

NZZ

# Jump-starting responsible algorithmic content recommendations for news

SDS2023

10th IEEE Swiss Conference on Data Science

**Cristina Kadar, PhD**

Data Science & Machine Learning Product Owner

@tweeting\_cris

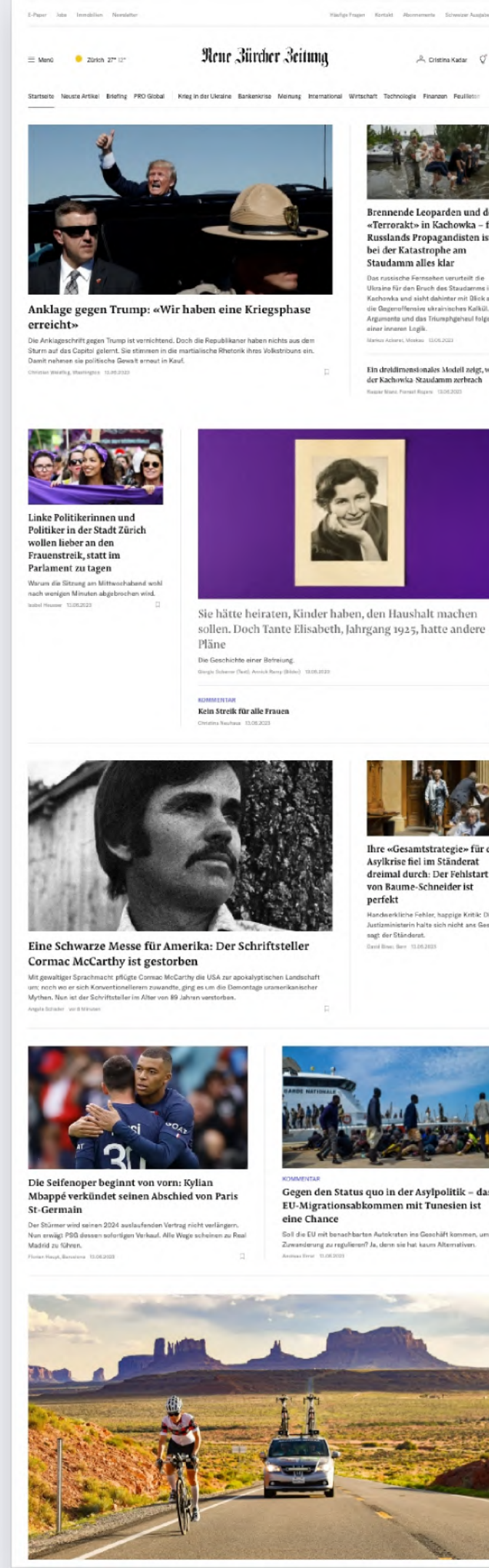
# Motivation

# Why use algorithmic recommendations?

A typical newspaper publishes ~100 stories/day.

For users is easy to miss important articles, due to the high frequency and limited space on our curated distribution channels (e.g. front pages).

💡 At their core, news recommender systems solve an **information filtering** problem.

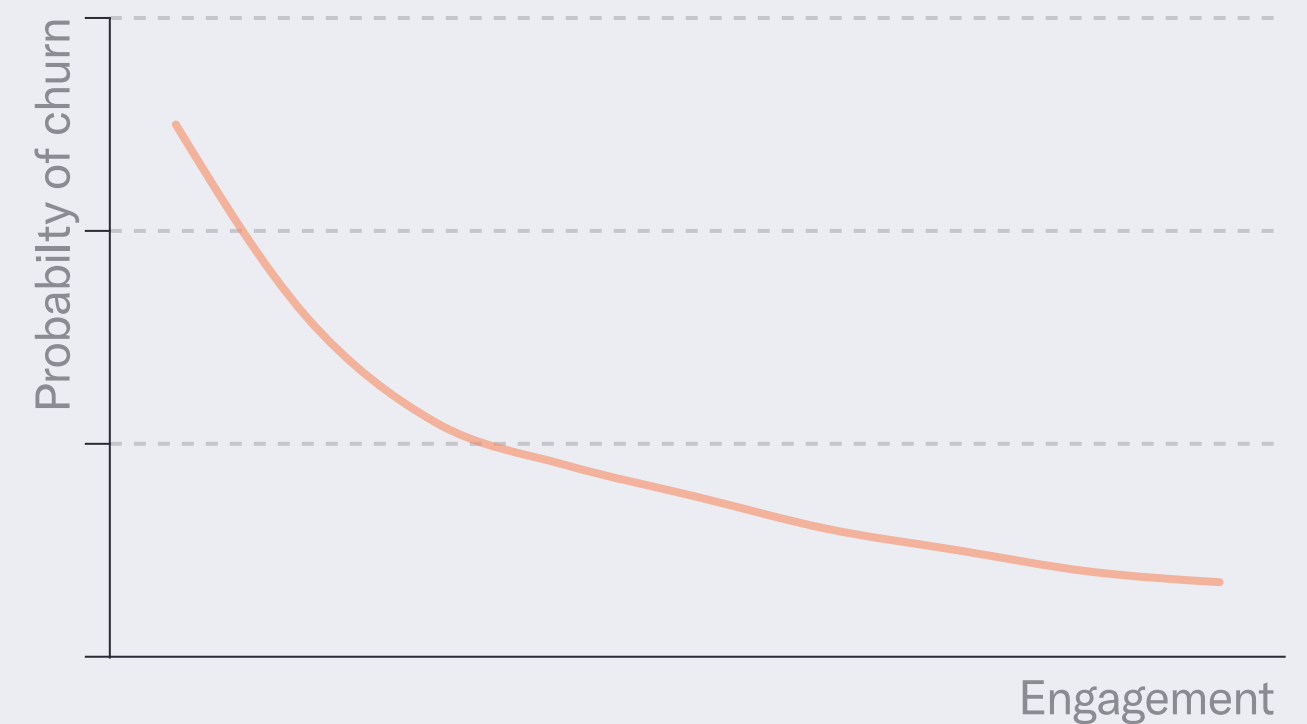
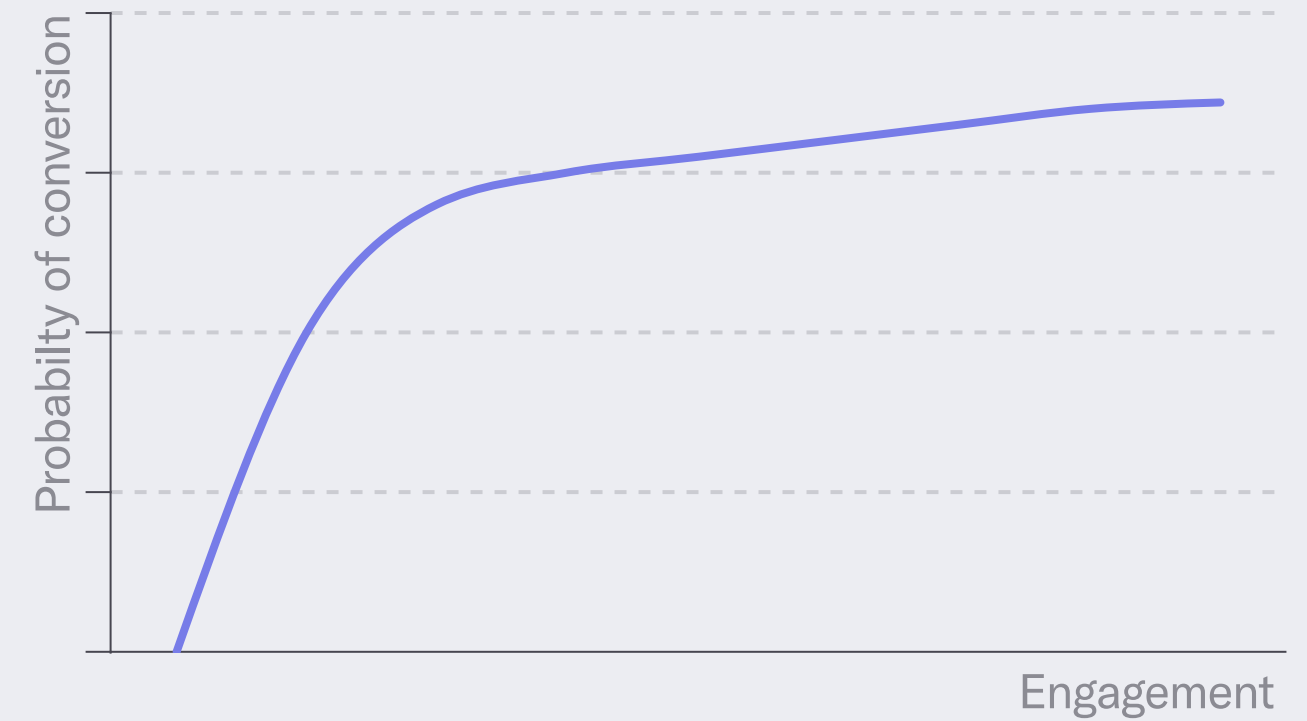


# Goals of algorithmic content recommendations in news

## Increase engagement and subscription / retention

Provide *relevant* recommendations so that users at all stages feel more informed and become more engaged with the product

- Newly registered users exposed to more and diverse content are more likely to subscribe
- Subscribers' engagement is negatively correlated with churn



## Responsibly address user needs and uphold editorial judgement

Design *responsible* recommendations, mindful of the user's voice and the newspaper's curation principles

FOMO



No personalization

Information overload

*What I am missing?*

Too much personalization

Fear of filter bubble

*What I am missing?*

# Success factors of a responsible content personalization strategy

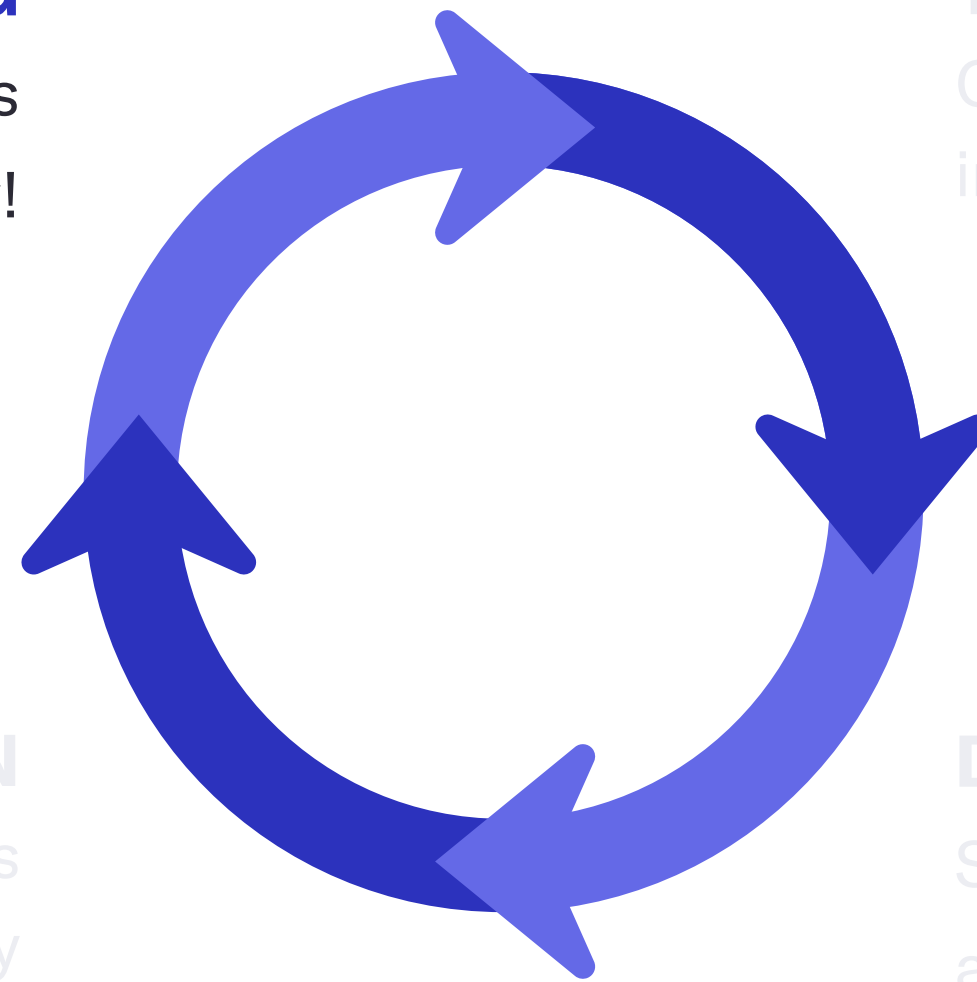
## Interdisciplinary, iterative, and experiment-driven approach

### **USER UNDERSTANDING**

Address real user needs & concerns  
Start smart and economically!

### TRANSPARENCY

Offer transparent value proposition,  
in line with editorial values



### EXPERIMENTATION

Define success metrics  
A/B test consistently

### DATA & ALGORITHMS

Setup tracking, cloud infrastructure,  
and in-house responsible algorithmic dev.



## Type of personalization

### Active personalization

- Uses explicit user feedback
- Requires high up-front costs

### Passive personalization

- Uses implicit signals from past behavior
- For easy gains, start here

## Use case

### Front page

- Already highly optimized
- Potentially, unclear value proposition or even unwanted by users

### Newsletters, article pages, user area

- Deliver value fast to users
- Require less cross-departmental coordination

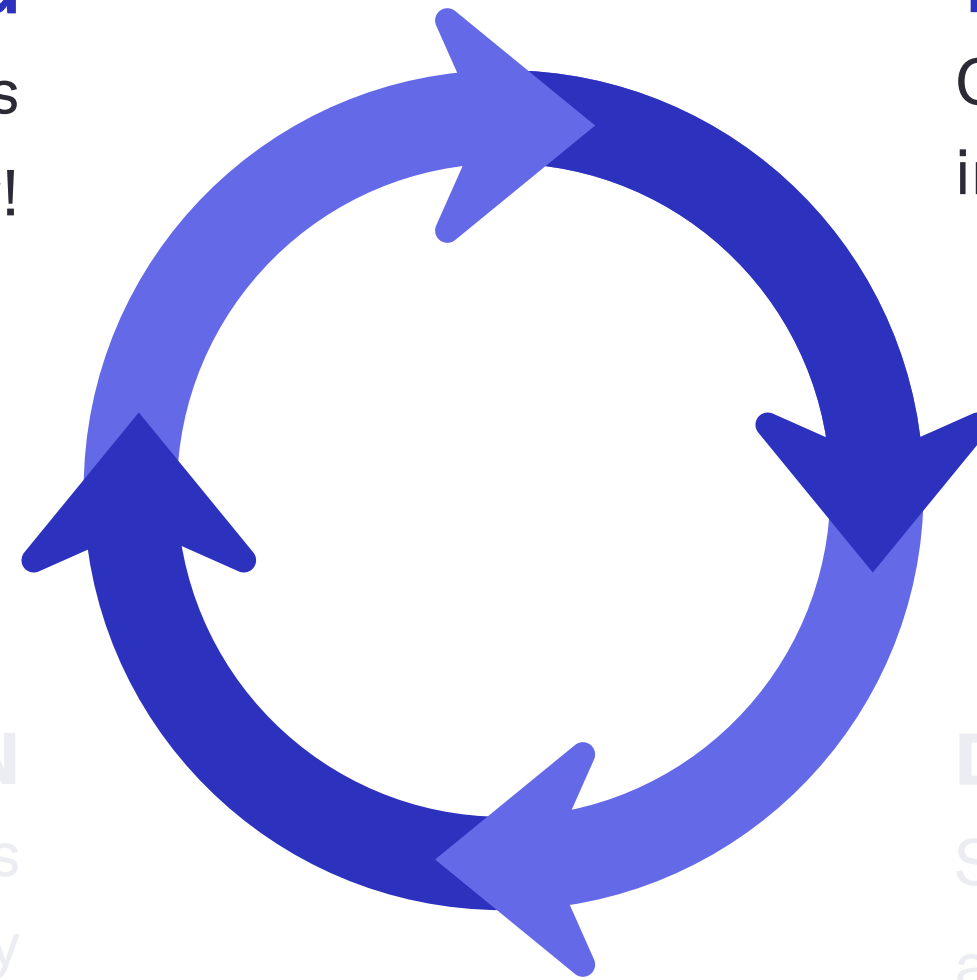
## Interdisciplinary, iterative, and experiment-driven approach

### **USER UNDERSTANDING**

Address real user needs & concerns  
Start smart and economically!

### **TRANSPARENCY**

Offer transparent value proposition,  
in line with editorial values



### **EXPERIMENTATION**

Define success metrics  
A/B test consistently

### **DATA & ALGORITHMS**

Setup tracking, cloud infrastructure,  
and in-house responsible algorithmic dev.

# Transparency

---

## 💡 Highlight personalization

- Clearly mark algorithmically curated areas in the product

## 💡 Explainability

- Design explainable approaches
- Keep users informed and connected

NZZAS.CH

### «Nur für Sie»: Wie unser Empfehlungs-Algorithmus funktioniert

Einerseits wollen wir Ihnen Artikel empfehlen, die zu Ihren Lesegewohnheiten passen, andererseits aber verhindern, dass Sie in einer Filterblase landen.

Marco Metzler und Cristina Kadar · 09.07.2019, 10.32 Uhr

---

 Merken  Drucken  Teilen

Unser Empfehlungs-Algorithmus heisst «Nur für Sie». Wie der Name sagt, erstellt dieser aus mehreren Ausgaben der «NZZ am Sonntag» eine nur für Sie persönlich zusammengestellte Leseliste.

**Nur für Sie >>**

Diese Empfehlungen beruhen auf Ihrem Leseverhalten. [Mehr erfahren](#)

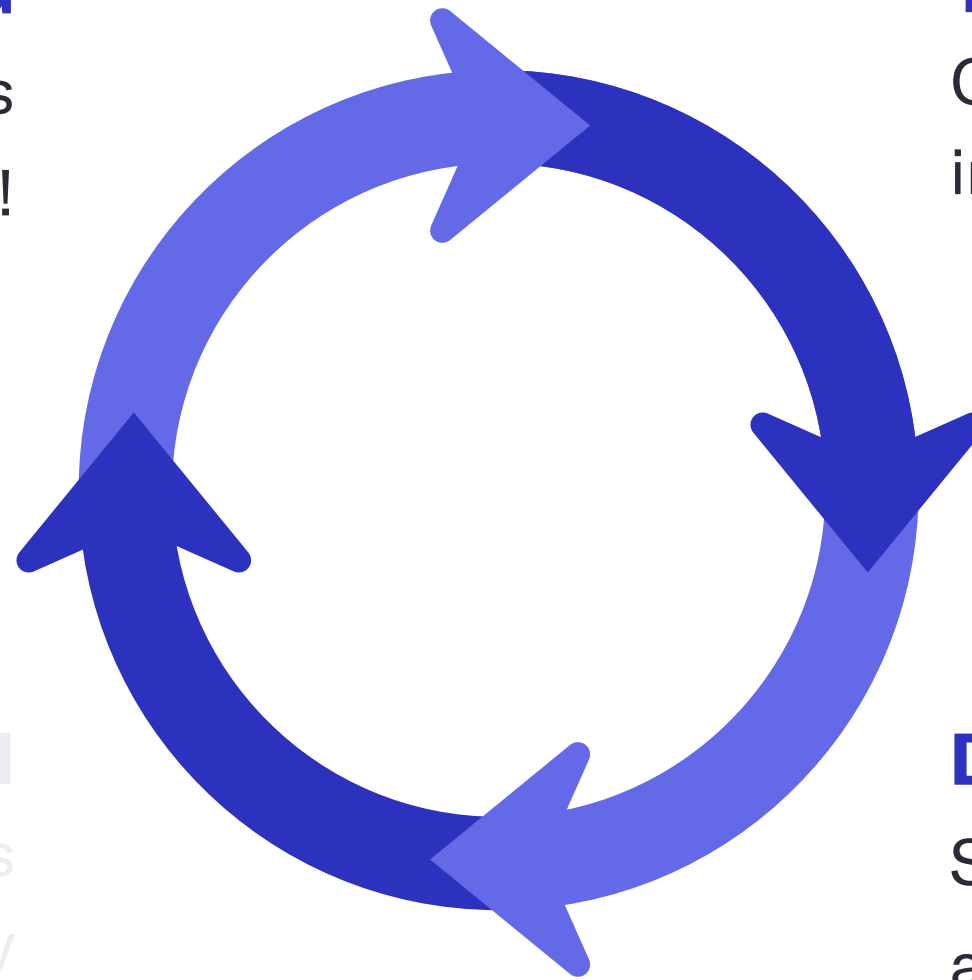
## Interdisciplinary, iterative, and experiment-driven approach

### **USER UNDERSTANDING**

Address real user needs & concerns  
Start smart and economically!

### **TRANSPARENCY**

Offer transparent value proposition,  
in line with editorial values



### **EXPERIMENTATION**

Define success metrics  
A/B test consistently

### **DATA & ALGORITHMS**

Setup tracking, cloud infrastructure,  
and in-house responsible algorithmic dev.

## Data & algorithms

### 💡 1st party data

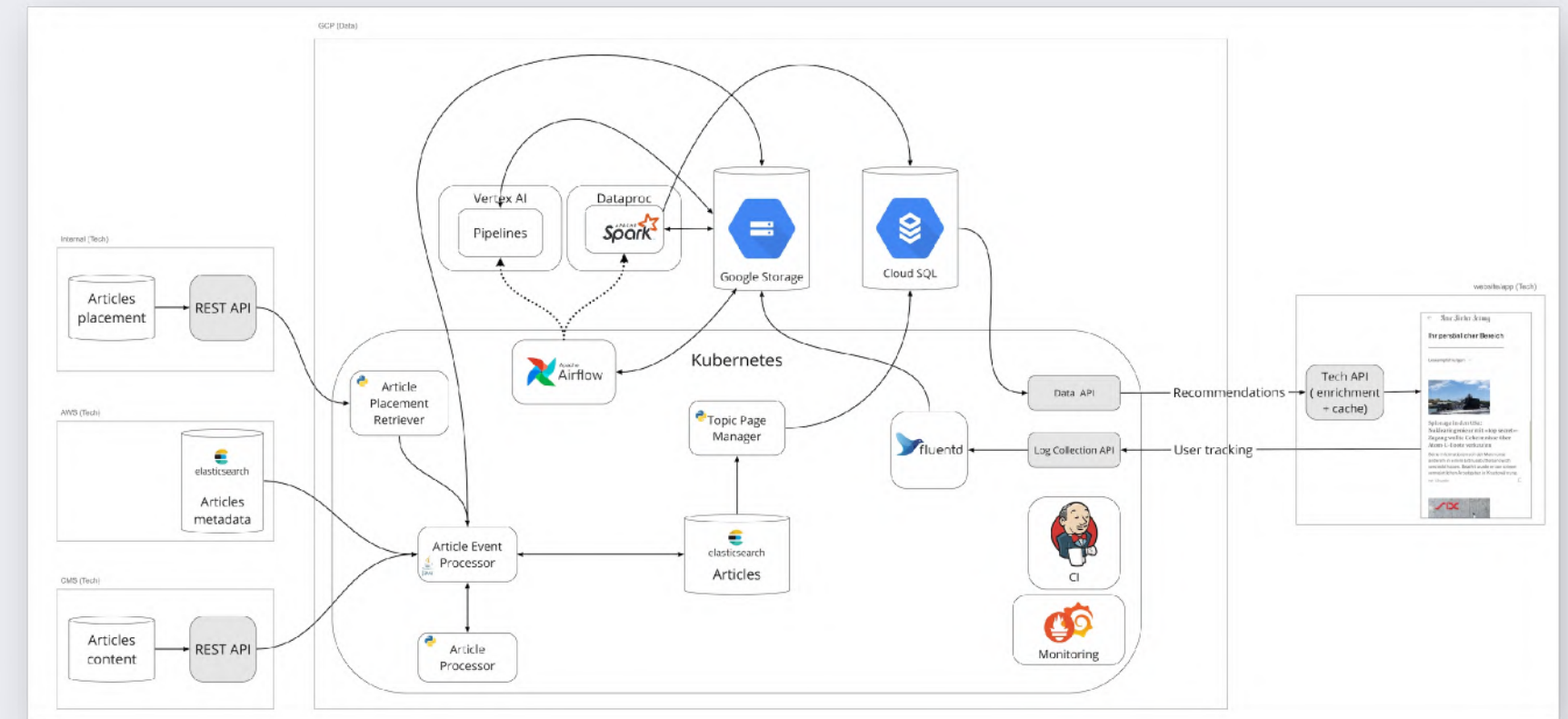
- tracking of rich behavioral data

### 💡 Infrastructure

- cloud storage and processing

### Development

- internal data science team for responsible algorithmic dev.



## Use hybrid personalized recommenders, sorting relevant articles along several explainable scores



### Personal Score

Prefers personally relevant articles

- **collaborative-based filtering:** based on the behavior of users with similar taste
- **content-based filtering:** based on *implicit* user history or *explicit* preferences



### Trending Score

Prefers articles currently popular



### Editorial Score

Prefers generally relevant articles,  
as per our journalists

## Interdisciplinary, iterative, and experiment-driven approach

### **USER UNDERSTANDING**

Address real user needs & concerns  
Start smart and economically!

### **TRANSPARENCY**

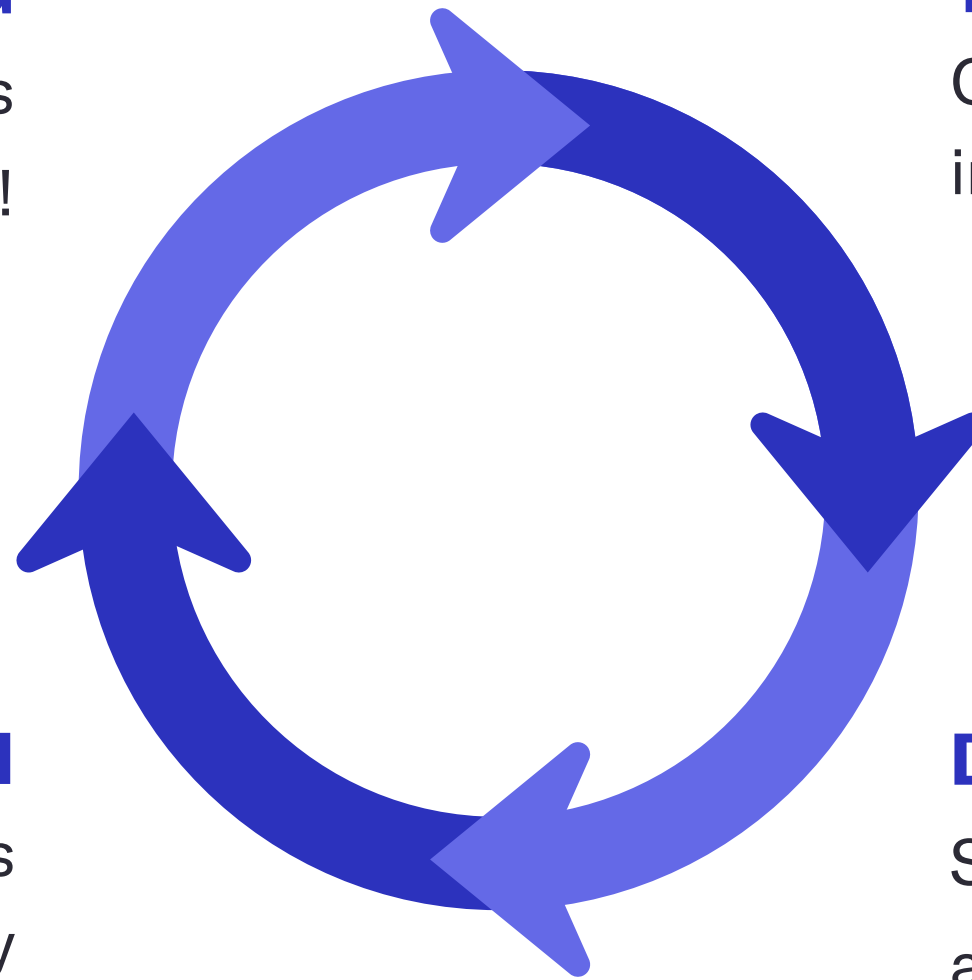
Offer transparent value proposition,  
in line with editorial values

### **EXPERIMENTATION**

Define success metrics  
A/B test consistently

### **DATA & ALGORITHMS**

Setup tracking, cloud infrastructure,  
and in-house responsible algorithmic dev.



# Experimentation

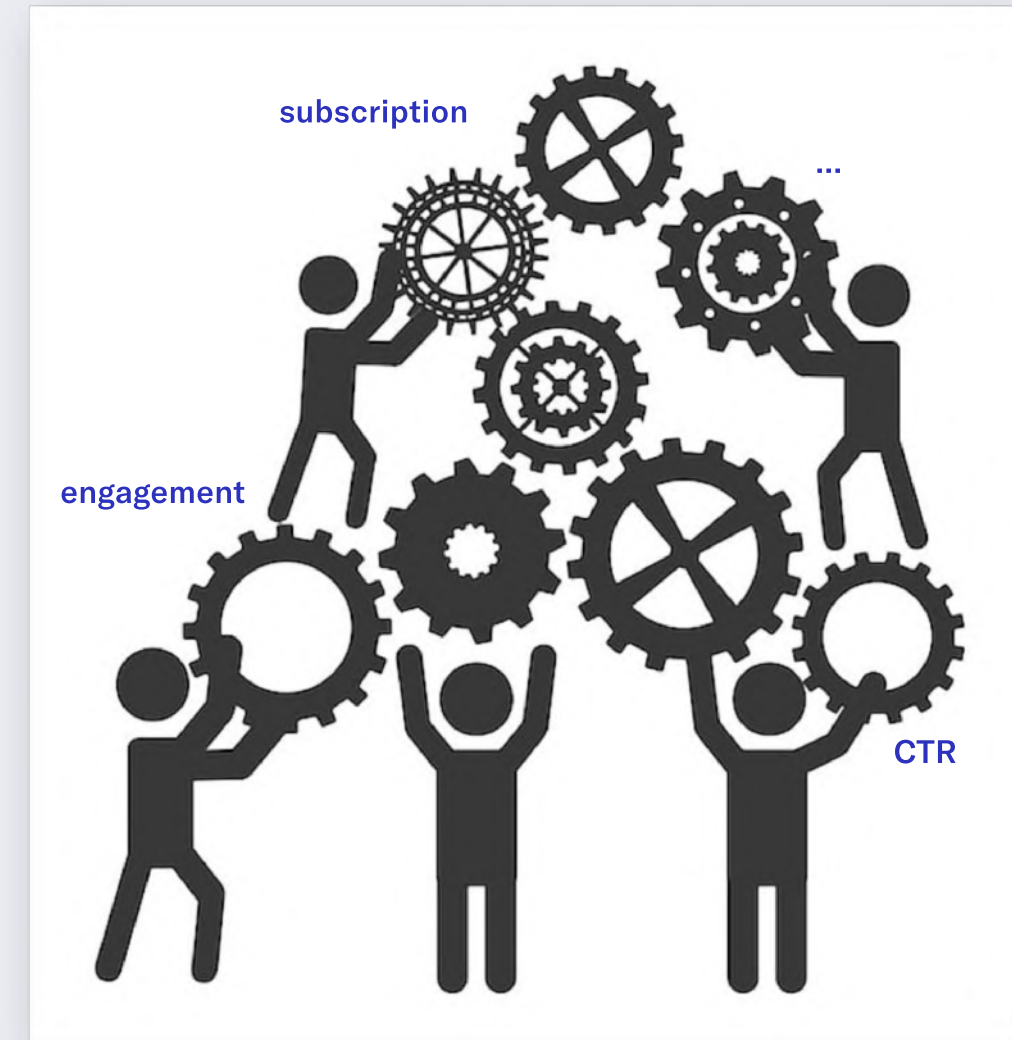
---

## 💡 Metrics

- Business metrics over algorithmic metrics
- Online experimentation over offline experimentation

## 💡 Process

- Invest early in culture and platform
- Feature changes at least as important as algorithmic changes





Use Case:

*Next Reads section on [nzz.ch](https://www.nzz.ch)*

## Fully algorithmically-driven next reads section on the article pages

- many A/B tests over the course of 1 year
- both algorithmic & feature improvements
- two major deployments

**Author Feed:** sources articles from the same author

**Topic Feed:** sources articles from the same topic

**User Feed:** uses a balanced, hybrid personalized approach to rank articles relevant to the user

um die Monetarisierung von Online-Welten und das Organisieren einer durch Nutzerinhalte geprägten digitalen Wirtschaft geht.



## baseline

### Mehr zum Thema



#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Die auf der Blockchain basierende Web3-Technologie soll das World Wide Web laut den Befürwortern wieder dezentraler und demokratischer machen. Wie genau muss man sich das vorstellen?

Philipp Gollmer vor 16 Minuten



#### Facebook setzt auf das Metaversum. Was ist damit gemeint?

Der Konzern von Mark Zuckerberg heisst ab sofort Meta. Er setzt darauf, dass das Metaversum, das vor dreissig Jahren noch Science-Fiction war, bald zum neuen Alltag wird.

Philipp Gollmer vor 16 Minuten



### Lesen Sie auch



#### «Sich selber aufnehmen und fotorealistisch woandershin transportieren – das ist unser Ziel für die nächsten fünf Jahre»

Otmar Hilliges erforscht an der ETH Zürich Technologien für virtuelle Räume. Das grosse Thema für die nächsten Jahre sei Telepräsenz, sagt er – allerdings auf einem ganz anderen Niveau, als das bis jetzt möglich ist.

Helga Rietz vor 16 Minuten



ANZEIGE

#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Jetzt Risiko senken



#### Metaverse, Riesenfernseher, autonomes Fahren: Das sind die Trends an der CES

In Las Vegas startet unter erschwerten Bedingungen die weltgrösste Elektronikmesse CES. Das sind die Highlights des zweiten Pandemiejahres 2022.

Helga Rietz vor 16 Minuten



ANZEIGE

#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Mehr erfahren



#### Nike wächst robust – und will nun den Boom im Metaversum fortsetzen

um die Monetarisierung von Online-Welten und das Organisieren einer durch Nutzerinhalte geprägten digitalen Wirtschaft geht.



## baseline

### Mehr zum Thema



#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Die auf der Blockchain basierende Web3-Technologie soll das World Wide Web laut den Befürwortern wieder dezentraler und demokratischer machen. Wie genau muss man sich das vorstellen?

Philipp Gollmer vor 16 Minuten



#### Facebook setzt auf das Metaversum. Was ist damit gemeint?

Der Konzern von Mark Zuckerberg heisst ab sofort Meta. Er setzt darauf, dass das Metaversum, das vor dreissig Jahren noch Science-Fiction war, bald zum neuen Alltag wird.

Philipp Gollmer vor 16 Minuten

### Lesen Sie auch



#### «Sich selber aufnehmen und fotorealistisch woandershin transportieren – das ist unser Ziel für die nächsten fünf Jahre»

Otmar Hilliges erforscht an der ETH Zürich Technologien für virtuelle Räume. Das grosse Thema für die nächsten Jahre sei Telepräsenz, sagt er – allerdings auf einem ganz anderen Niveau, als das bis jetzt möglich ist.

Helga Rietz vor 16 Minuten



#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Jetzt Risiko senken



#### Metaverse, Riesenfernseher, autonomes Fahren: Das sind die Trends an der CES

In Las Vegas startet unter erschwerten Bedingungen die weltgrösste Elektronikmesse CES. Das sind die Highlights des zweiten Pandemiejahres 2022.

Helga Rietz vor 16 Minuten



#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Mehr erfahren



#### Nike wächst robust – und will nun den Boom im Metaversum fortsetzen

um die Monetarisierung von Online-Welten und das Organisieren einer durch Nutzerinhalte geprägten digitalen Wirtschaft geht.



## v1

### Passend zum Artikel



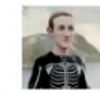
#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

13.01.2022



#### Facebook setzt auf das Metaversum. Was ist damit gemeint?

29.10.2021



#### Metaversum: Diese Firmen konkurrieren mit Meta um eine aufstrebende Milliardenindustrie

02.11.2021

Average CTR uplift 48%

Average completion ratio uplift 61%

### Mehr zum Thema Metaversum

Alle Artikel zum Thema >



#### «Sich selber aufnehmen und fotorealistisch woandershin transportieren – das ist unser Ziel für die nächsten fünf Jahre»

22.01.2022



#### Metaverse, Riesenfernseher, autonomes Fahren: Das sind die Trends an der CES

06.01.2022



#### Nike wächst robust – und will nun den Boom im Metaversum fortsetzen

26.12.2021

Weitere Themen [Facebook](#) [Virtual Reality](#) [NFT](#) [Second Life](#)

### Für Sie empfohlen

Weitere Artikel >



#### Schon wieder verdirbt die Pandemie Millionen von Menschen das chinesische Neujahrsfest

vor 1 Stunden



#### Ein «System der Doppelmoral»: 125 Mitarbeiter der katholischen Kirche outen sich öffentlich – und kritisieren die Diskriminierung durch ihren Arbeitgeber

vor 2 Stunden

um die Monetarisierung von Online-Welten und das Organisieren einer durch Nutzerinhalte geprägten digitalen Wirtschaft geht.



## baseline

### Mehr zum Thema



#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Die auf der Blockchain basierende Web3-Technologie soll das World Wide Web laut den Befürwortern wieder dezentraler und demokratischer machen. Wie genau muss man sich das vorstellen?

Philipp Gollmer vor 16 Minuten



#### Facebook setzt auf das Metaversum. Was ist damit gemeint?

Der Konzern von Mark Zuckerberg heisst ab sofort Meta. Er setzt darauf, dass das Metaversum, das vor dreissig Jahren noch Science-Fiction war, bald zum neuen Alltag wird.

Philipp Gollmer vor 16 Minuten

### Lesen Sie auch



#### «Sich selber aufnehmen und fotorealistisch woandershin transportieren – das ist unser Ziel für die nächsten fünf Jahre»

Otmar Hilliges erforscht an der ETH Zürich Technologien für virtuelle Räume. Das grosse Thema für die nächsten Jahre sei Telepräsenz, sagt er – allerdings auf einem ganz anderen Niveau, als das bis jetzt möglich ist.

Helga Rietz vor 16 Minuten



#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Jetzt Risiko senken



#### Metaverse, Riesenfernseher, autonomes Fahren: Das sind die Trends an der CES

In Las Vegas startet unter erschwerten Bedingungen die weltgrösste Elektronikmesse CES. Das sind die Highlights des zweiten Pandemiejahres 2022.

Helga Rietz vor 16 Minuten



#### Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

Mehr erfahren



#### Nike wächst robust – und will nun den Boom im Metaversum fortsetzen

um die Monetarisierung von Online-Welten und das Organisieren einer durch Nutzerinhalte geprägten digitalen Wirtschaft geht.



## v1

### Passend zum Artikel



#### ERKLÄRT Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

13.01.2022



#### Facebook setzt auf das Metaversum. Was ist damit gemeint?

29.10.2021



#### Metaversum: Diese Firmen konkurrieren mit Meta um eine aufstrebende Milliardenindustrie

02.11.2021

### Mehr zum Thema Metaversum

Alle Artikel zum Thema >



#### «Sich selber aufnehmen und fotorealistisch woandershin transportieren – das ist unser Ziel für die nächsten fünf Jahre»

22.01.2022



#### Metaverse, Riesenfernseher, autonomes Fahren: Das sind die Trends an der CES

06.01.2022



#### Nike wächst robust – und will nun den Boom im Metaversum fortsetzen

26.12.2021

### Weitere Themen

Facebook Virtual Reality NFT Second Life

### Für Sie empfohlen

Weitere Artikel >



#### Schon wieder verdirbt die Pandemie Millionen von Menschen das chinesische Neujahrsfest

vor 1 Stunden



#### Ein «System der Doppelmoral»: 125 Mitarbeiter der katholischen Kirche ouden sich öffentlich – und kritisieren die Diskriminierung durch ihren Arbeitgeber

vor 2 Stunden

Average CTR uplift 48%

Average completion ratio uplift 61%

um die Monetarisierung von Online-Welten und das Organisieren einer durch Nutzerinhalte geprägten digitalen Wirtschaft geht.



## v2

### Passend zum Artikel



#### ERKLÄRT Web3: Die Internet-Revolution, die aus der Abhängigkeit von Big Tech führen soll

13.01.2022



#### Facebook setzt auf das Metaversum. Was ist damit gemeint?

29.10.2021

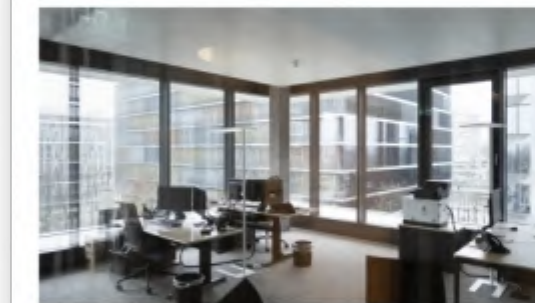


#### Metaversum: Diese Firmen konkurrieren mit Meta um eine aufstrebende Milliardenindustrie

02.11.2021

### Mehr von Philipp Gollmer (phg.)

Weitere Artikel >



#### Hybrid Work: After Corona, not everyone wants to go back to the office. For this to work well, there are a few things to consider

vor 1 Stunden



#### Mit der VR-Brille auf dem Kopf vom Home-Office direkt ins Büro

vor 2 Stunden



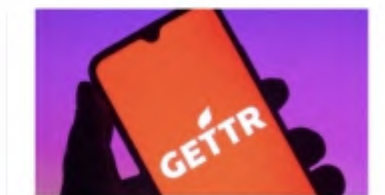
#### Mit Empathie gegen Hassrede: ETH-Experiment zeigt, wie wirksam dieses Vorgehen ist

02.06.2021



#### Wie aus Facebooks internen Unterlagen erst die «Facebook Files» und dann die «Facebook Papers» wurden

08.11.2021



#### Gettr: Was hat der Twitter-Klon zu bieten, und mit welchen Problemen kämpft er?

04.01.2022

### Mehr zum Thema Metaversum

Alle Artikel zum Thema >



#### Metaverse, Riesenfernseher, autonomes Fahren: Das sind die Trends an der CES

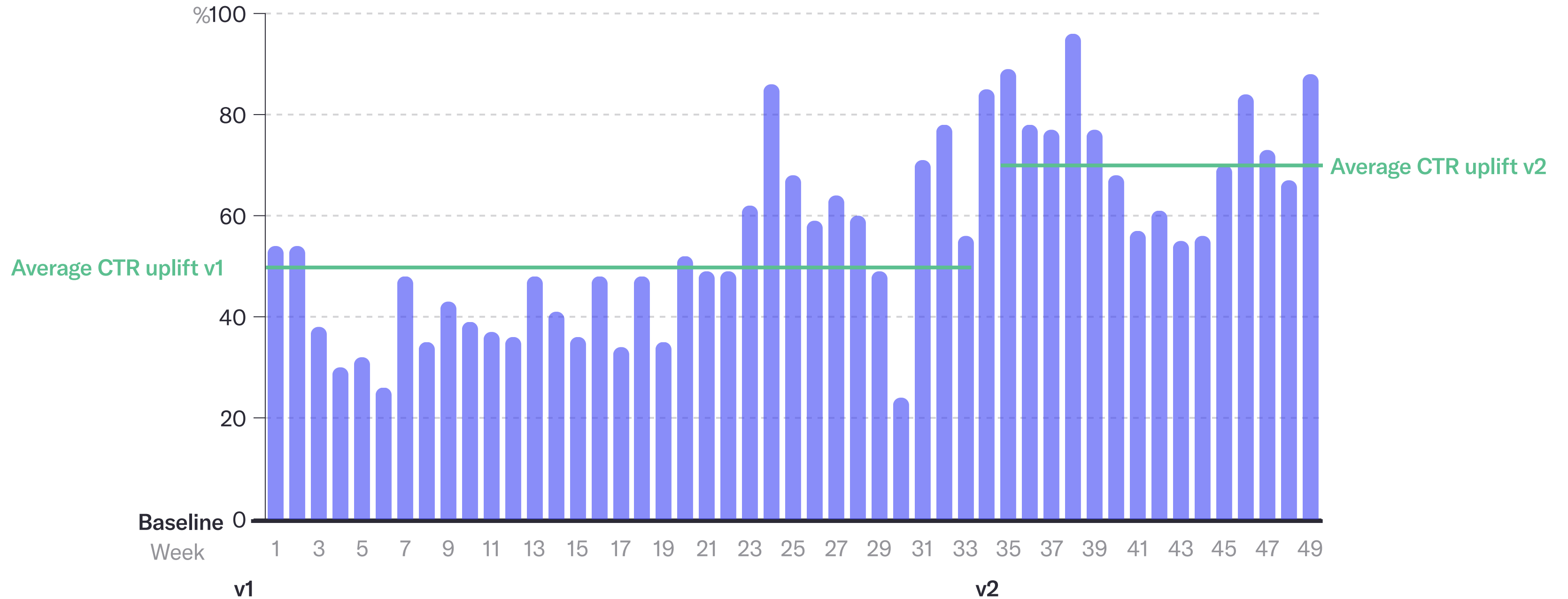
06.01.2022

Average CTR uplift 73%

Average completion ratio uplift 78%

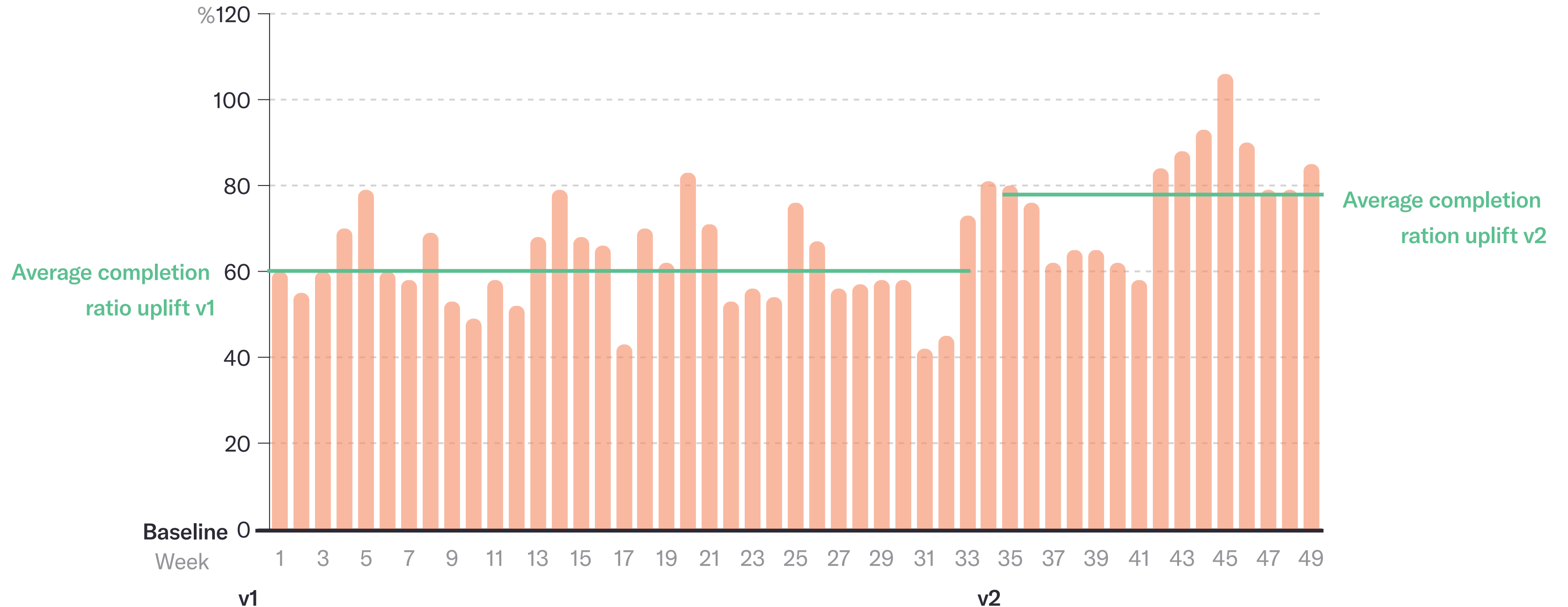
# Weekly CTR on content teasers in the entire section improved on average by **48%** and **73%**, respectively\*

\* The uplifts were computed based on the average weekly local metric after the deployments in comparison to the baseline week set ahead.



# Weekly completion ratio of recommendations in the personalized user feed improved on average by **61%** and **78%**, respectively\*

\* The uplifts were computed based on the average weekly local metric after the deployments in comparison to the baseline week set ahead.



# Further resources



## A history of responsible content personalization at NZZ in four blogs by the data team

- 2021: [Next-reads section on nzz.ch](#)
- 2019: [Nur Für Sie on nzzas.ch](#)
- 2018: [Meine NZZ personalized newsletter from nzz.ch](#)
- 2017: [NZZ Companion App beta product](#)

NZZ

Thank you!  
Questions?

**Cristina Kadar, PhD**

Data Science & Machine Learning Product Owner

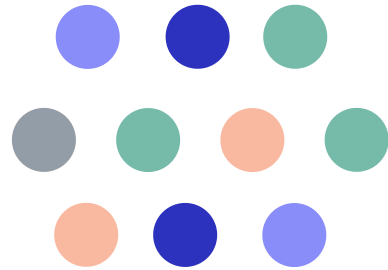
@tweeting\_cris

NZZ

Back-up

# Algorithmic design of personalized recommendations

## All recommender systems consists of these three components



### Candidate Pool

Create pool of eligible articles

e.g. by publication date, site section



### Ranking

Sort articles by relevance

personalized (user-specific) or not  
hybrid (several scores) or not



### Post-processing

Apply filters and business rules

diversity guardrails or not  
time decay or not

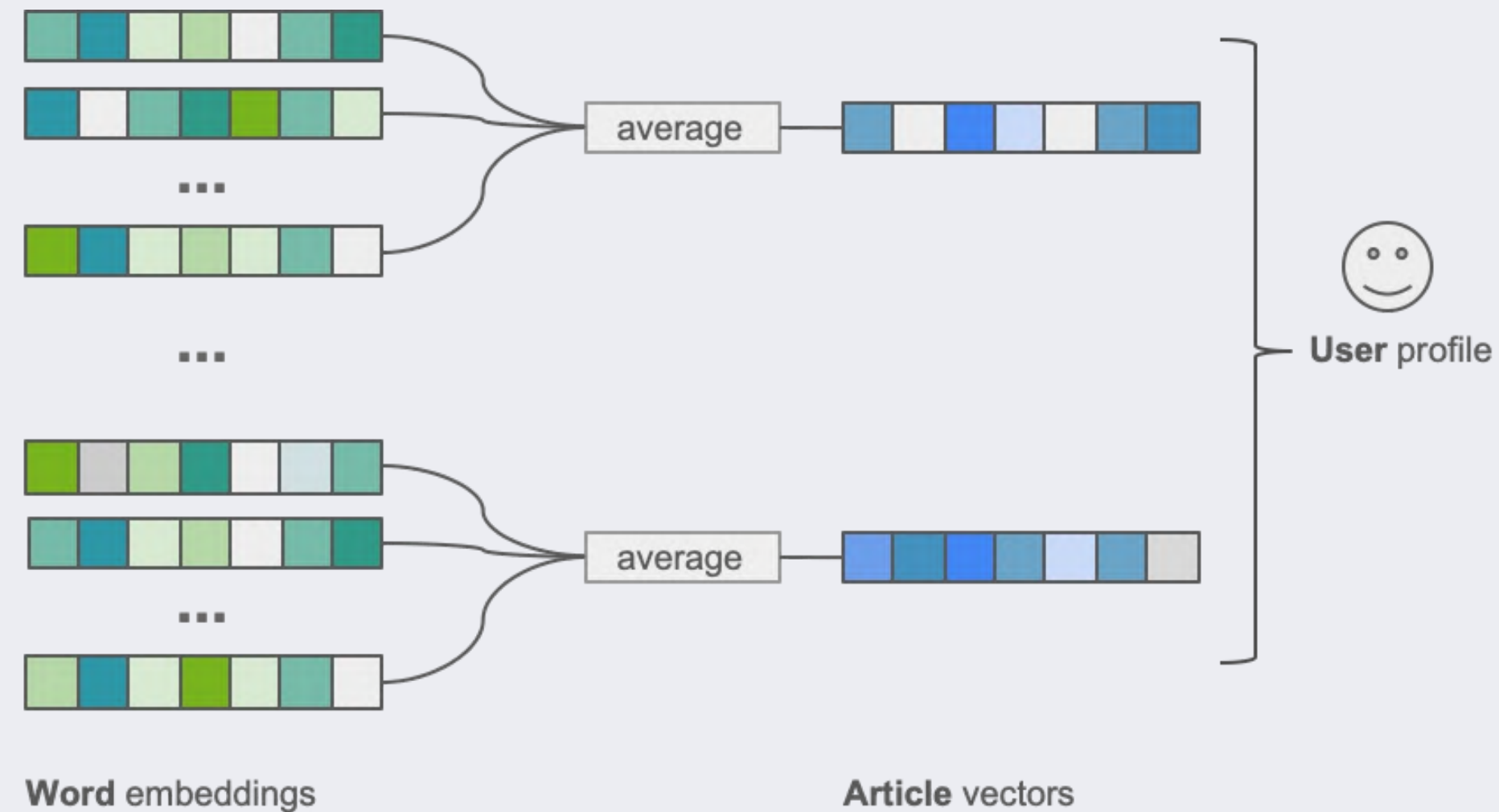
## Content-based approach for the personal score

For each user, we compute daily their **implicit user profile** = a set of 300D vectors representing articles the user has read

💡 FastText word embeddings


For each candidate article, we compute regularly its **distance** to the user profile

💡 Euclidean distance



## Final score used in the personalized feed

The personal, trending, and editorial scores are summarized in a **weighted sum** and further business logic is applied

 Reflects the principles of the newspaper & the characteristics of the use case

$$\begin{aligned} \text{score article, user, time} &= w1 * \text{personal\_score article, user, time} \\ &+ w2 * \text{trending\_score article, time} \\ &+ w3 * \text{editorial\_score article, time} \end{aligned}$$

*apply time decay w.r.t. article's publication date*

*filter out too similar articles*