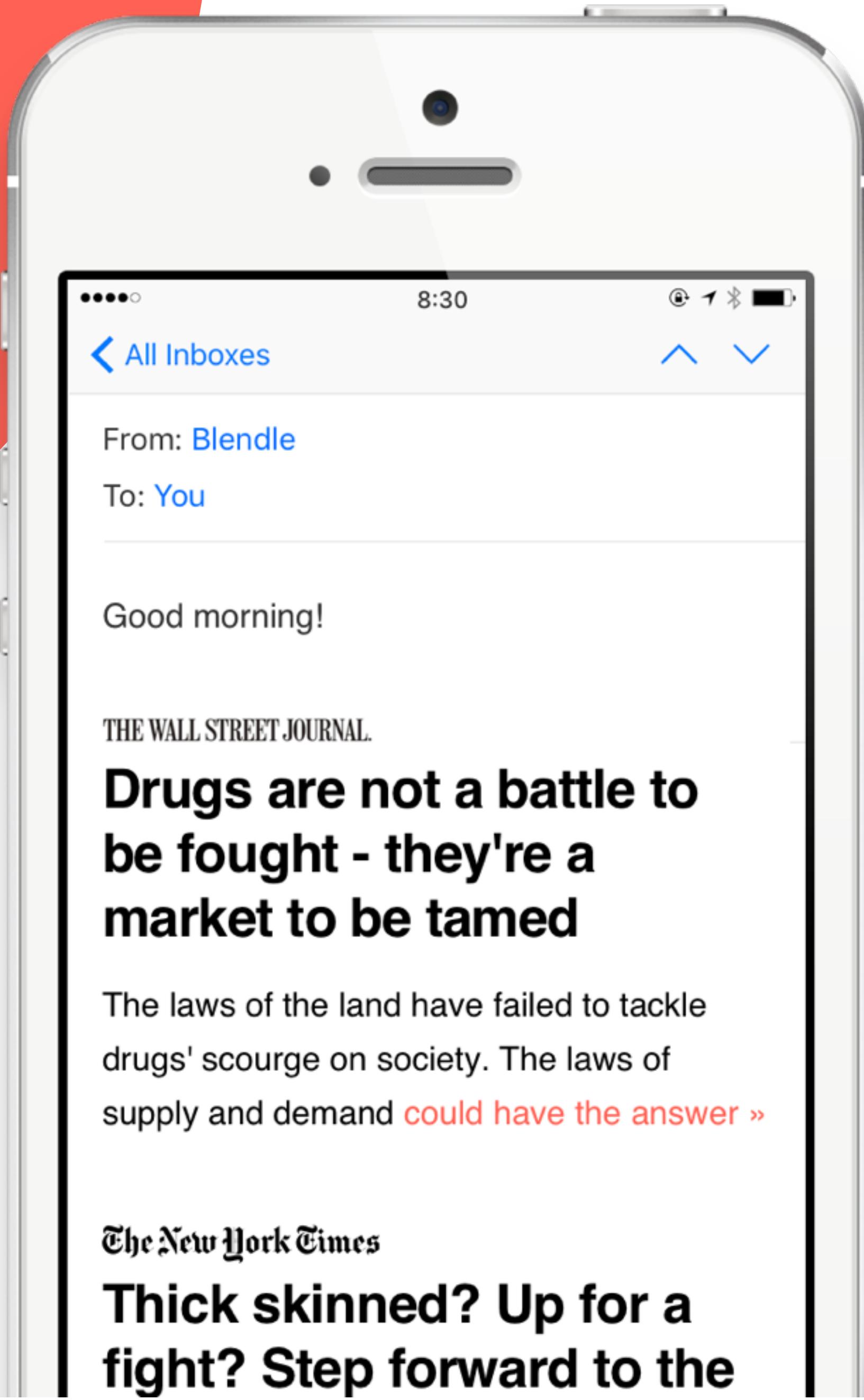


# Timing problem



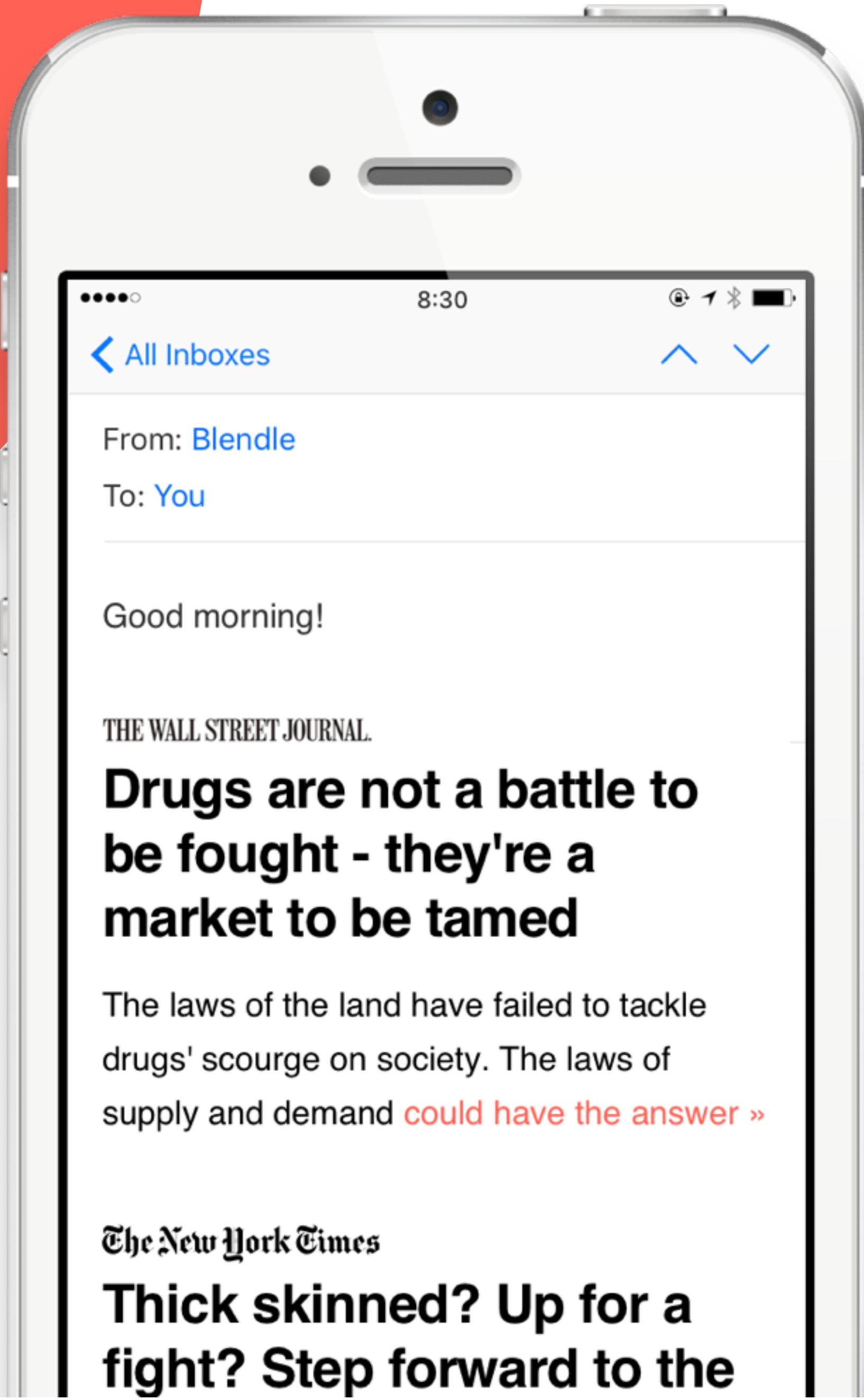
# Timing problem

- 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

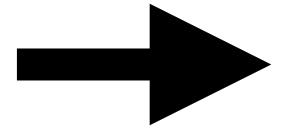


# Timing problem

- 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

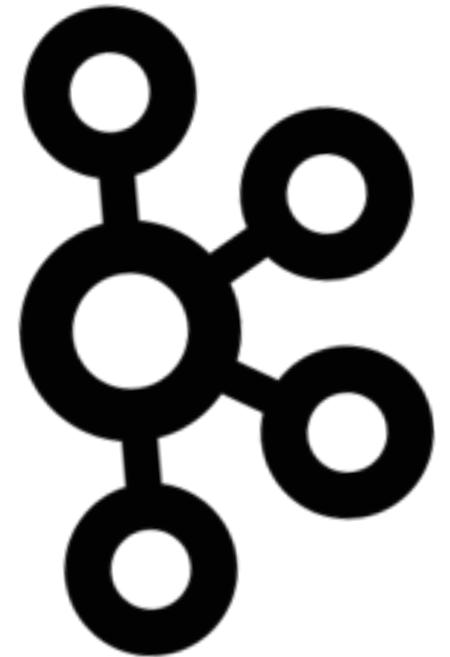
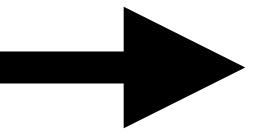
articles

published  
article



published article

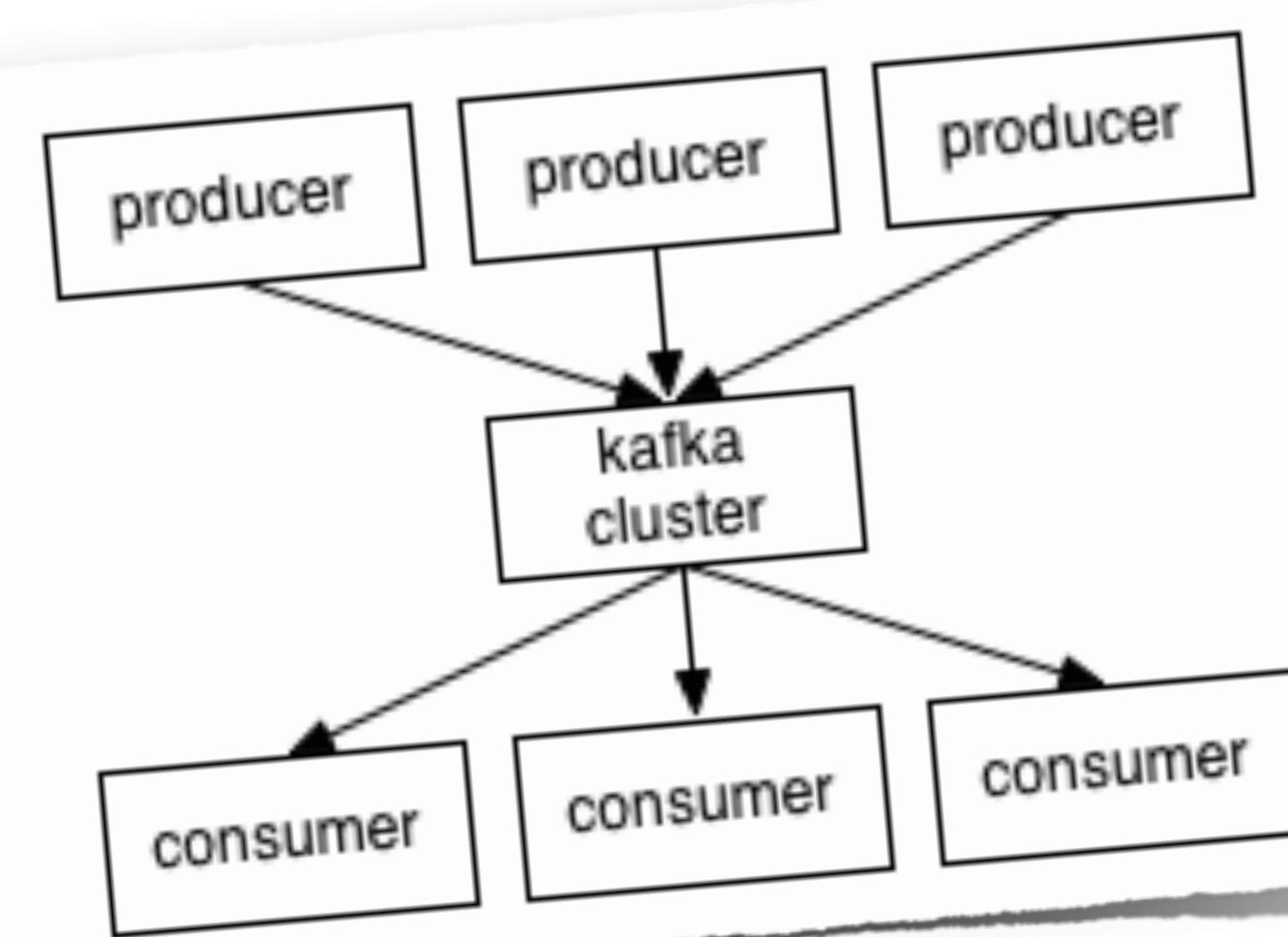
articles



# Kafka

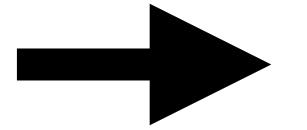
## 1.1 Introduction

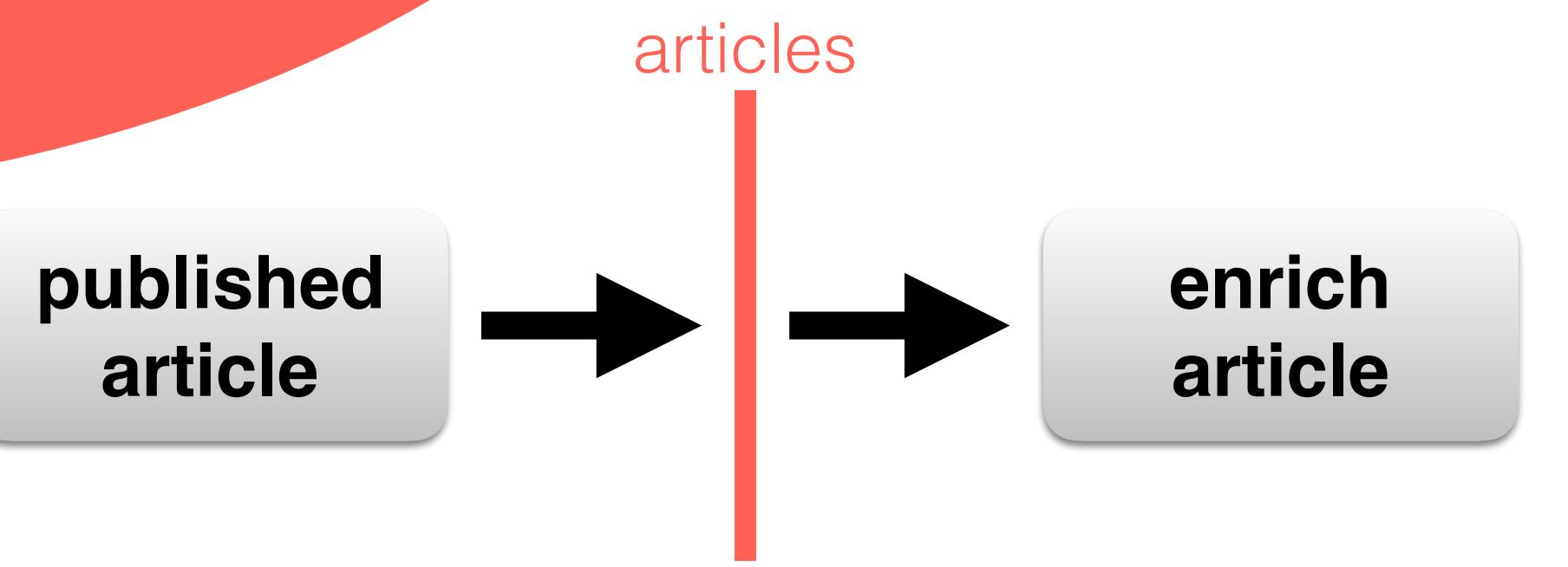
Kafka is a distributed, partitioned, replicated commit log service. It provides the functionality of a messaging system, but with a unique design.

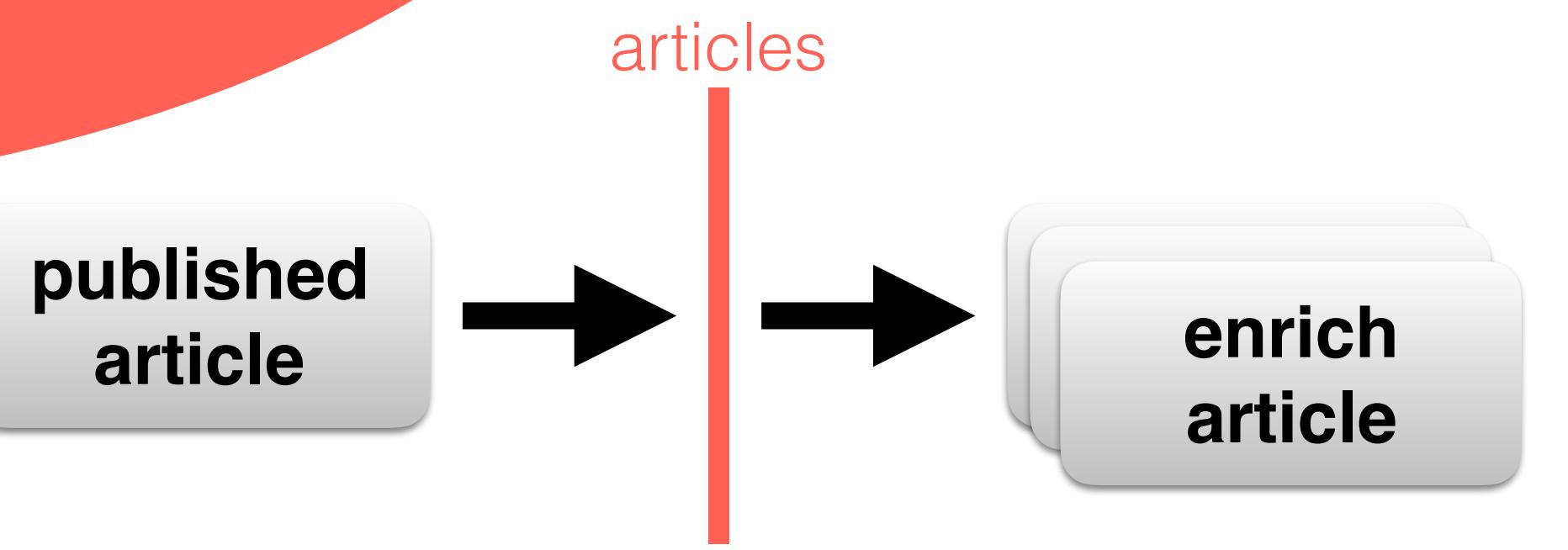


articles

published  
article







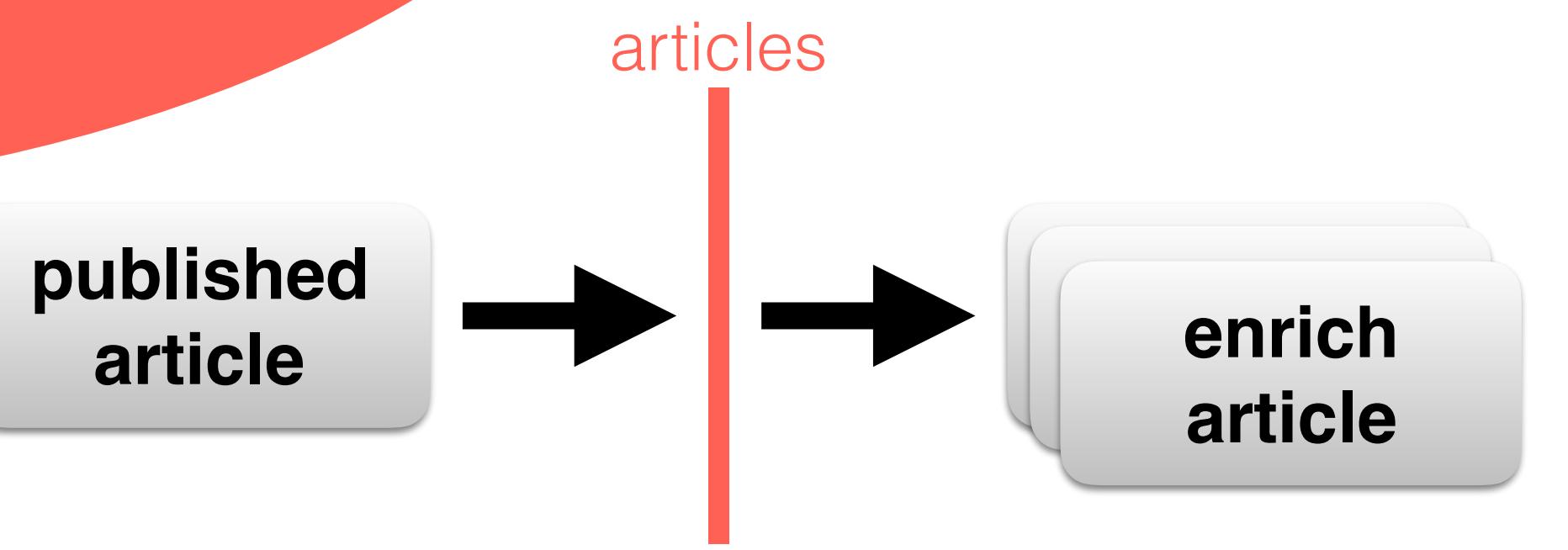


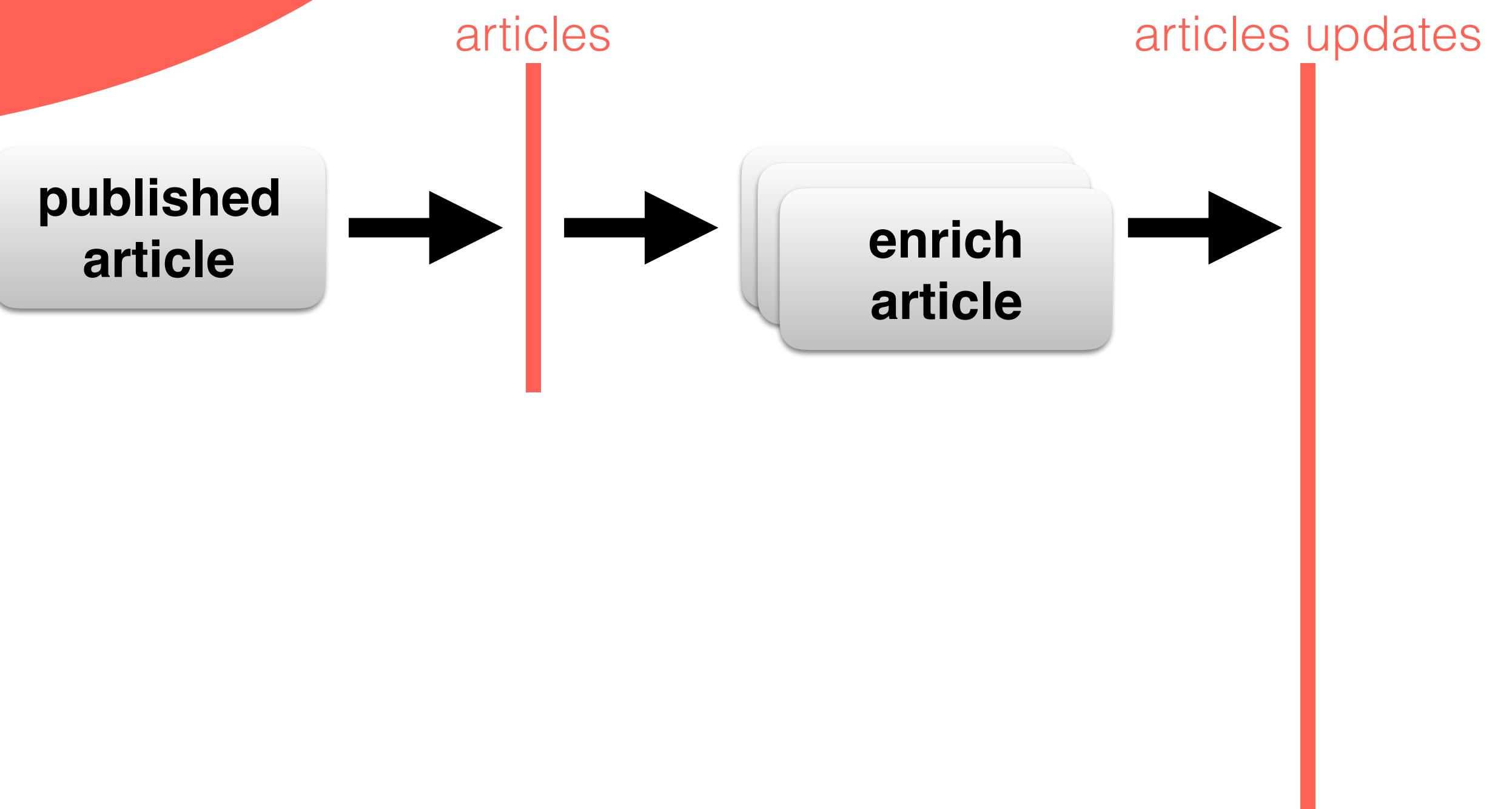
```
import argparse
import os
import sys

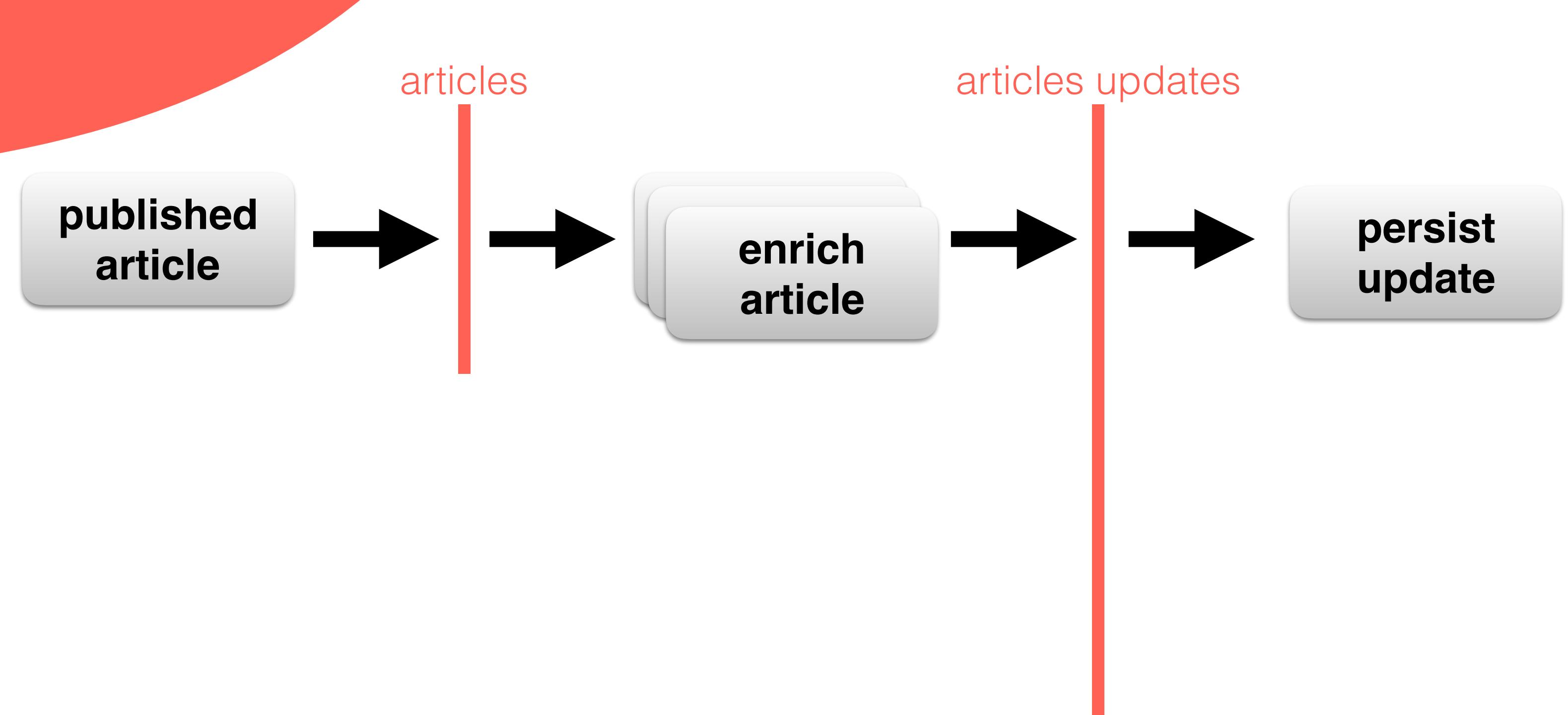
from blendle.pipes.task import Task
from blendle.tasks.enrich.itemprofiles import Pipeline
ENRICH_PIPELINE = os.environ.get('ENRICH_PIPELINE',
                                  'language,tokenizer,postagger,ner,stylometry,author').split(',')
```

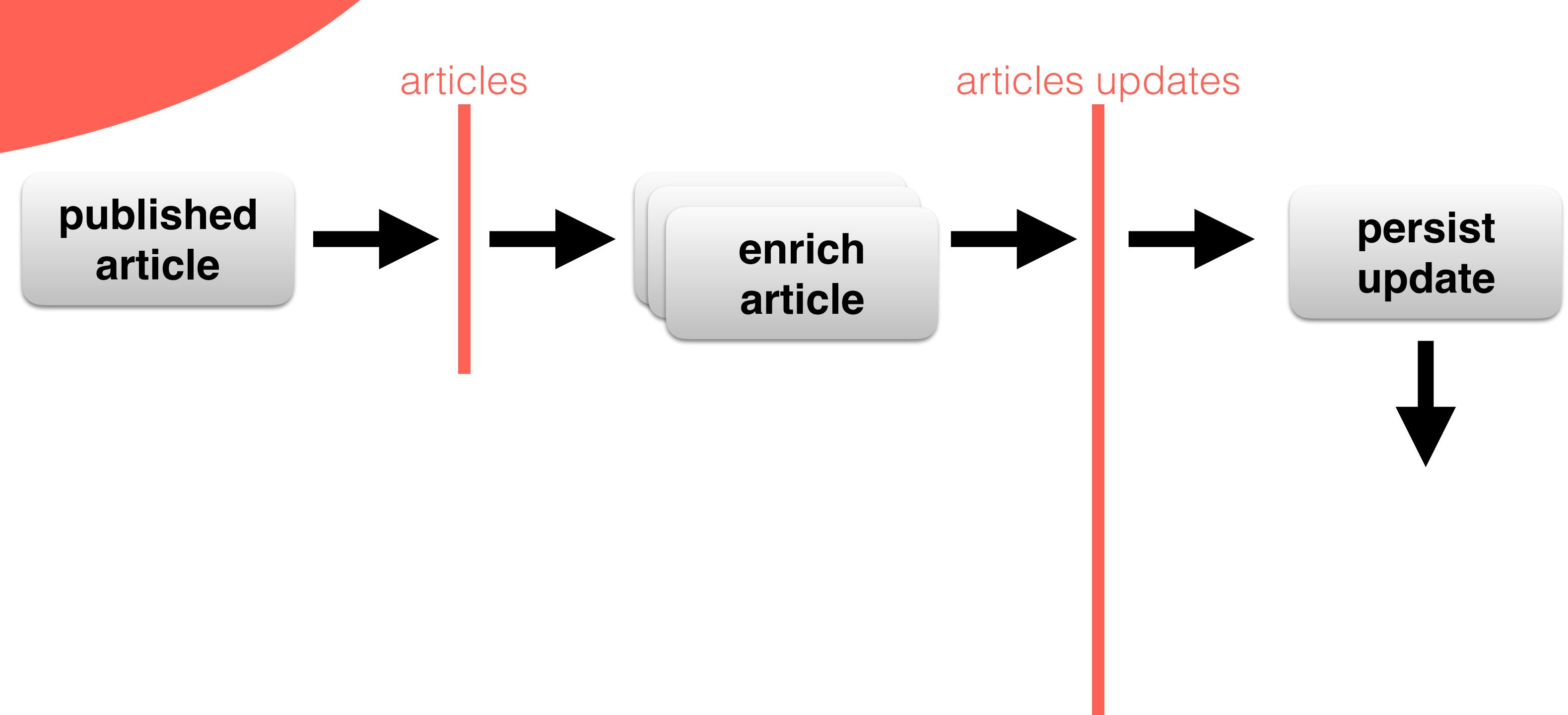
```
class Enrich(Task):
    def __init__(self, in_file, out_file, error_file, **kwargs):
        self.enricher = kwargs.get('enricher', ENRICH_PIPELINE)
        self.pipeline = Pipeline(self.enricher)
        super().__init__(in_file, out_file, error_file, **kwargs)

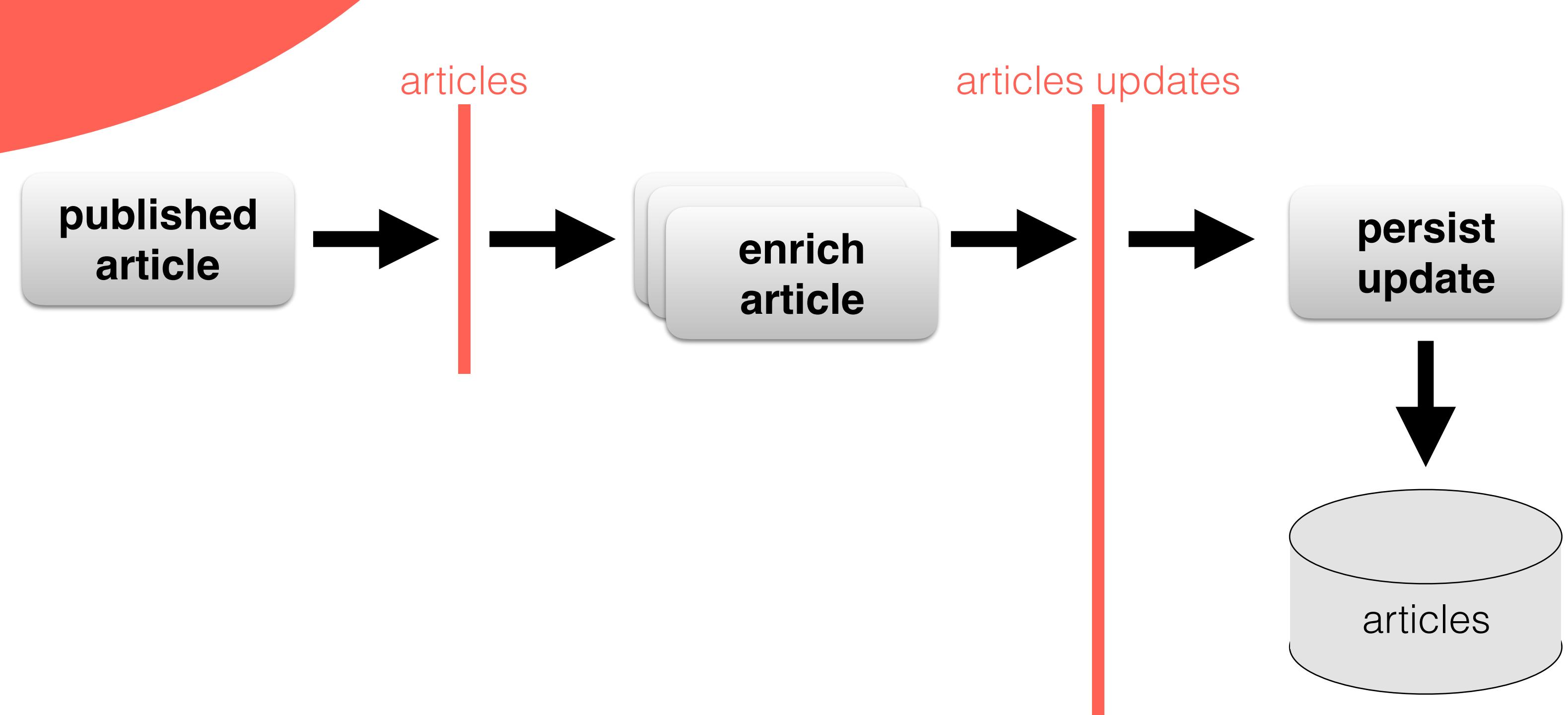
    def process(self, item):
        enrichments = self.pipeline(item).get('enrichments', {})
        yield {'id': item['id'],
               'item-metadata': {'enrichments': enrichments}}
```

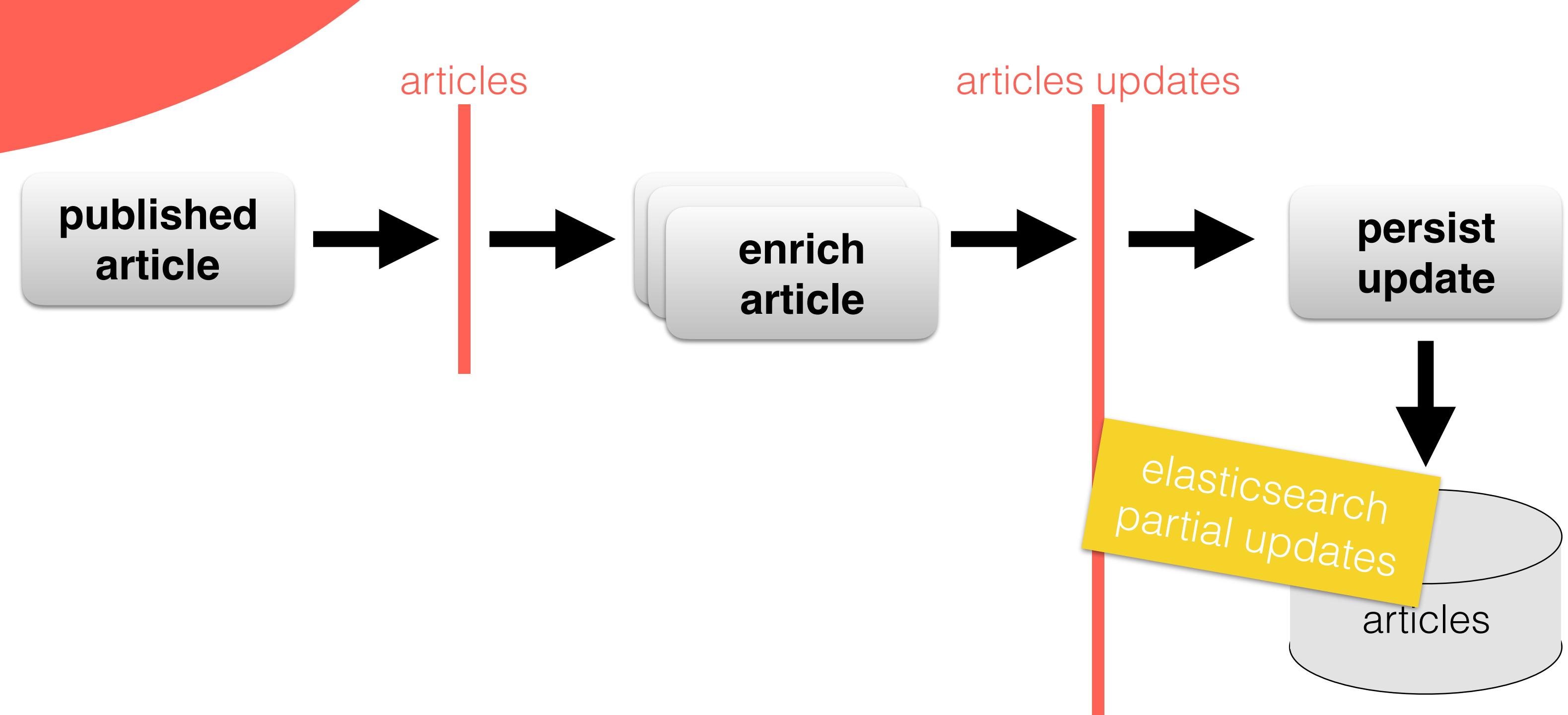


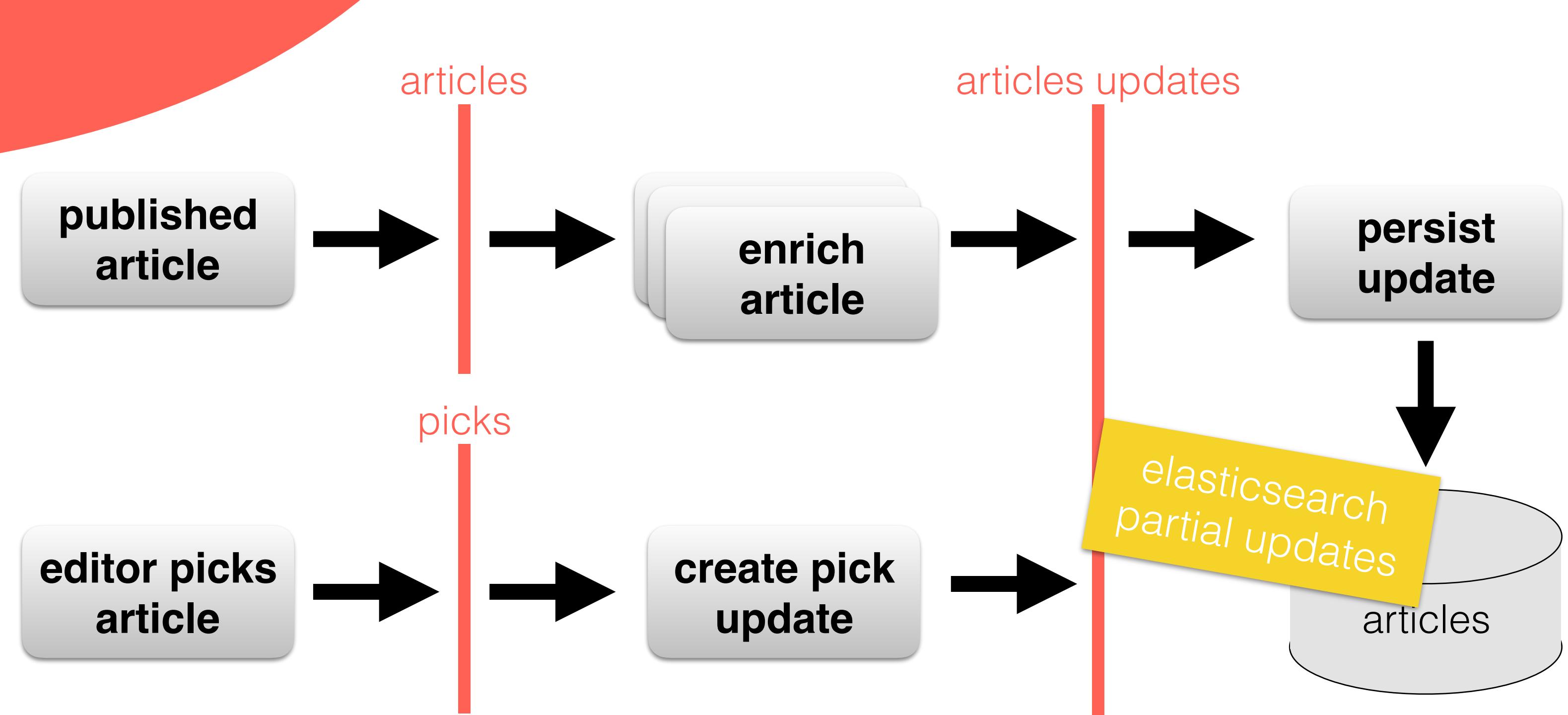


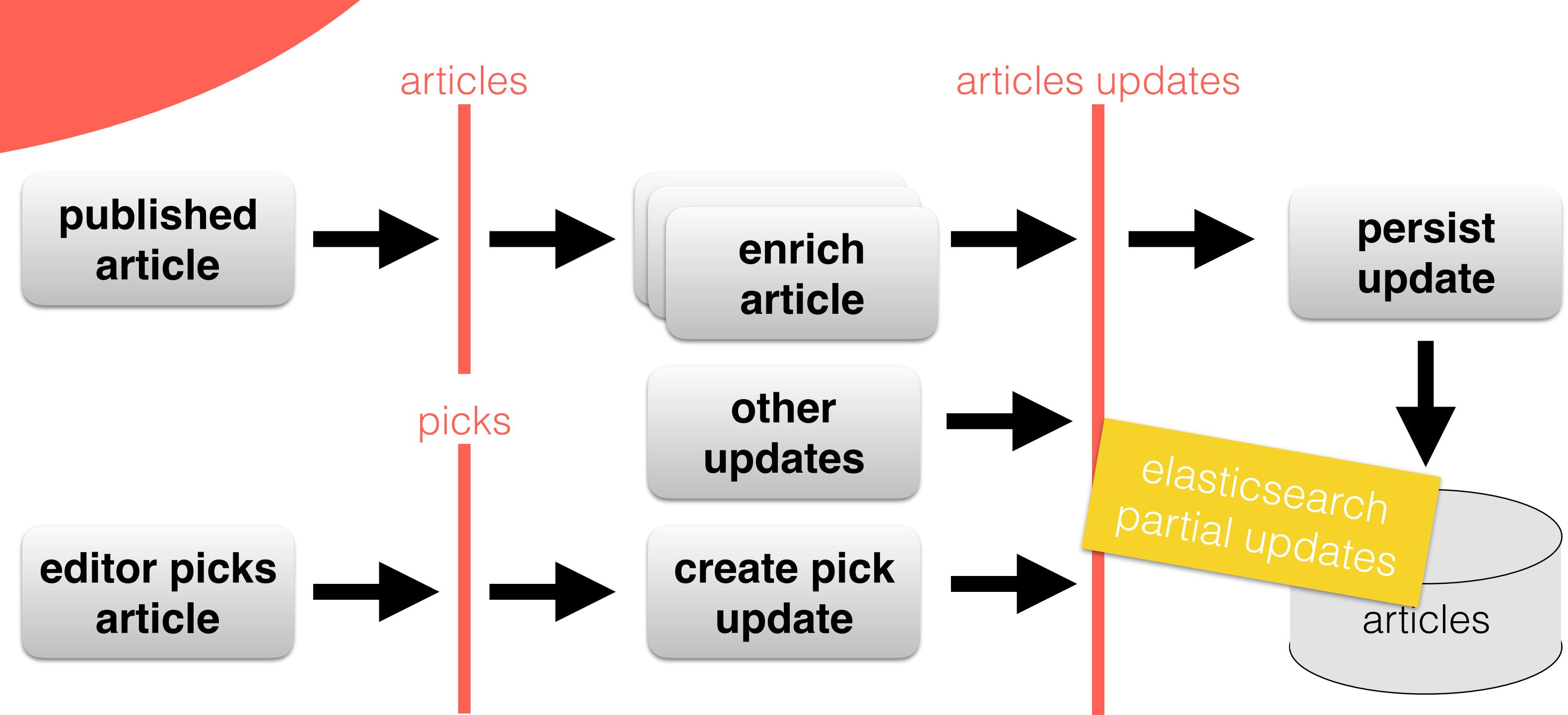


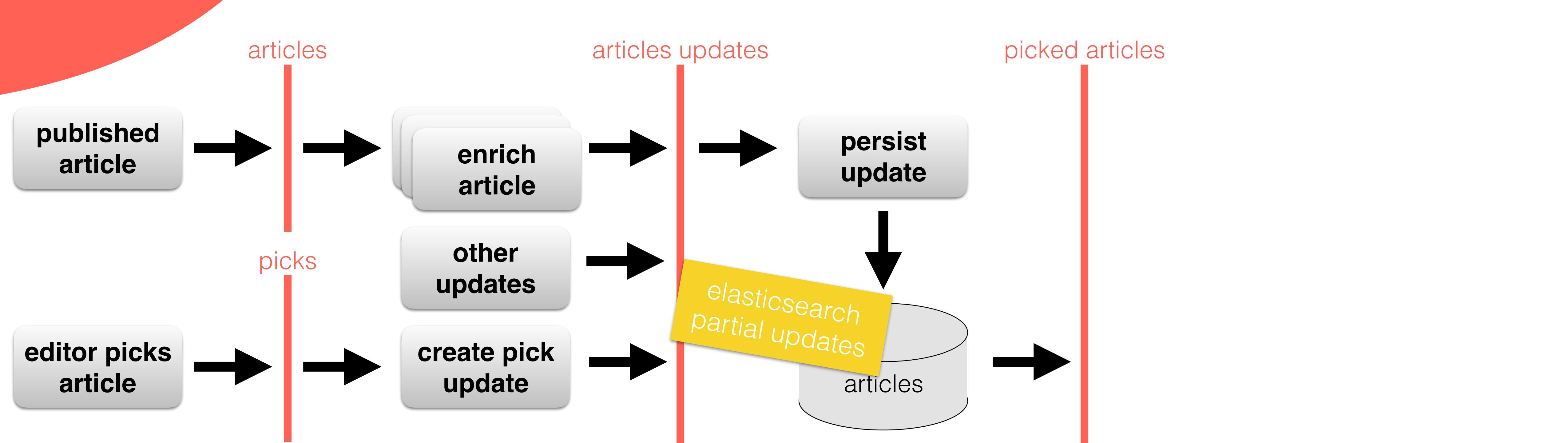


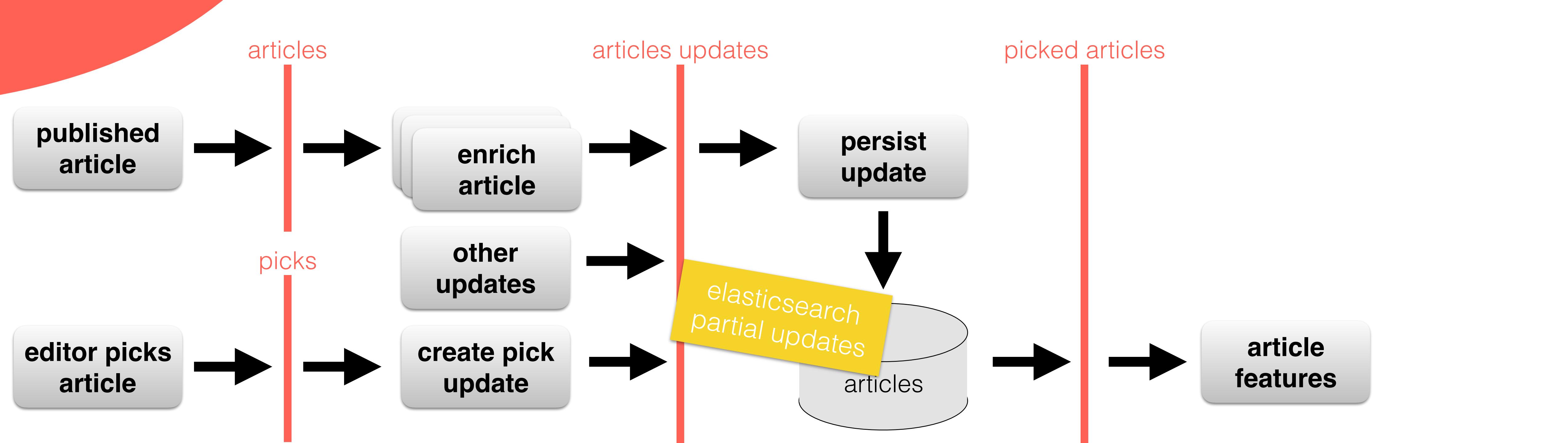


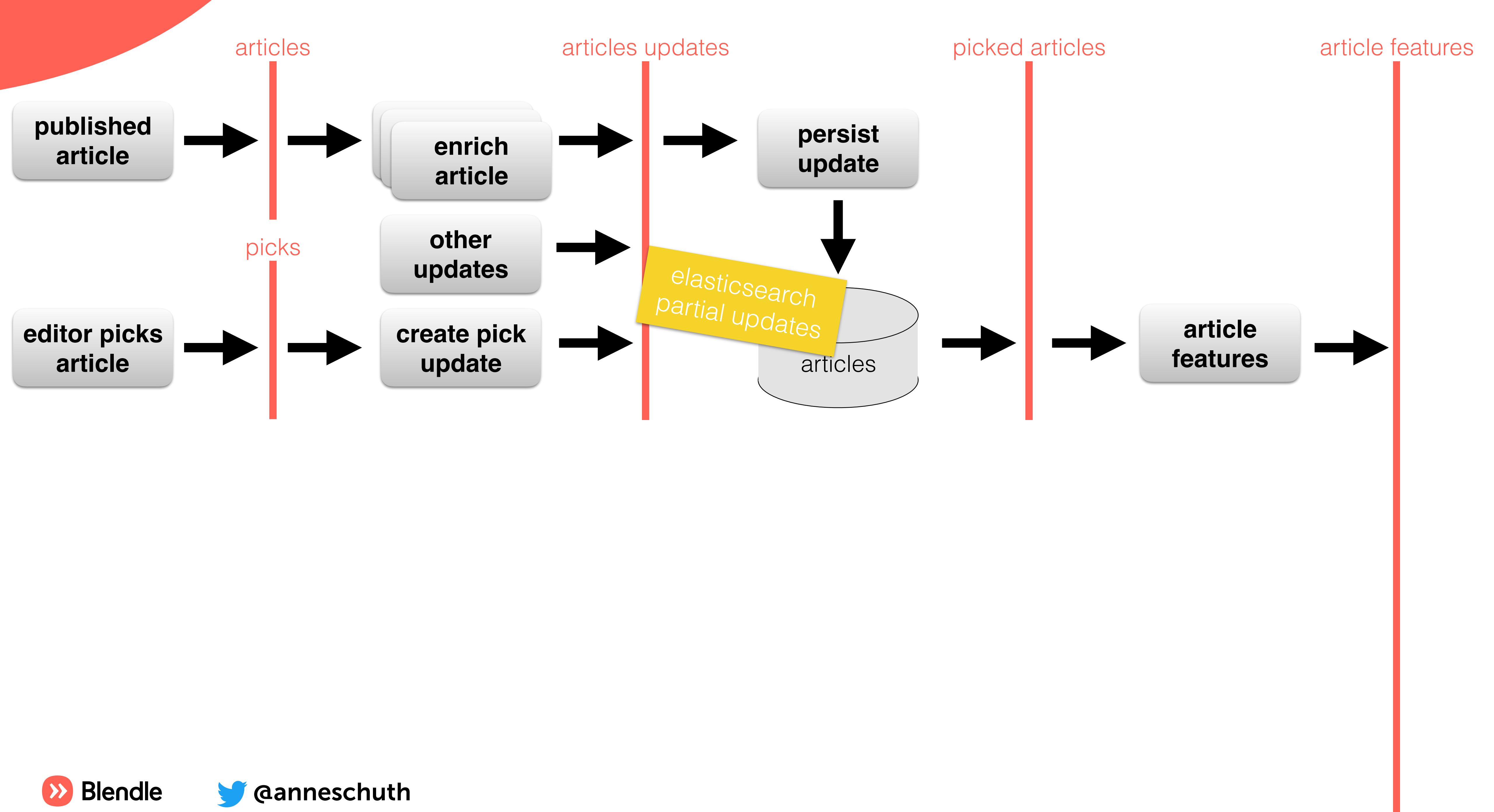


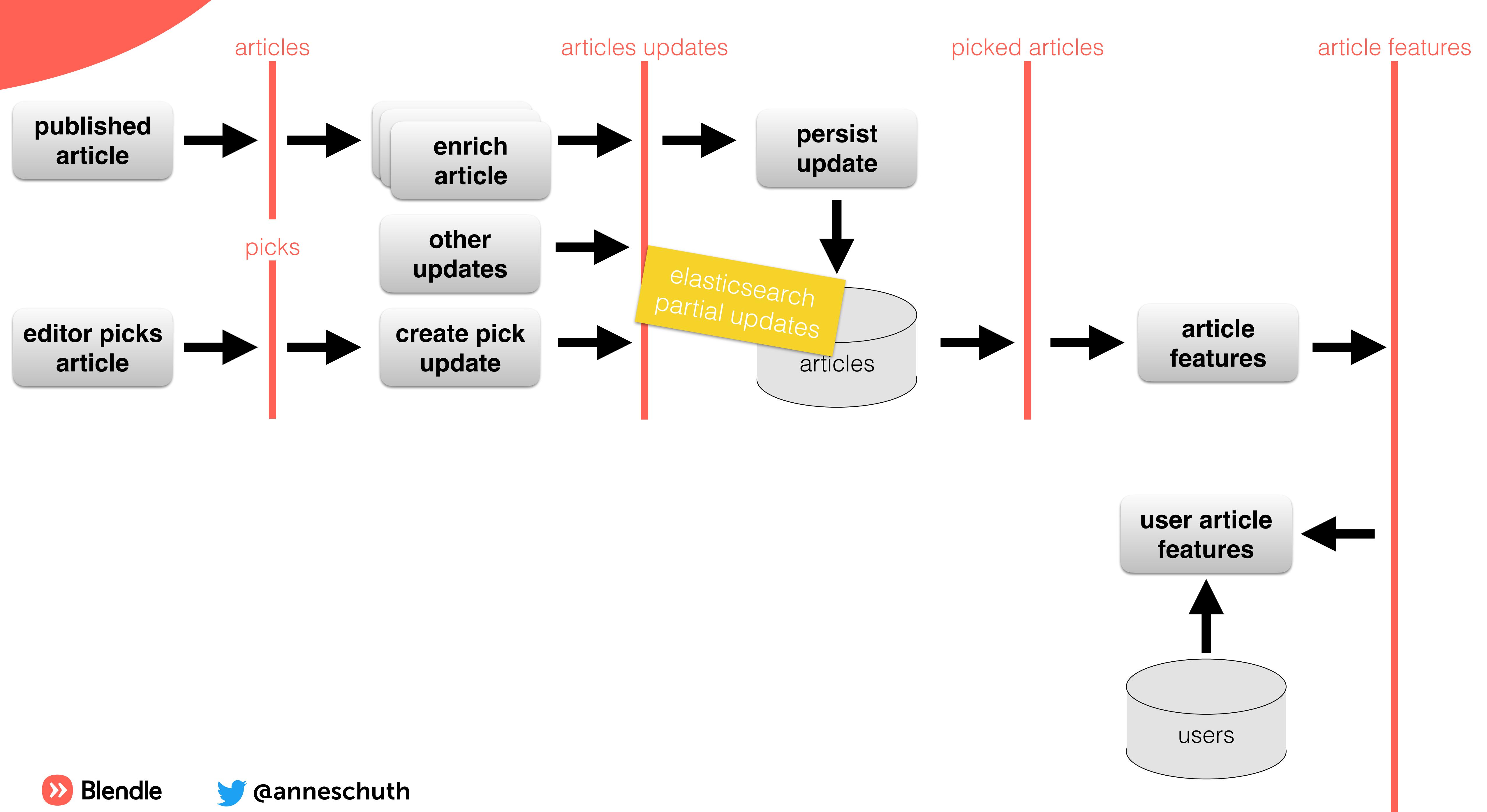


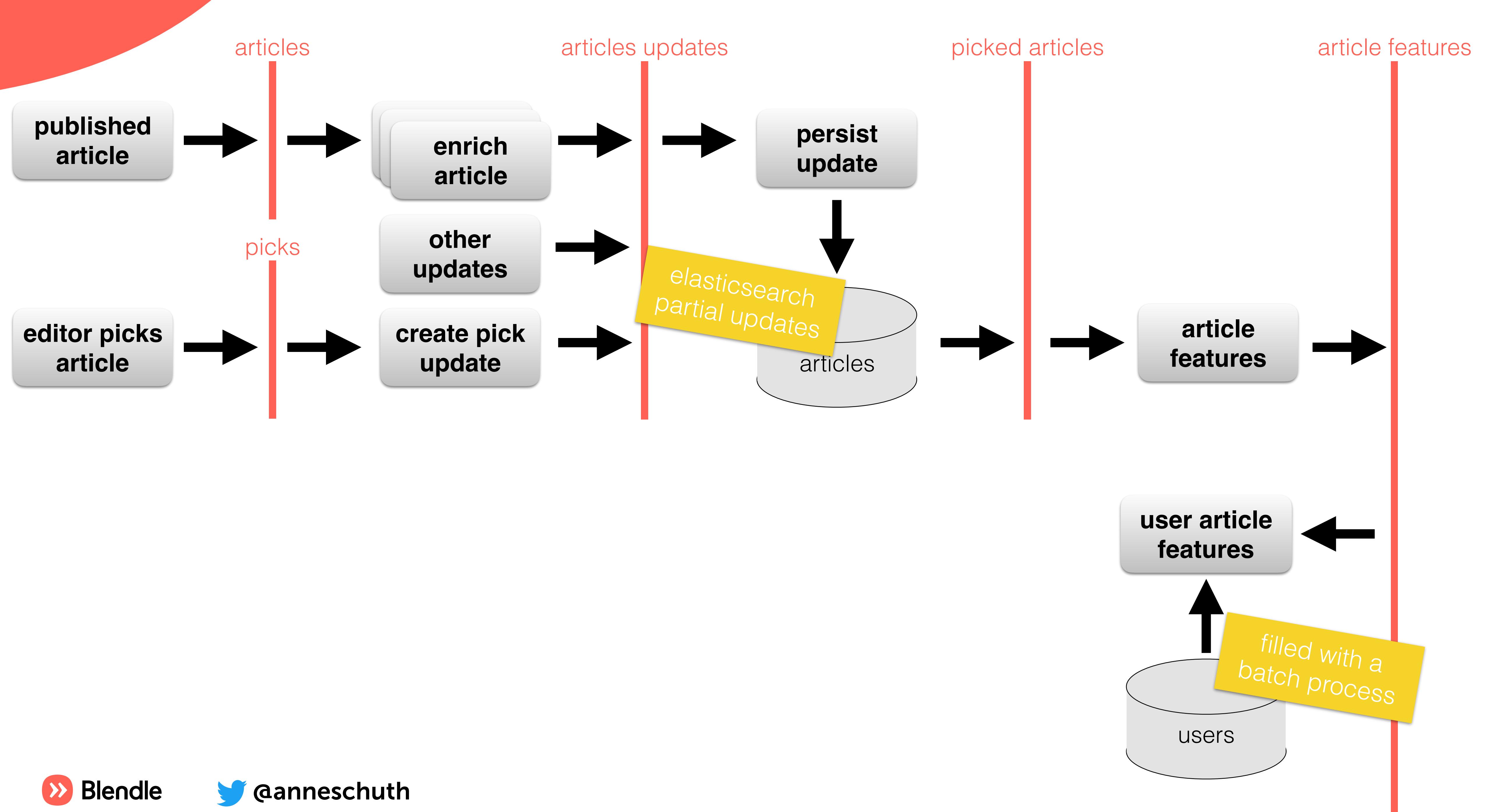


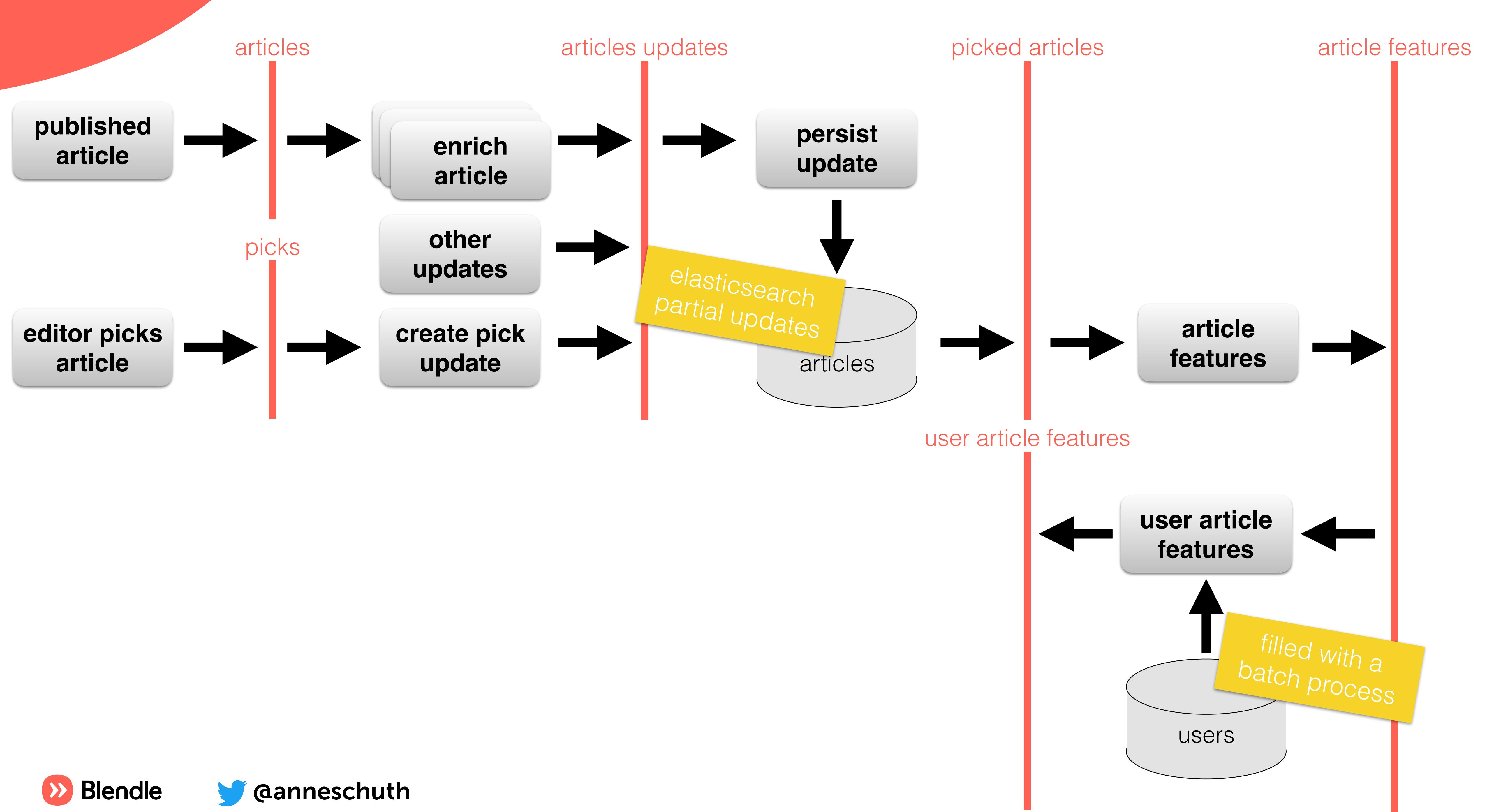


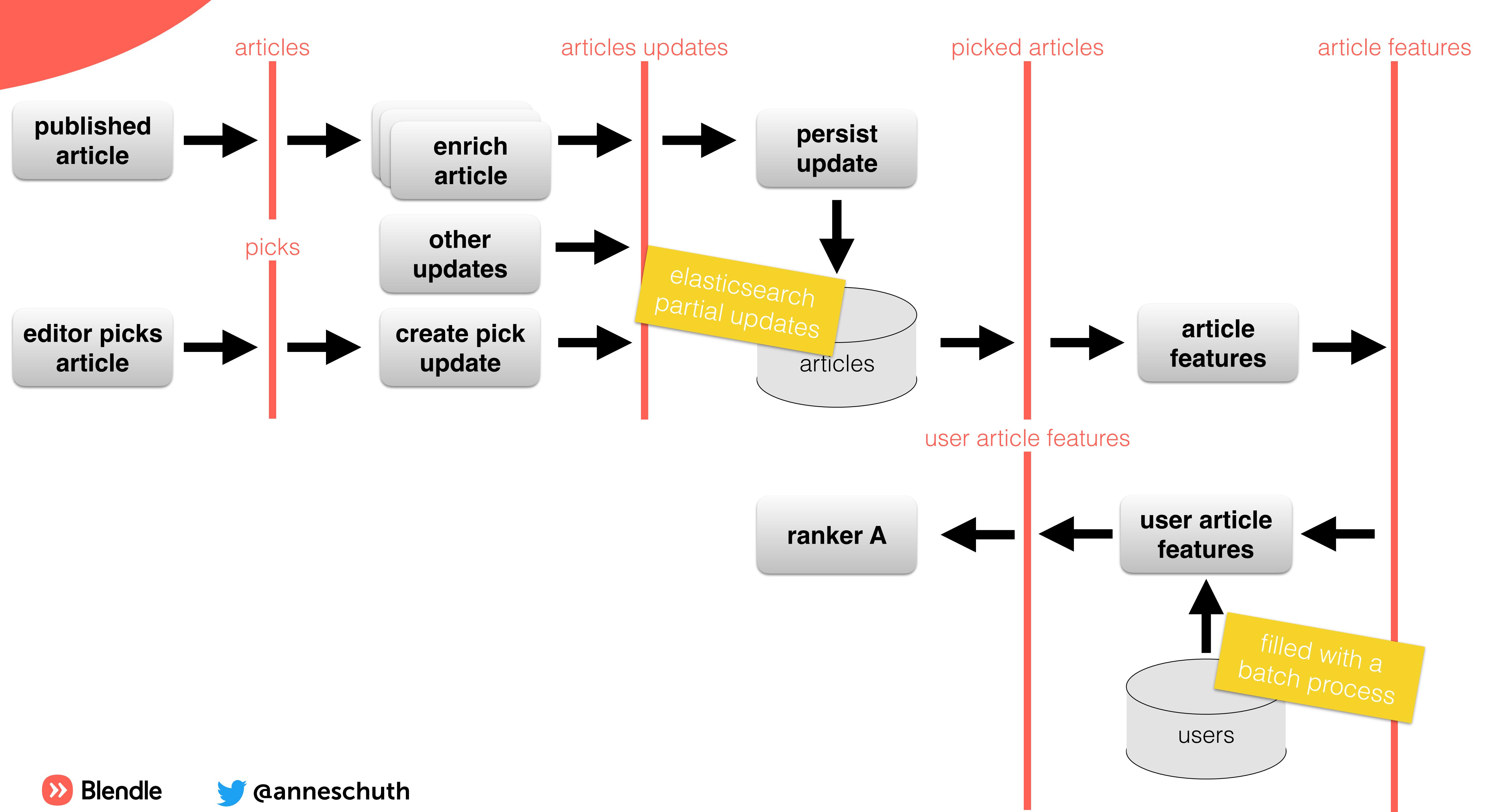


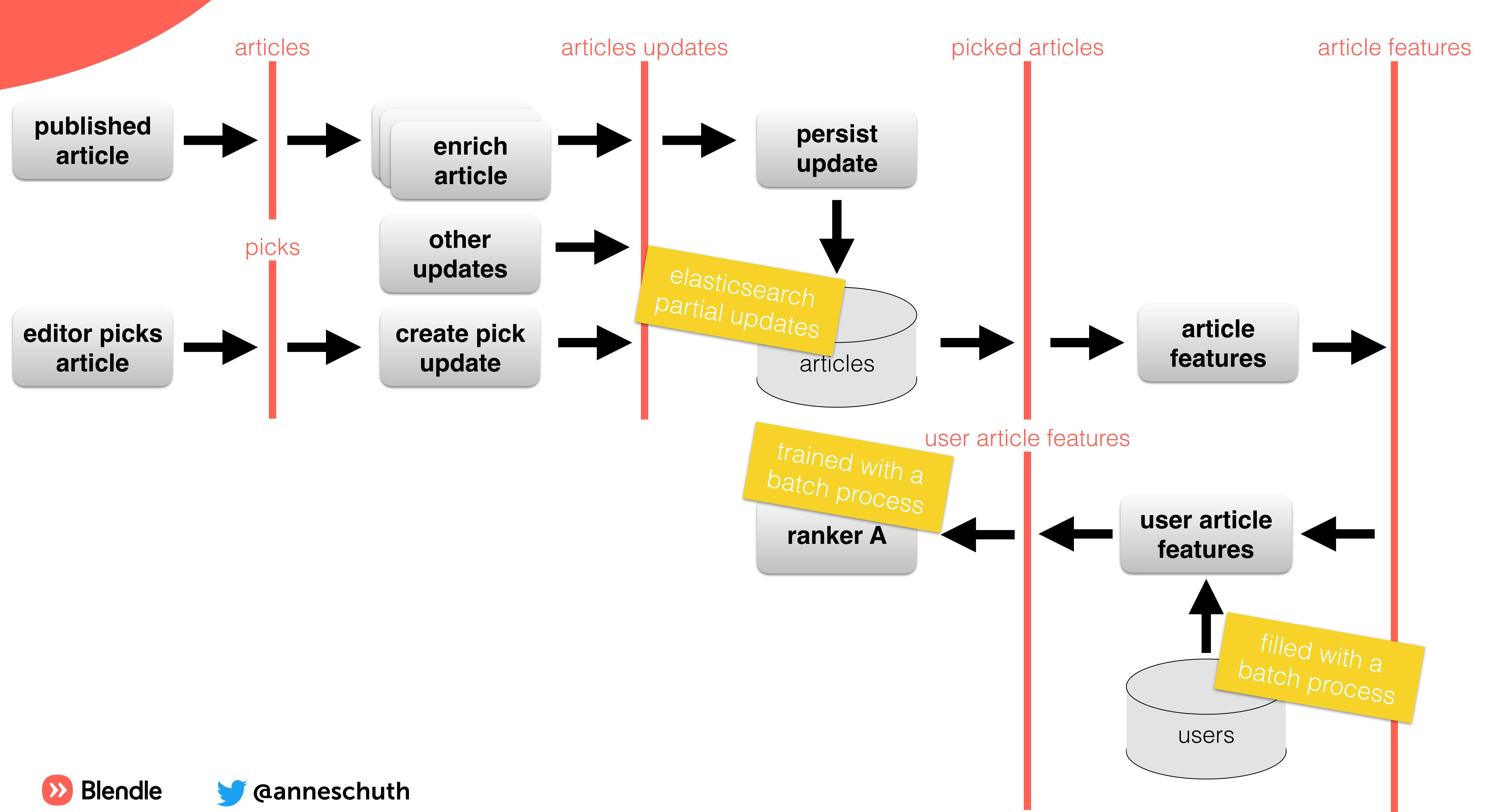


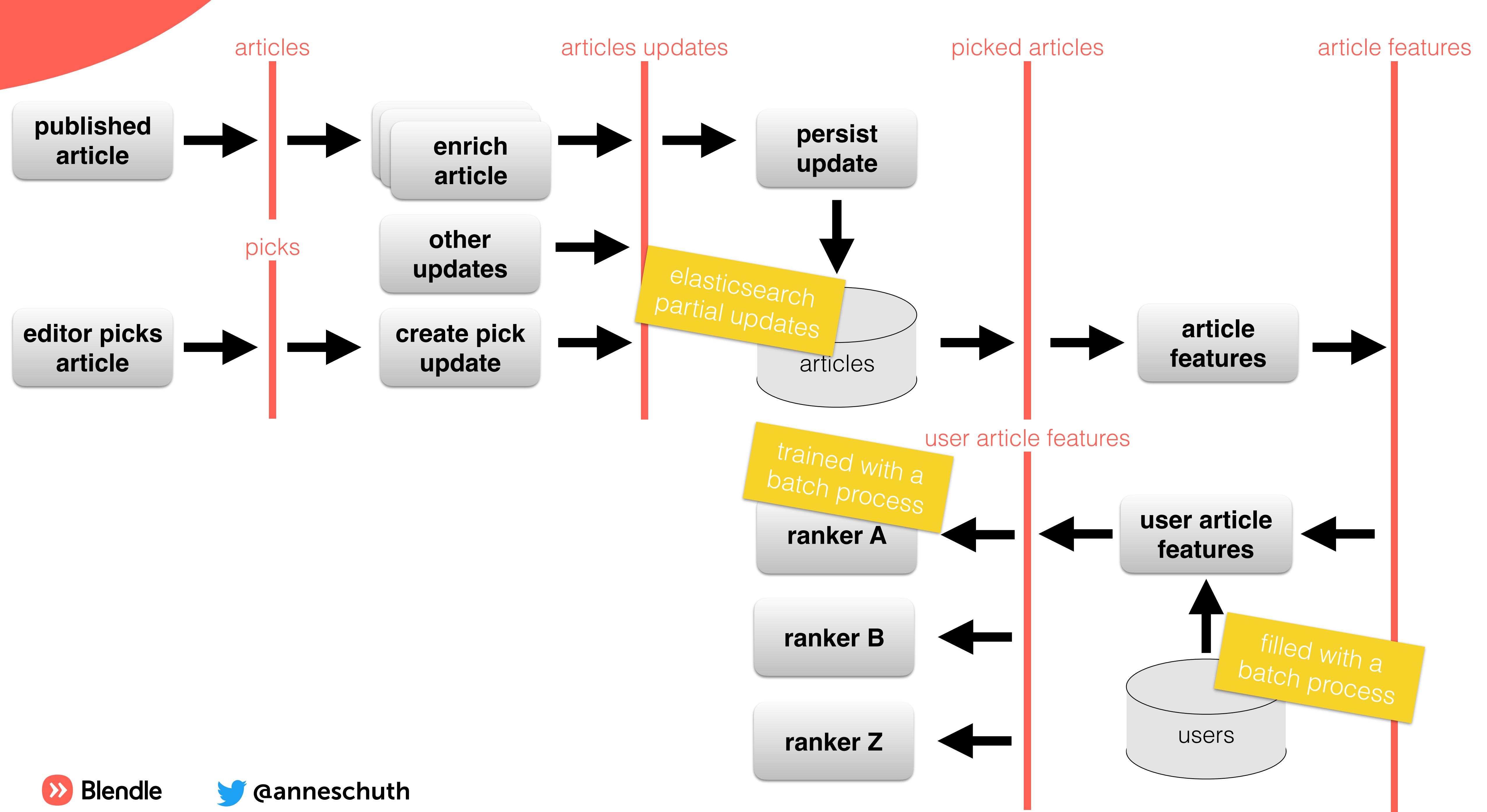


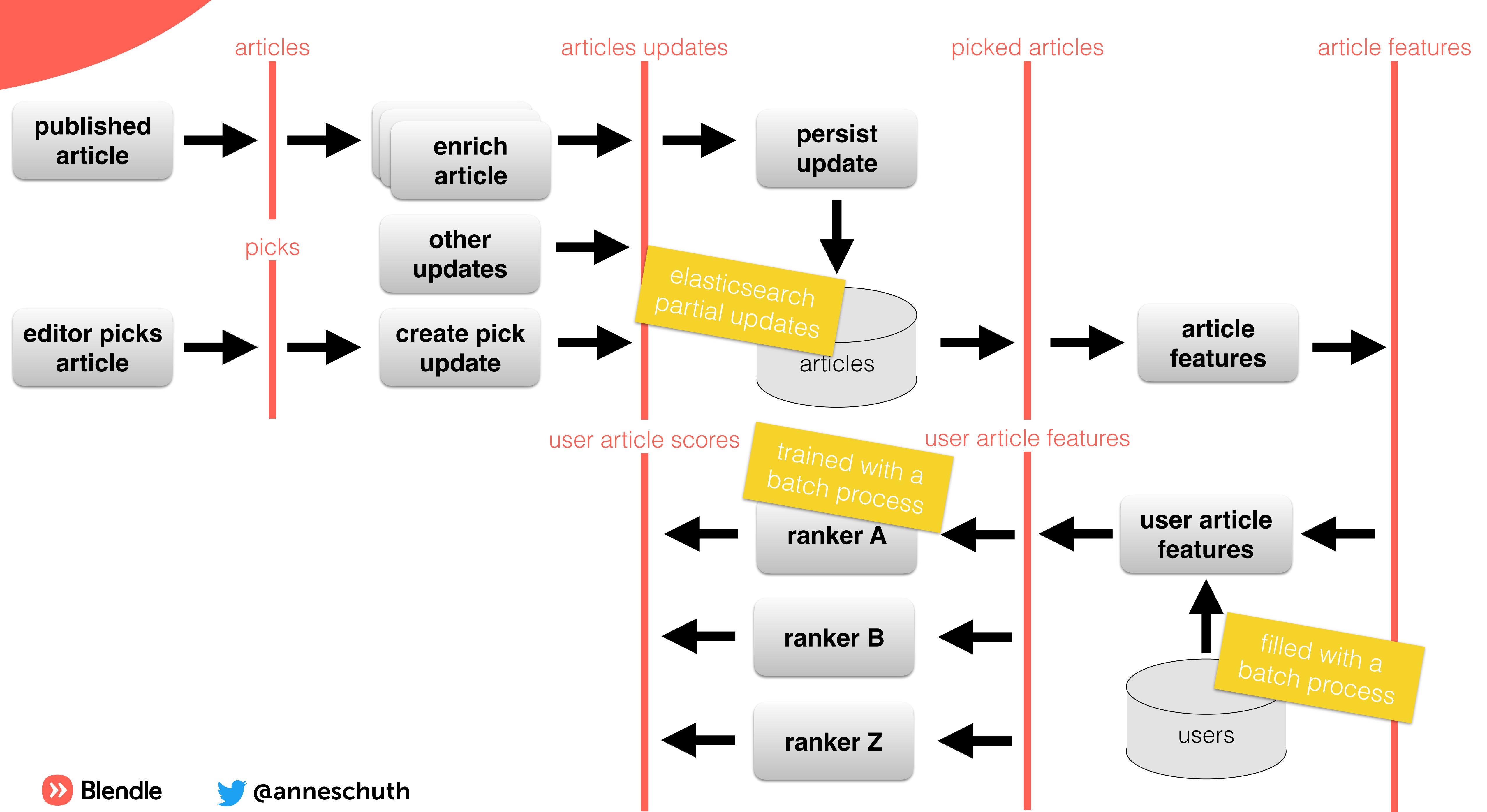


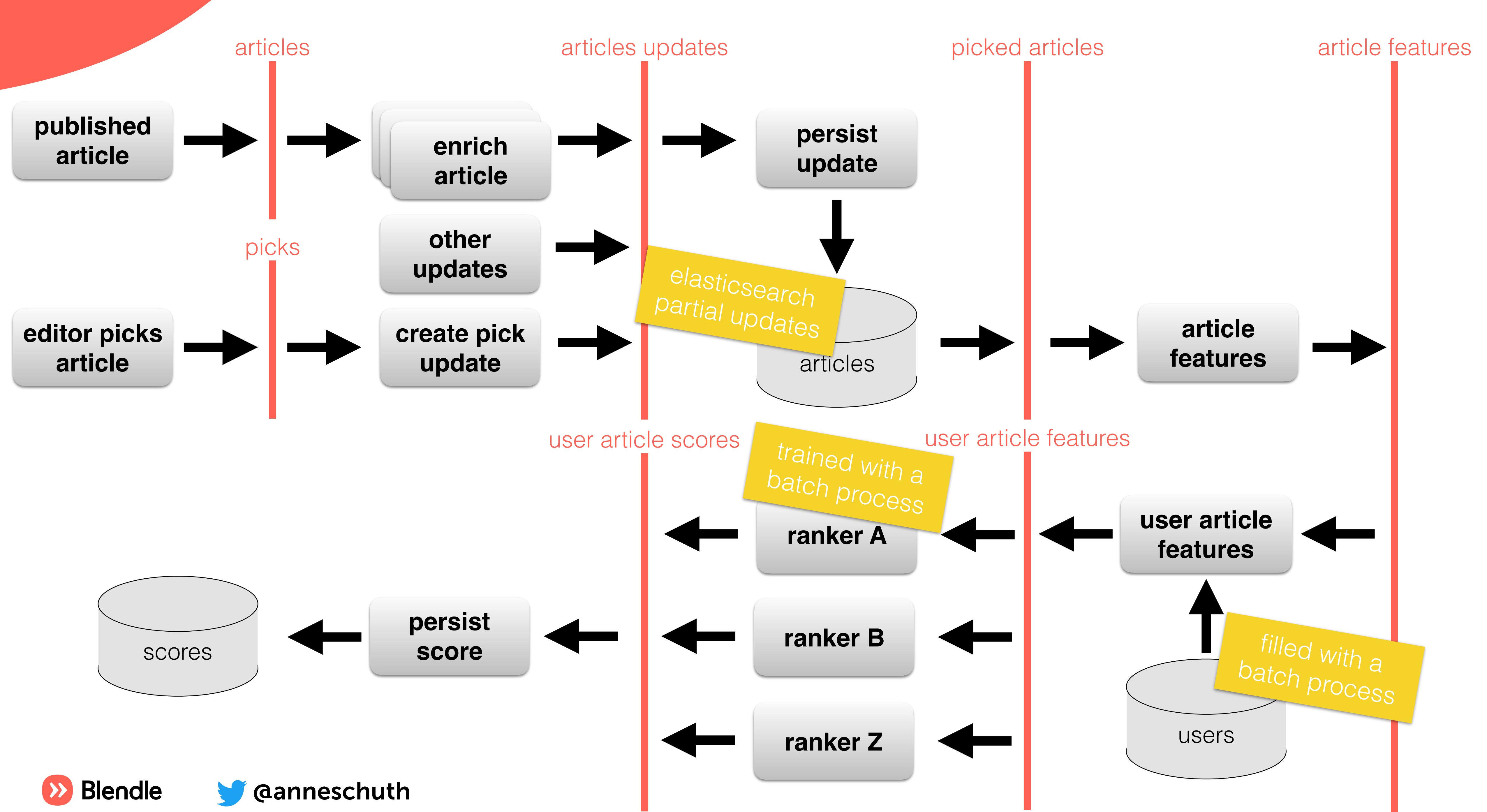


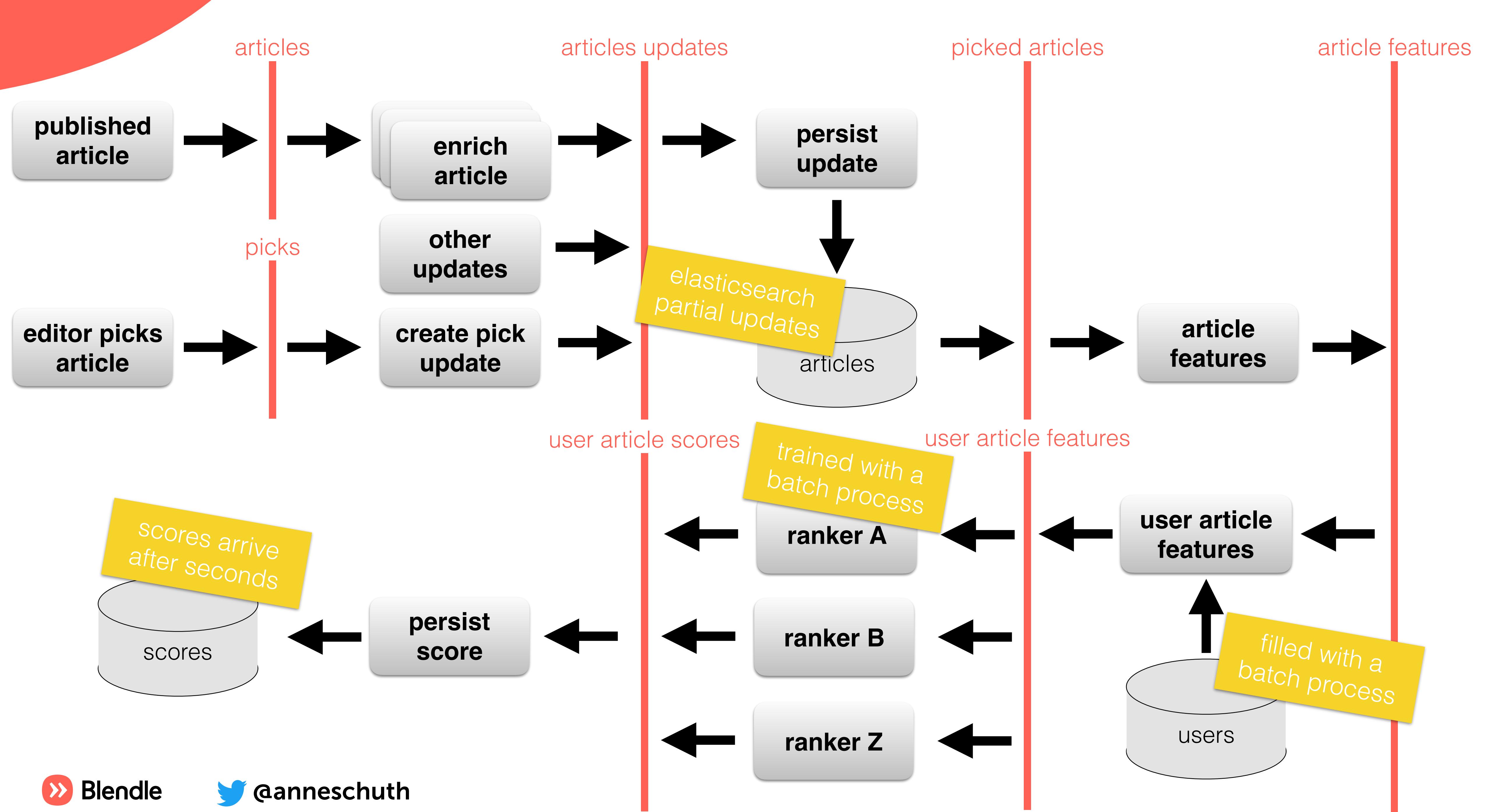


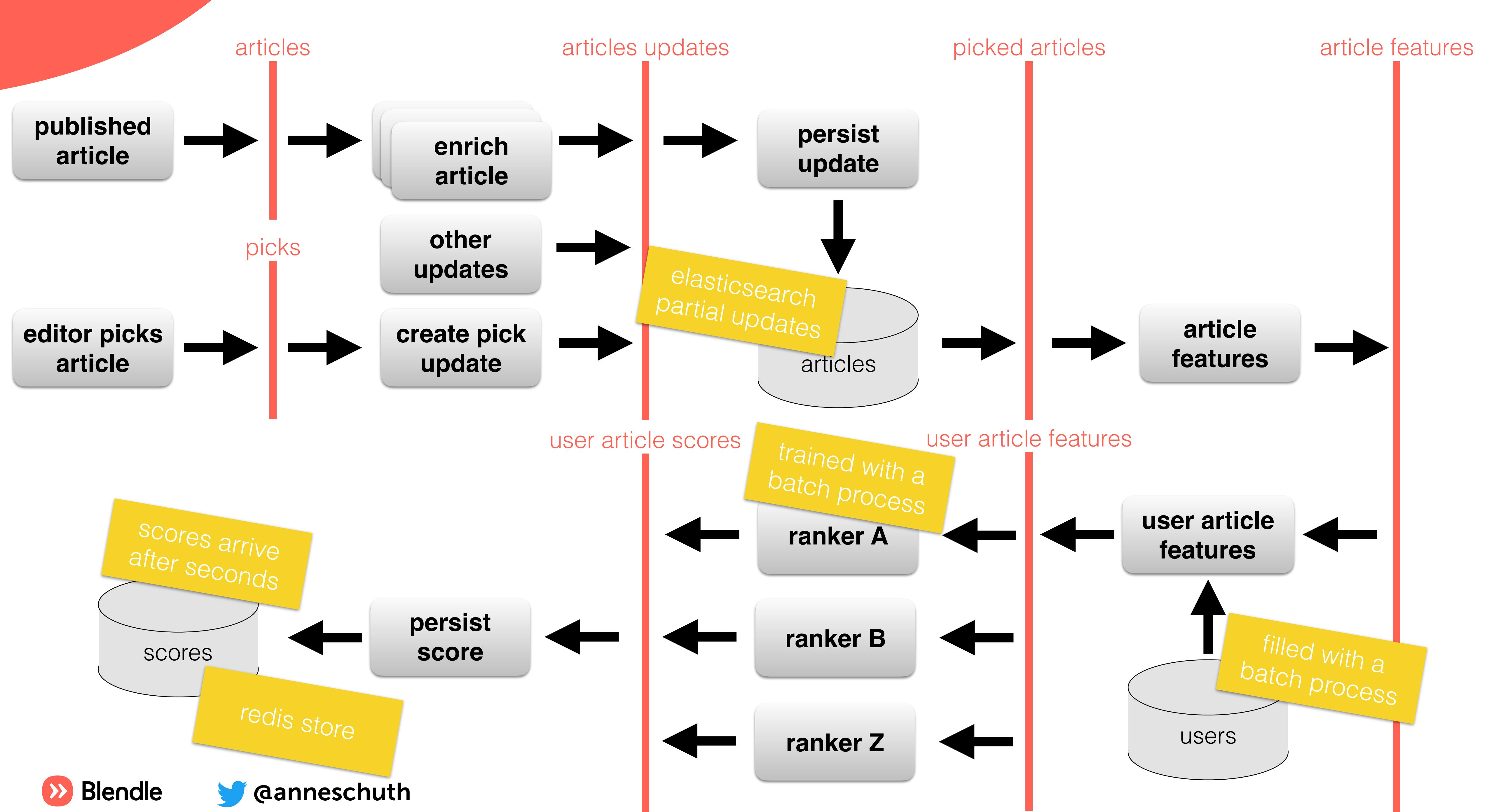


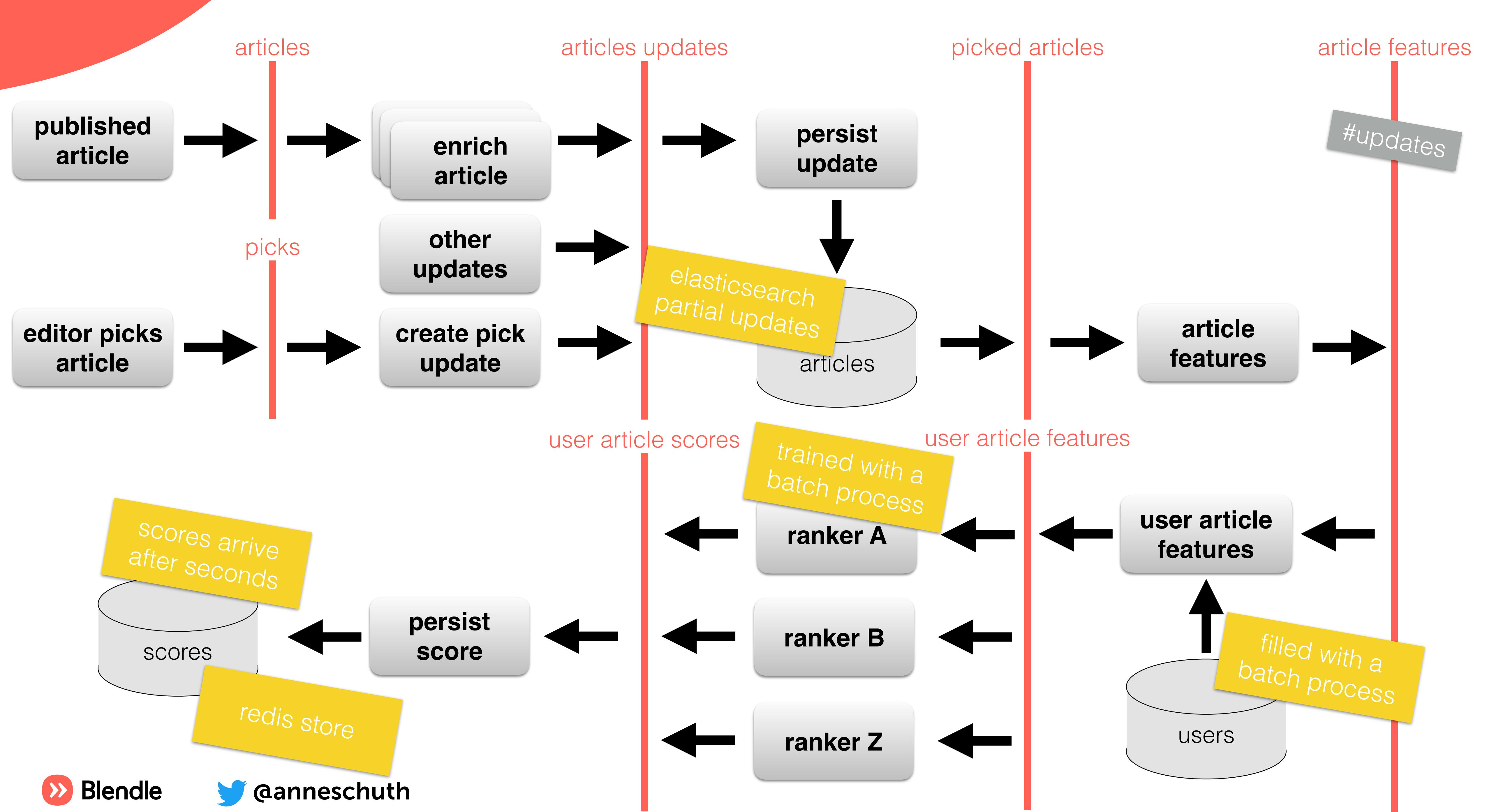


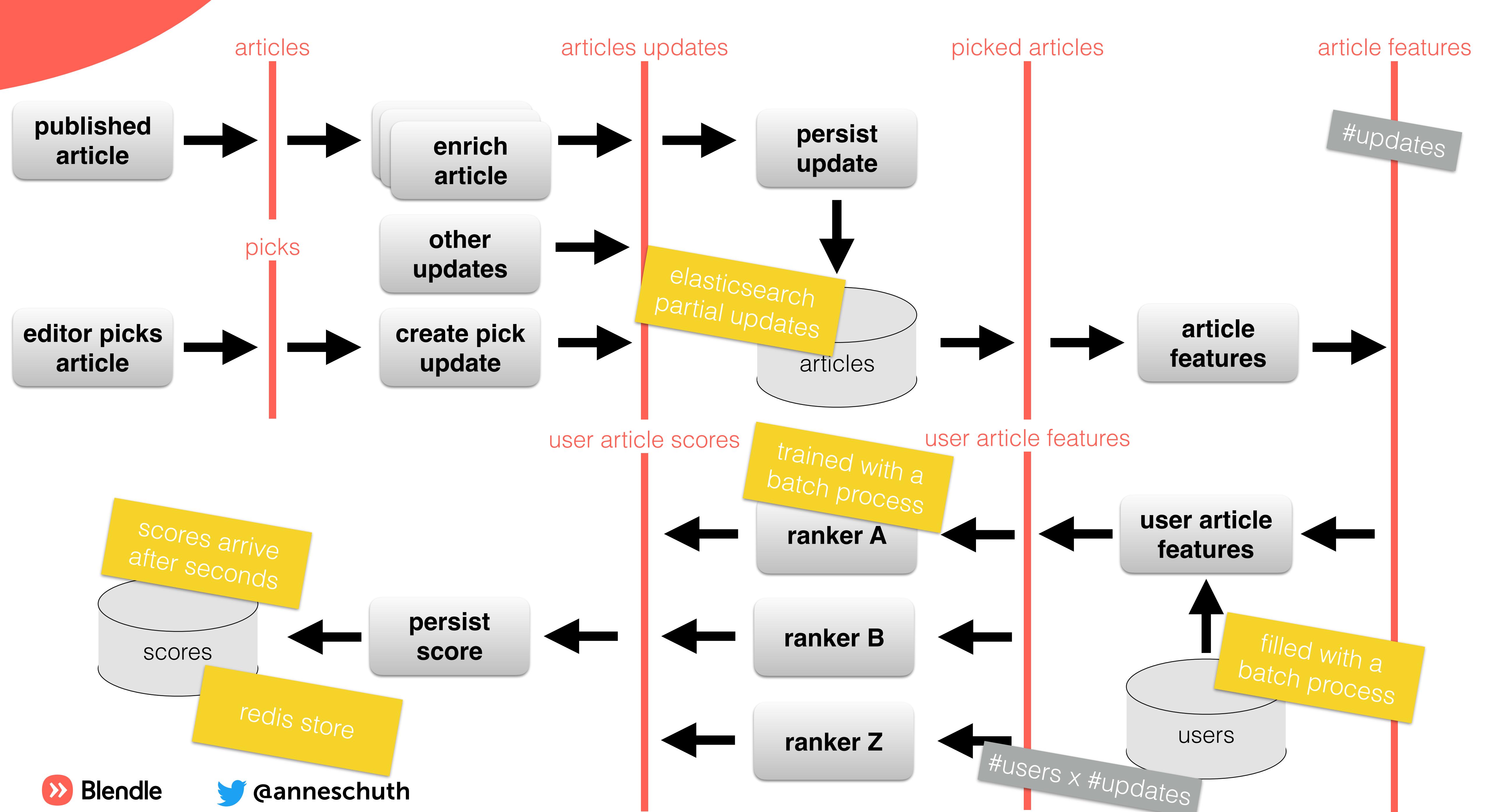


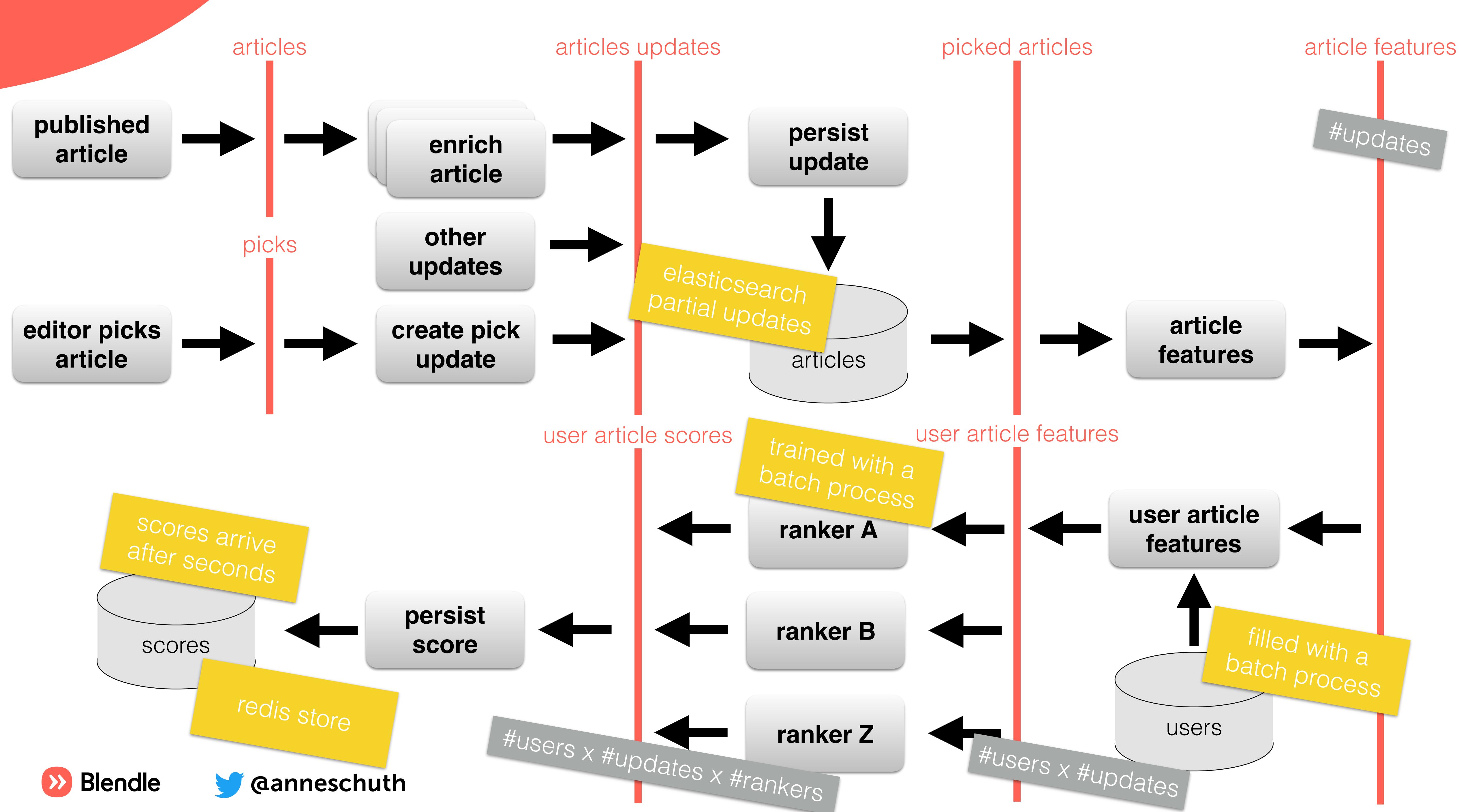




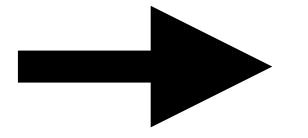




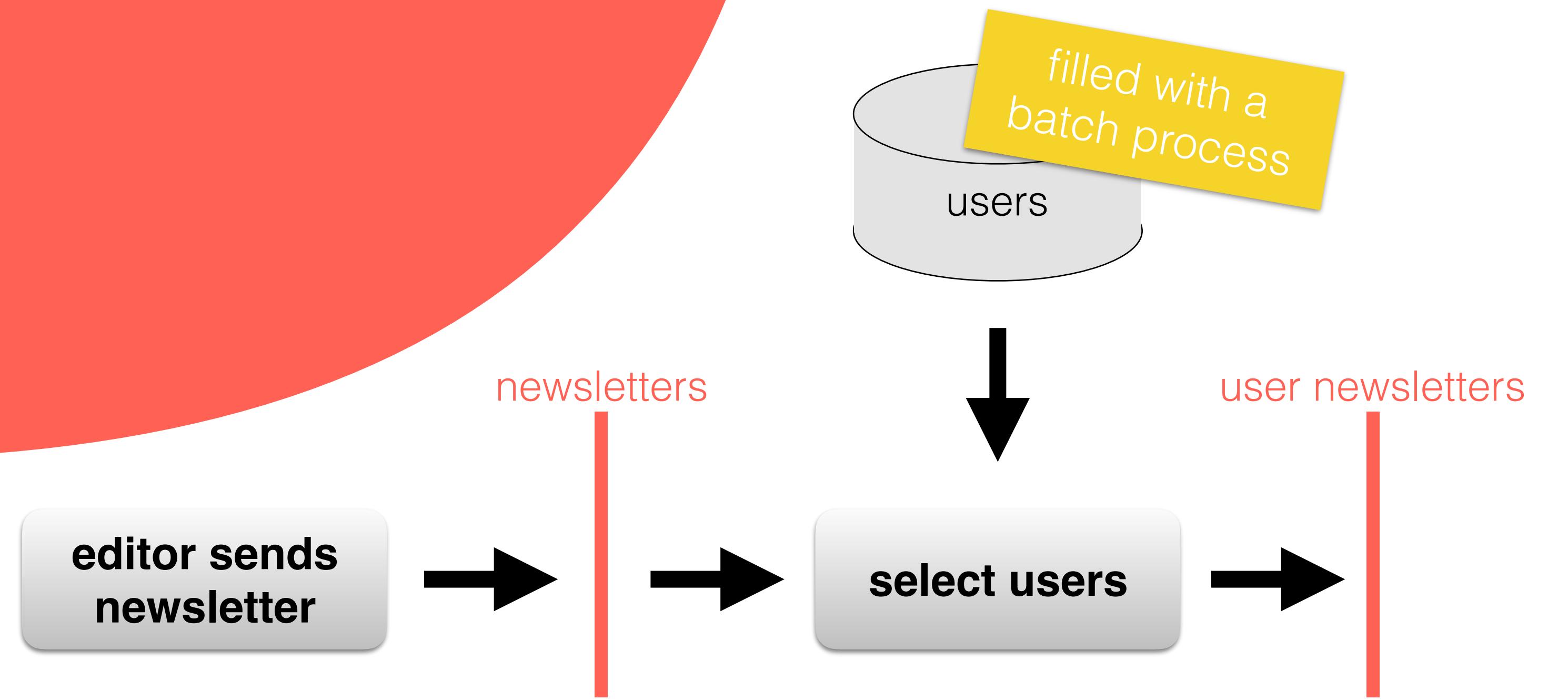


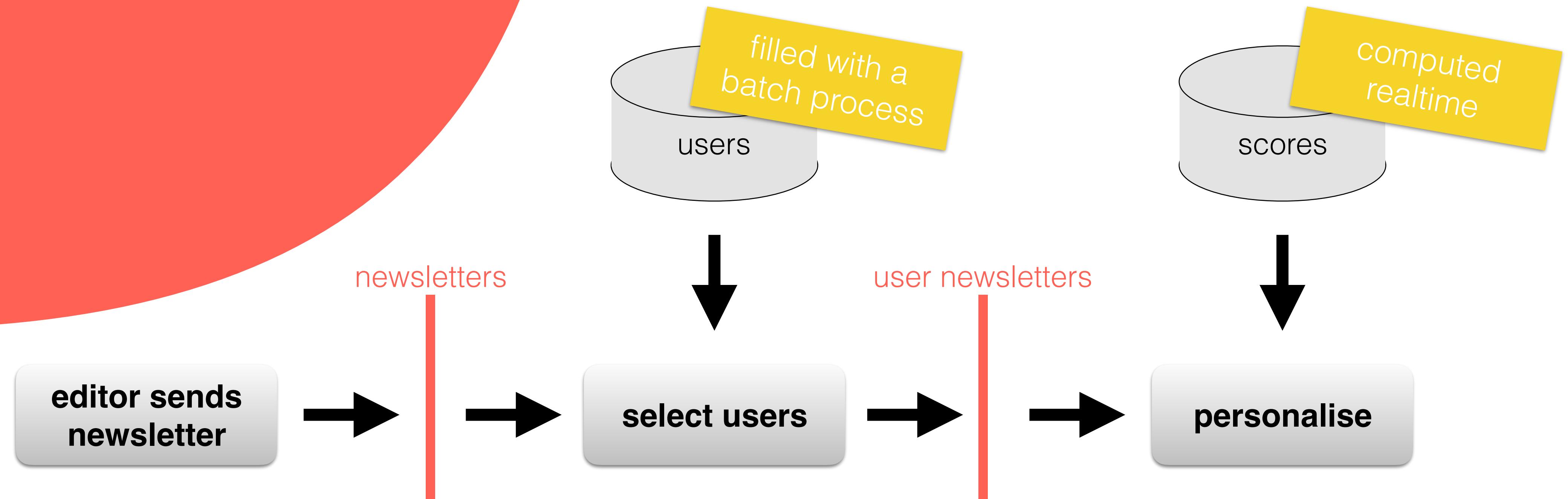


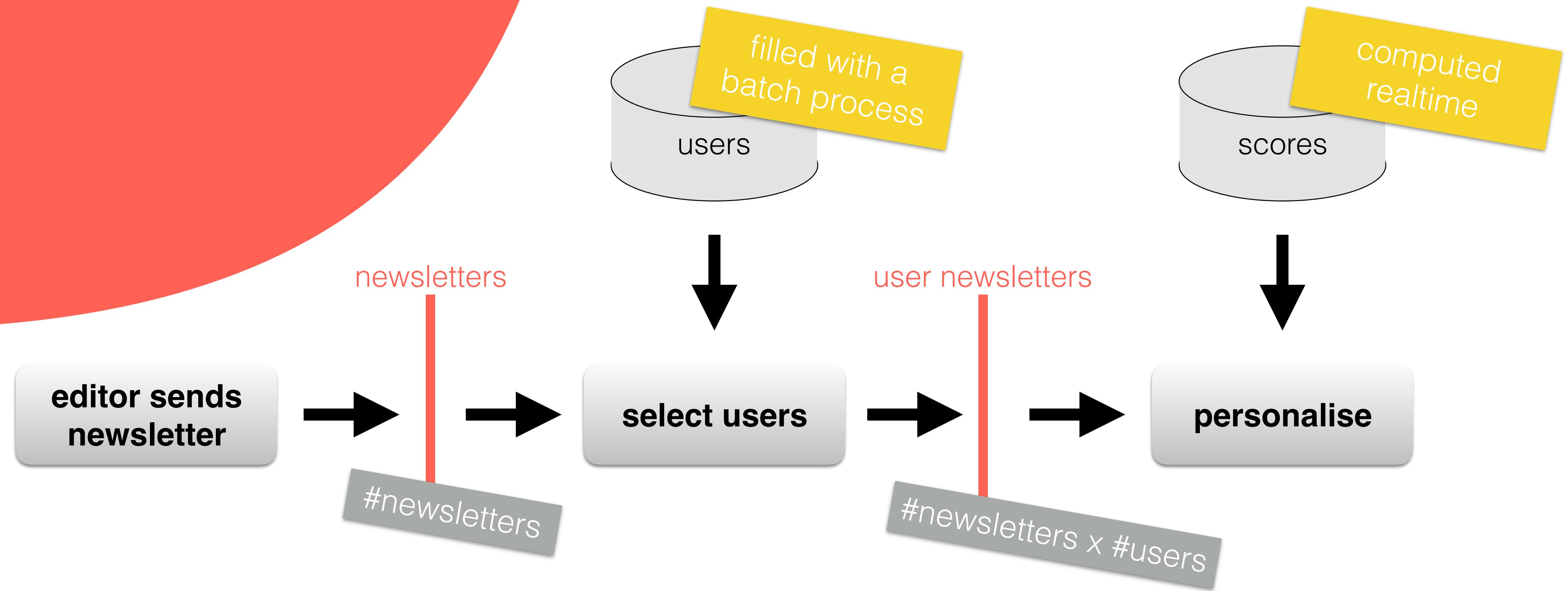
**editor sends  
newsletter**

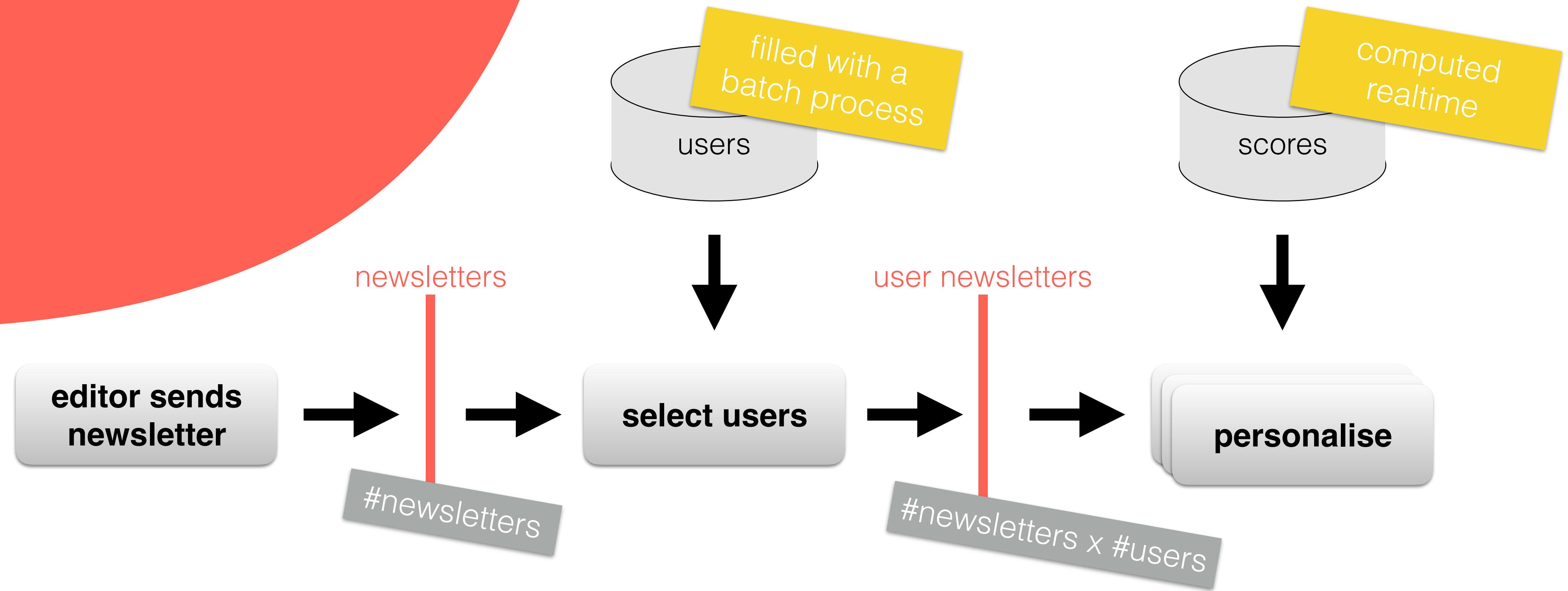


newsletters





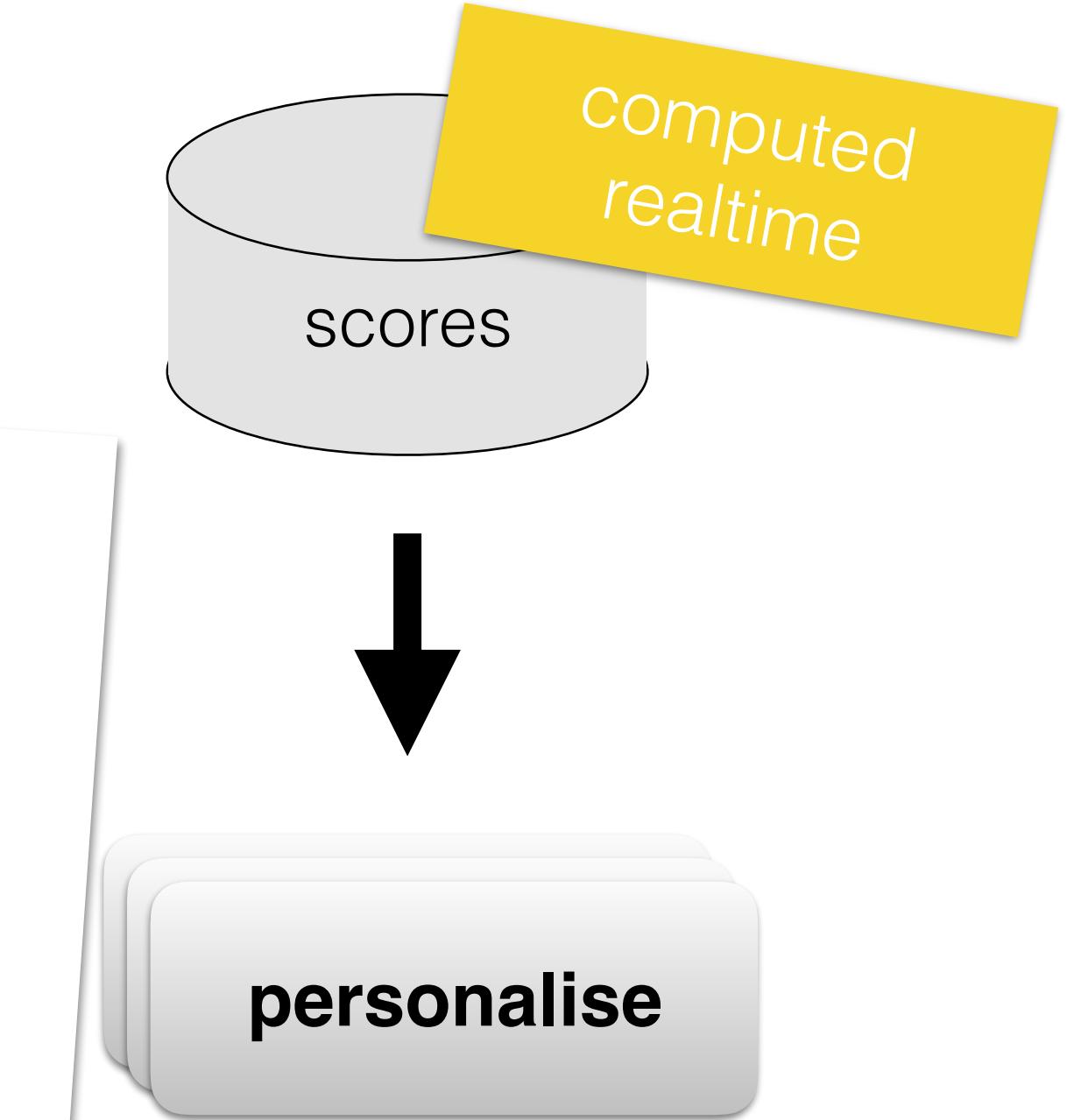






# Docker

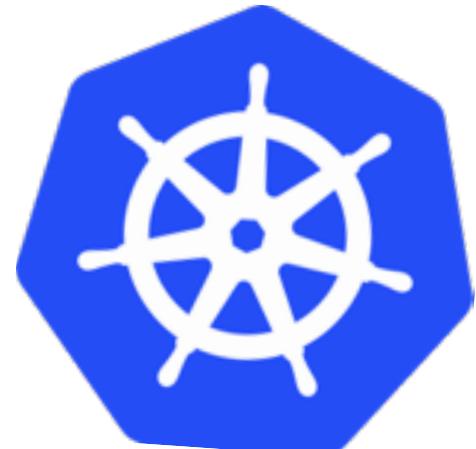
- containerise tasks
- easy dev/test/deploy





# Docker

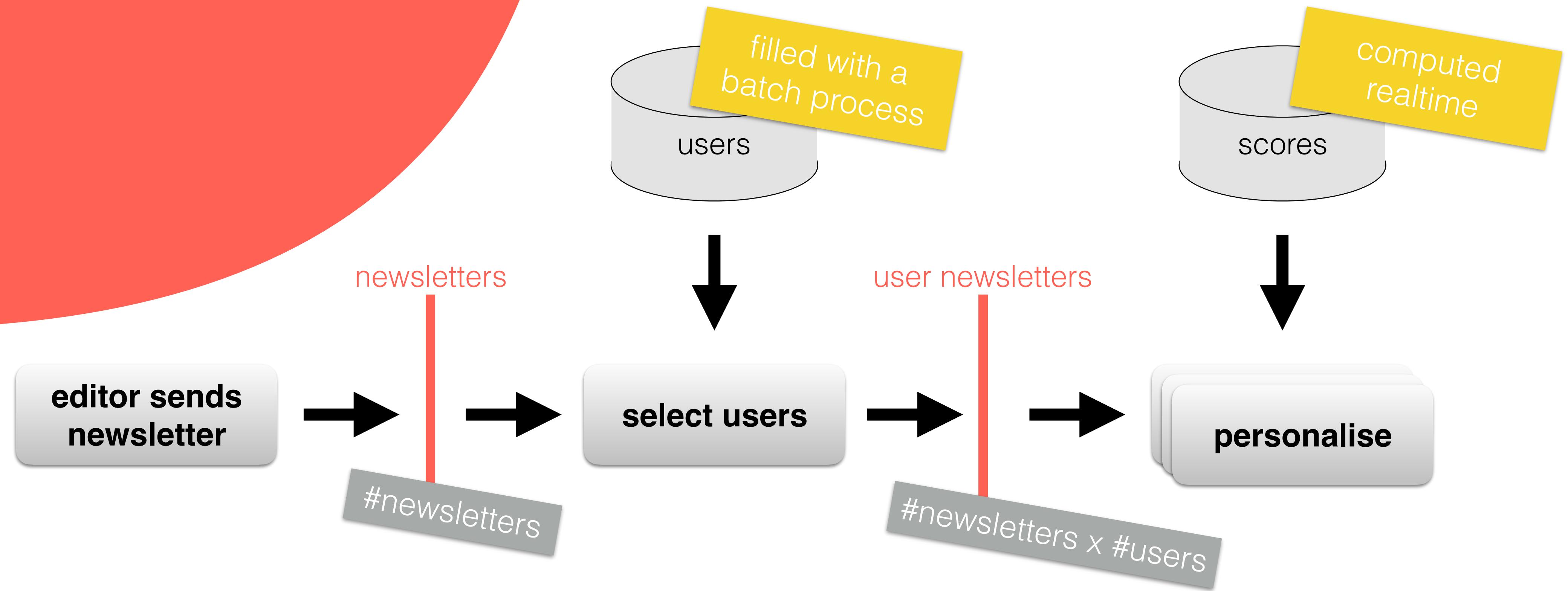
- containerise tasks
- easy dev/test/deploy

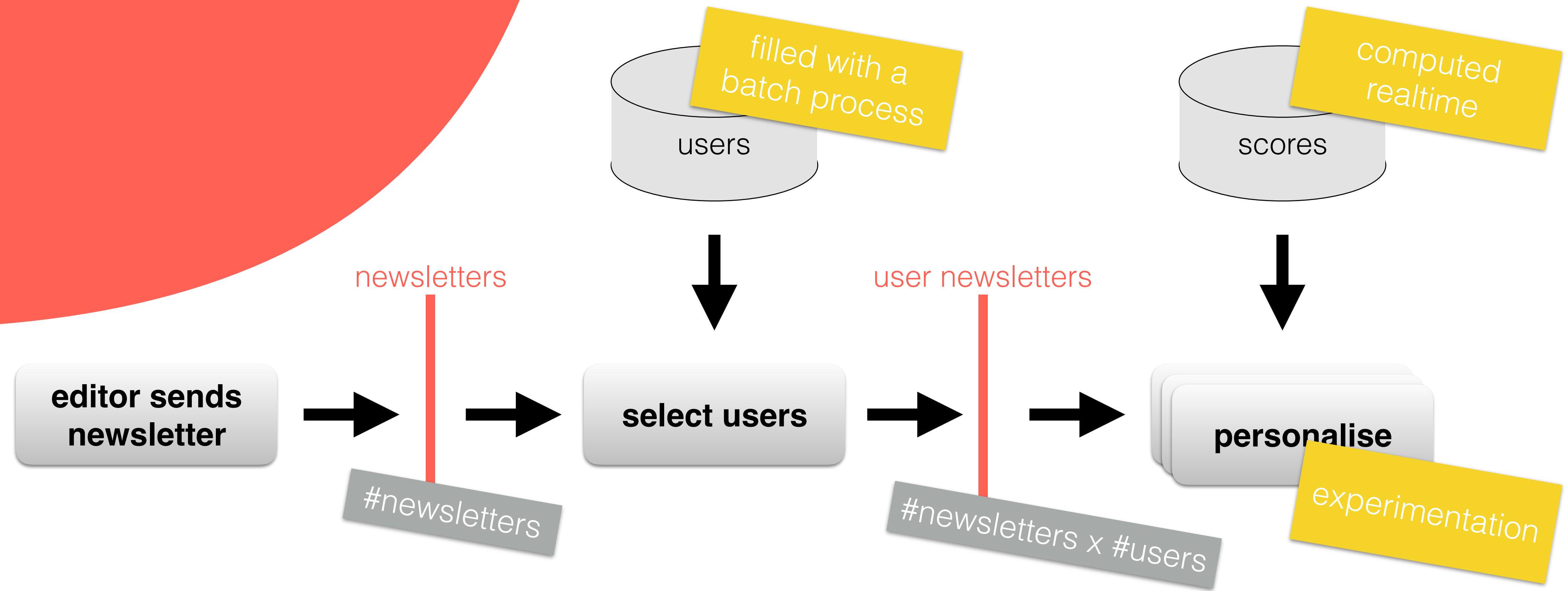


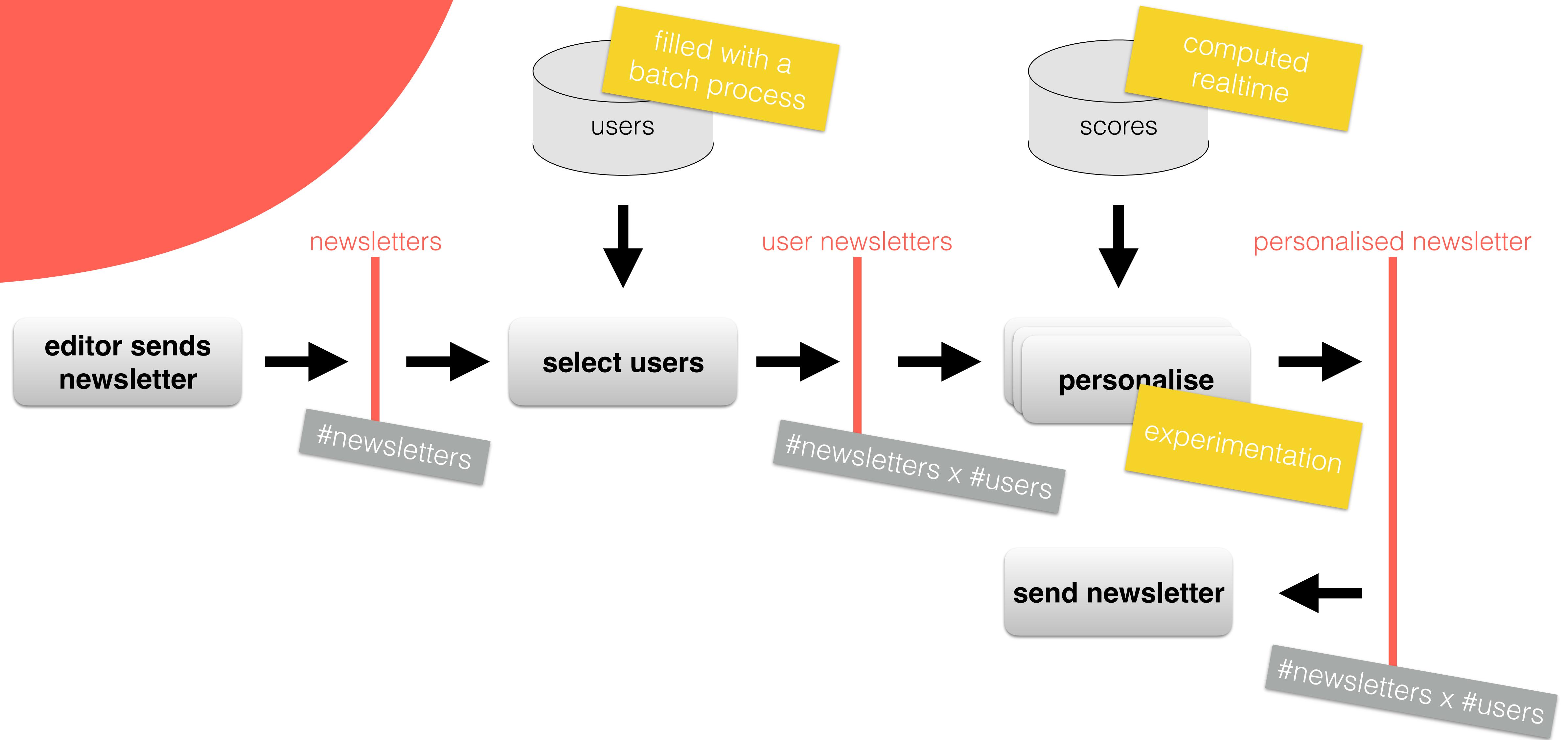
# Kubernetes

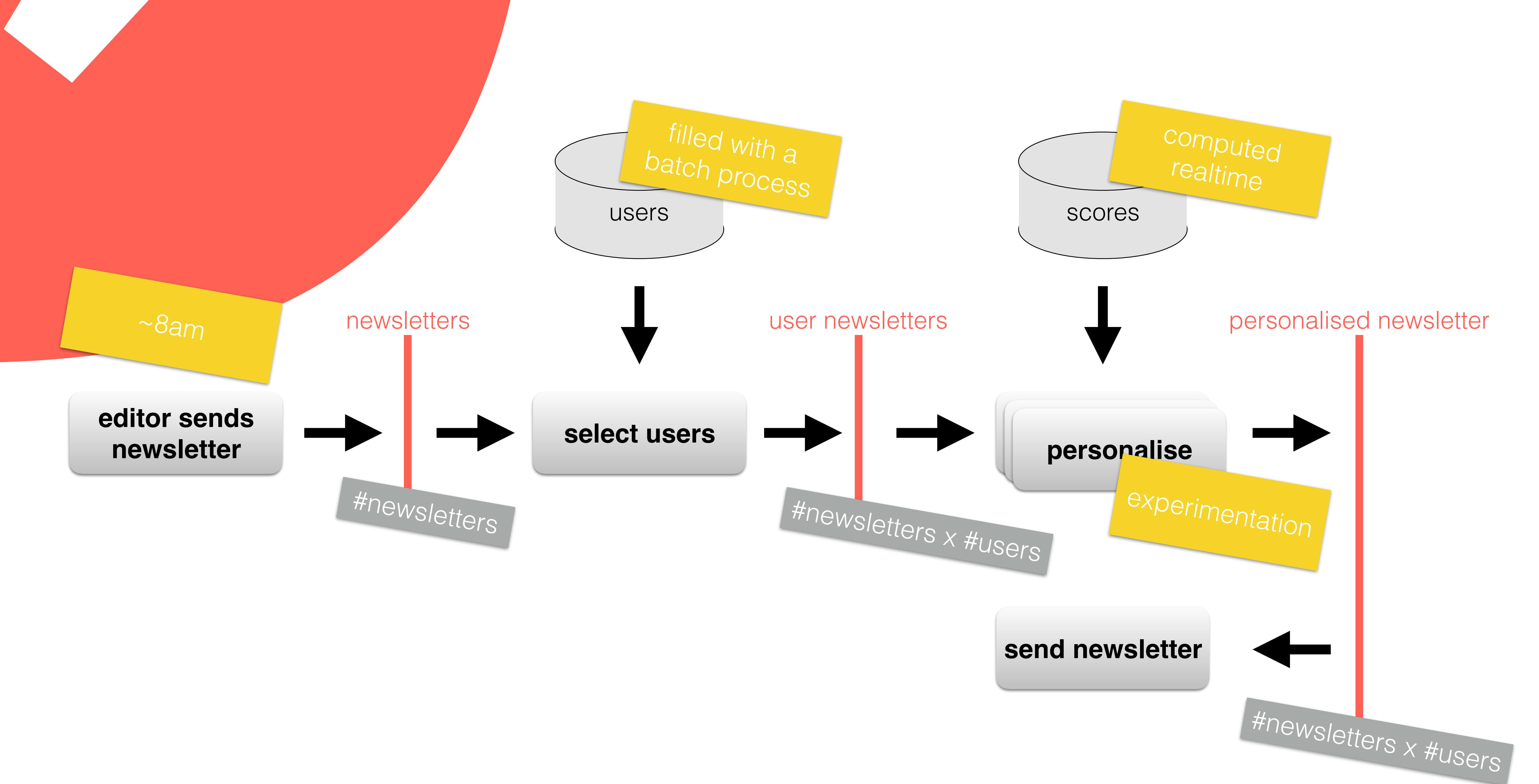
- orchestrate deployment
- autoscaling

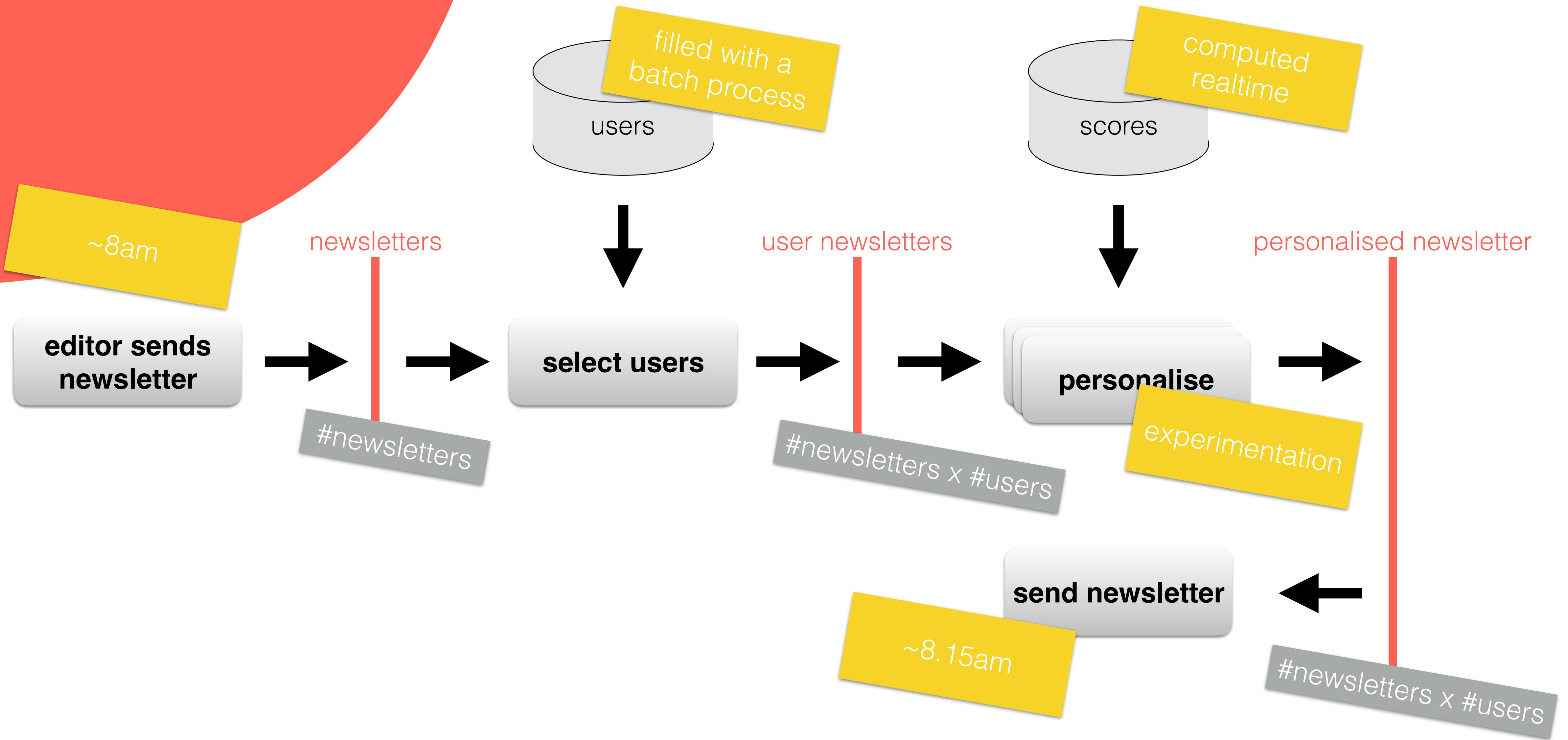












# Components we use for Machine Learning



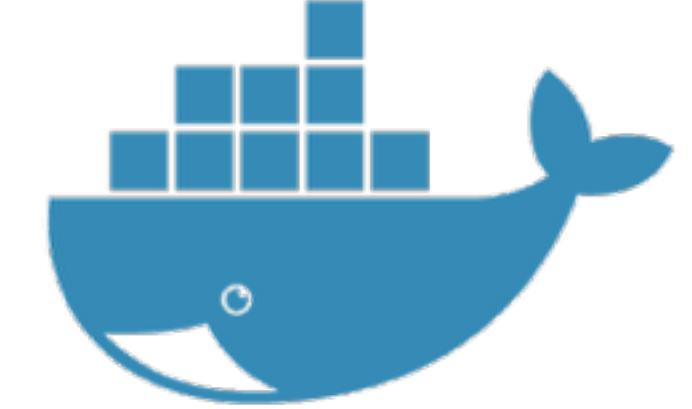
*Luigi*  
- plumbing



*Cucumber*  
- BDD development

*cucumber*

*Scikitlearn*  
- experimental ML



*Docker + Kubernetes*  
- containerisation



*Redis*  
- fast access



*redis*



*Kafka*  
- streaming



*S3*  
- raw content



*Spark*  
- processing large data



*elastic*

*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

# 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

# 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

# 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

# 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

The Blendle app interface is visible on the left, showing a list of articles from TIME. The first article is titled "Big Money's Tiny Impact on the Presidential Campaign" by Michael Scherer. The TIME magazine digital edition is displayed on the right, featuring a large "TIME" logo at the top. Below it is a black and white photograph of a man in a suit. To the right of the photo, there are several news stories, including "Cover Story: Code Case" and "The Loungesome Death of Mikhail Lezin".

Blendle BETA

Shared

Browse

Popular

Time – Yesterday

TIME

The Brief

Big Money's Tiny Impact on the Presidential Campaign

Michael Scherer

This was supposed to be a great year to buy an election. Thanks to recent court rulings, after all, billionaires are allowed to spend freely to elect their chosen candidate. But so far, money has not seemed to matter. As the numbers reveal, there is little correlation between independent ad spending (which often comes from super PACs and special-interest groups) and primary results, especially for Donald Trump.

SOURCE: CENTER FOR RESPONSIVE

TIME  
MARCH 20, 2017

Cover Story  
**Code Case**

The FBI wants to review the next Linux distro. Say Valley wants to protect your privacy. Apple CEO Tim Cook speaks out about his role in the fight

By Lew Greenman 12

The Loungesome Death of Mikhail Lezin

Vladimir Putin's former protégé died last month dead in a D.C. hotel room in November. The cause? Blunt trauma to the head. TIME reports

By Simon Shuster 50

Putin and Lezin during happier times

Bittersweet Success

Black hole model shows unique challenge

By Karim Abdallah 52

The President and The Ballerina

Barack Obama and Misty Copeland in conversation

By Maya Rhodan 56

On the cover: The Code is here! See our Apple roundup article in our special section on page 50. Photograph by Michael S. Lewis for TIME

© 2017 TIME INC.