Architecting for news recommendations: The many hidden colors of the Gray Lady

Katerina Iliakopoulou, Senior Software Engineer, The New York Times
Who am I

- Electrical Engineer turned Journalist turned Software Engineer
- 2.5 years at The New York Times
- Algorithmic Recommendations group
- Architectural Review Board (ARB)
- Machine Learning, Natural Language Processing, Systems Design
- Traveler. Can you tell where the picture was taken?
LEARN ABOUT THE MAIN COMPONENTS OF A PRODUCTION BASED RECOMMENDER PLATFORM FOR NEWS AND THE DESIGN CHALLENGES WHEN BUILDING FOR AVAILABILITY, RESPONSIVENESS AND SCALABILITY.
Overview

- Legacy
- Challenge
- Recommendations Platform
- Architecture Trade Offs
- The Path Forward
Overview

- Legacy
- Challenge
- Recommendations Platform
- Architecture Trade Offs
- The Path Forward
First steps towards a personalized experience
Recommendations Page
nytimes.com/recommendations

Recommended for You
Recommendations are based on what you've viewed recently.

MOVIES
Oscar Nominations: ‘Roma,’ ‘The Favourite’ and ‘Black Panther’ Break Ground
by BROOKS BARNES | Jan 22nd 2019
Netflix got its first picture nod. The 91st Academy Awards are scheduled for Sunday, Feb. 24.

U.S.
Trump Announces Deal to Reopen Government, in Surprise Retreat From Wall
by SHERRY GAY STOLBERG, NICHOLAS FANDOS and PETER BAKER | Jan 22nd 2019
The plan will fund the government until Feb. 15 while Republicans and Democrats negotiate over border security. The plan includes none of the money for the border wall that he had demanded.

U.S.
Fuller Picture Emerges of Viral Video of Native American Man and Catholic Students
by JAHANMERIPOUR and EMILY S. RUBEN | Jan 22nd 2019
A convergence of racial and religious groups sets the stage for the viral moment. “I stepped between to pray,” Nathan Phillips said on Sunday.

MOVIES
Oscar Snubs and Surprises: Bradley Cooper, Netflix and ‘Black Panther’
by KYLE BUCHANAN | Jan 22nd 2019
The director of “A Star Is Born” was left off the ballot and faces stiff competition as
What our stack looked like

- CTPF model (Collaborative Filtering + LDA)
- Python
- PHP → Go
- Luigi framework for scheduling
- Redis for storing User’s History
- AWS
Click Data Consumer

Article Data Consumer

Redis

User History

Rest API

Jobs/Tasks

Edge Server

DynamoDB
Problems with this architecture

- We started building a monolith.
- Tight coupling.
- Hard to build new models.
- Redis was not the right tool.
- Didn’t allow for building recommenders for different areas in the platform.
- It became harder to add business rules that allows the newsroom to control what’s being recommended.
“THE NEW HOMEPAGE ALLOWED TO THINK CREATIVELY ON HOW WE CAN EXPERIMENT WITH DIFFERENT AREAS AND OPTIMIZE HIGHLY CURATED POOLS OF CONTENT FOR OUR READERS.”
Overview

- Legacy
- Challenge
- Recommendations Platform
- Architecture Trade Offs
- The Path Forward
Challenge

- Build/Ship models fast
- Combine any surface of content with any model
- Support many different clients and platforms
Business Logic

Newsroom control over what’s being recommended

- Publication date
- Expiration date
- Remove duplicate content
- Lock articles in certain positions
Business Logic

Newsroom control over what’s being recommended

- Publication date
- Expiration date
- Remove duplicate content
- Lock articles in certain positions
Business Logic

Newsroom control over what’s being recommended

- Publication date
- Expiration date
- Remove duplicate content
- Lock articles in certain positions
Overview

- Legacy
- Challenge
- Recommendations Platform
- Architecture Trade Offs
- The Path Forward

Recommendations Platform
Decisions on the Architecture

● Use microservices.

● Only Go and Python.

● Get rid of Luigi. Start simple with plain cron jobs.

● Build our own CMS to give newsroom control.

● Make User History its own service.

● Isolate the business logic in one single service.
SLOs for the new platform

- All services should be available 99.5% of the time.
- Training cron jobs should run successfully 99.5% of the time.
- Response time for a client request for recommendations should be answered under 100ms for 90% of the requests.
- Response time for retrieving a user’s history should be answered under 30ms for 90% of the requests.
How does our platform look like now
Many more models to try

- Matrix Factorization (ALS)
- Thompson Sampling
- Contextual Bandits
- Different popularity rankings
Let’s see how that works
View of all lists shows main characteristics.

<table>
<thead>
<tr>
<th>Display Name</th>
<th>Name</th>
<th>Daysback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opinion</td>
<td>home-opinion</td>
<td>3</td>
</tr>
<tr>
<td>Personal Tech</td>
<td>personal-tech</td>
<td>30</td>
</tr>
<tr>
<td>Your Money</td>
<td>your-money</td>
<td>30</td>
</tr>
<tr>
<td>newsletter-opinion</td>
<td>newsletter-opinion</td>
<td>30</td>
</tr>
<tr>
<td>Signature Journalism</td>
<td>signature-journalism</td>
<td>30</td>
</tr>
<tr>
<td>Smarter Living</td>
<td>home-living</td>
<td>30</td>
</tr>
<tr>
<td>Newsletter News Analysis</td>
<td>newsletter-news-analysis</td>
<td>30</td>
</tr>
<tr>
<td>Published Assets BQ</td>
<td>published-assets-bq</td>
<td>7</td>
</tr>
<tr>
<td>Watching</td>
<td>watching</td>
<td>7</td>
</tr>
</tbody>
</table>
Editors can expire an asset or lock an asset at a certain position.
The Business Service queries the Admin frequently to refresh the list and asset data in cache.
We do A/B Testing to measure the performance of different models.

EDGE SERVER

Incoming Request for Ranked Content

A/B

Model Allocation for user

Update list and asset data in memory

CMS Admin

Business
Upon request for a user and given a model allocation the business service performs all necessary operations to rank the content for the reader.
<table>
<thead>
<tr>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smarter Living</td>
</tr>
<tr>
<td><strong>5 Cheap(ish) Things to Combat Winter Cabin Fever</strong></td>
</tr>
<tr>
<td>Smarter Living</td>
</tr>
<tr>
<td><strong>To Fight Fatty Liver, Avoid Sugary Foods and Drinks</strong></td>
</tr>
<tr>
<td>Eat</td>
</tr>
<tr>
<td><strong>At Hotels, Space That’s Like the Office, ‘but Cooler’</strong></td>
</tr>
<tr>
<td>Business</td>
</tr>
<tr>
<td><strong>Artificial Sweeteners, Not Good, Not Bad</strong></td>
</tr>
<tr>
<td>Eat</td>
</tr>
<tr>
<td><strong>What to Cook This Week</strong></td>
</tr>
<tr>
<td>Food</td>
</tr>
</tbody>
</table>
ALLOWS FOR EASY EXPERIMENTATION ON DIFFERENT SURFACES USING DIFFERENT MODELS.
Surfaces in 2018

Your Weekly Edition
Discover the best stories you might have missed in this newsletter, curated by Times editors and personalized for you with machine learning.

Best of The Times Signature journalism of the last week

9 Democrats Already Running

All But Certain to Run
 Likely to Run

Nine Democrats Have Entered the 2020 Presidential Race. Here’s Who Is Running So Far.

An Indiana mayor became the latest Democrat to announce his candidacy in what is shaping up to be an already crowded field.

Editors' Picks

How Connected Is Your Community to Everywhere Else in America?

Arthur Mitchell Is Dead at 84; Showed the Way for Black Dancers

Debunking 5 Viral Rumors About Christine Blasey Ford, Kavanaugh’s Accuser
Overview

- Legacy
- Challenge
- Recommendations Platform
- Architecture Trade Offs
- The Path Forward
Architecture Trade Offs

- Availability
- Observability
- Agility
- Throughput
- Extensibility
- Scalability
Architecture Bottlenecks

- **Scaling** is hard and slow because all data is loaded to memory. Got worse after we moved to GKE.

- **Extensibility**: Online models work, but computing recommendations in batch? Not so much.
But overall we perform pretty good...
High System Availability

System Performance

![Graph showing system performance over time with high system availability]
System Performance

Average response time for 90%tile is 75ms.
We saw some important CTR lifts in how readers engaged with our content!

- **Smarter Living** (baseline: editorial) - 38%
- **Editor’s Picks** (baseline: random) - 57%
- **Midterm Elections** (baseline: random) - 82%
Overview

- Legacy
- Challenge
- Recommendations Platform
- Architecture Trade Offs
- The Path Forward
Extend our ML Platform using Google Cloud ML and Dataflow

Your Weekly Edition

Discover
Coverage on topics we think you'll like

How to Make a Millennial Feel Cozy in Just One Beverage

What We Saw at London Fashion Week Men’s
At a time of identity crisis for London, many designers, including Craig Green, and Hussein Chalayan, defiantly flew the flag for British men’s wear.

The Death of the Sick Day
For many office workers, “working from home” has replaced a day spent recovering under the covers.

Google Cloud Composer
We are only at the beginning
Homepage
Consider these recent examples: Heat records were toppled from Norway to Algeria last year. In Australia, a drought has gone on so long that a child in kindergarten will hardly have seen rain in her lifetime. And California saw its most ruinous wildfires ever in 2018, triggering a bankruptcy filing this week by the state’s largest utility, Pacific Gas and Electric.

Is it climate change?

Heat and drought extremes are consistent with scientific consensus: More greenhouse gas emissions in the atmosphere bring a greater likelihood of abnormally high temperatures. Also, broadly speaking, scientists say, a hotter planet makes extreme weather more frequent and more intense.

The real-life numbers bear out the climate models. Concentrations of carbon dioxide in the atmosphere are higher than they have been in 800,000 years, and average global temperatures have risen. The last four years have been the hottest on record, according to the World Meteorological Organization, and the 20 warmest years on record have
Your Weekly Edition
Discover the best stories you might have missed in this newsletter, curated by Times editors and personalized for you with machine learning.

Best of The Times  Signature journalism of the last week

9 Democrats Already Running

All But Certain to Run        Likely to Run

Nine Democrats HaveEntered the 2020 
Presidential Race. Here’s Who Is 
Running So Far.

An Indiana mayor became the latest Democrat to announce his candidacy in 
what is shaping up to be an already crowded field
● Homepage
● Story page
● Email
● Follow
Thank You.

We are hiring!

aikaterini.iliakopoulou@nytimes.com
www.nytimes.com/careers

@matzika00
Rate today’s session