

УДК 004.838

ВИМОГИ ДО ВЕБ-ДОДАТКІВ З ПЕРСОНІФІКОВАНИМ КОНТЕНТОМ

DOI 10.36994/2788-5518-2021-02-02-20

Кисельов Г.Д., к.т.н., доц, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна. g.kyselov@gmail.com

Александрова Т. О., Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна. aleksandrovatata.d.o@gmail.com

Анотація. Розглянуто основні вимоги до веб-додатків з персоніфікованим контентом. Також було розглянуто яким чином будуються персоніфіковані підбірки. Надані переваги і недоліки рекомендаційних алгоритмів, що використовуються в веб-додатках найбільш популярних соціальних мереж.

Ключові слова: персоніфікований, контент, користувач, метадані, підбірка, дані, інформація, додаток.

REQUIREMENTS FOR WEB APPLICATIONS WITH PERSONALIZED CONTENT

Gennadiy Kyselov, Ph.D., Ass. Prof. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. g.kyselov@gmail.com

Tetiana Aleksandrova, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. aleksandrovatata.d.o@gmail.com

Abstract. To write this article, we studied many social applications that use personalized content in their applications. The main subtleties of personalization in each given application were studied, and also all lack, both from the technical point of view and from user experience are considered. At the moment, personalization plays almost a major role in keeping users on the platform and being active users. In 2016, the Youtube team presented a large scientific article on how the neural network of the application works to create a personalized selection of information and what data is used for this. Giving the user some of the "best" recommendations from the list of candidates requires an accurate picture of the user's preferences. The most important indicator for personalization on Instagram is probability. That is: what is the possibility that a person will comment on this post or video; that she will switch to an account from this publication; that she likes this post or video; that this user will want to save this record; what type of content this user prefers. This article presented two examples of websites with personalized content. In preparation, many more systems were analyzed: Spotify, Facebook, Youtube Music, Instagram, Youtube. Based on all the systems that have been analyzed - Youtube provides the most accurate personalized content, in which all subsequent videos are offered based on the user's previous wishes or

the current video. Based on the information received, we proposed our own list of data that will be used to personalize the content in the application: what will be developed in the future. And the own simplified architecture of the application is given as an example. All decisions were made taking into account the positive and negative aspects of existing applications.

Keywords: *personified, content, user, metadata, collection, data, information, application.*

Вступ

На даний момент існує певна кількість популярних веб-сайтів, що мають контент, який персоніфікується під кожну людину, яка відвідує сайт. Таке представлення контенту дуже поширюється, але опираючись на власний досвід використання таких систем, можна зрозуміти, що багато з них працюють не найкращим чином. Наприклад, часто системи враховують старі запити або те, що вони вважають актуальним для людини, хоча цей контент може вже не цікавити користувача. Або ж враховуються не всі метадані для побудови персоніфікованого контенту. В цьому випадку, контент сайту стає для користувача занадто однотипним та може набридати. Це призводить до того, що кількість активних користувачів сайту зменшується.

Ця тема розвивається останніх років 10, особливо в соціальних мережах Youtube (Youtube Music), Instagram, Facebook, Spotify. Ці мережі мають певну власну аудиторію та провели багато досліджень, щоб вдосконалити власні системи. Але кожне рішення має власні недоліки, що можуть пригнічувати відвідувачів.

Як приклад можна привести Youtube Music. Персоніфікований контент з'являється у підбірках пісень, або при запуску «радіо» від певної пісні. Якщо запускати «радіо» зі своєї підбірки, то більшість пісень потрапить у чергу саме з цієї підбірки. Але, як правило, користувач, запускаючи радіо, хоче почути щось схоже, а не теж саме. Врахування цього підходу до побудови контенту забезпечить відповідному веб-сайту велику аудиторію активних користувачів, які завжди будуть зацікавлені у відвідуванні веб-сайту.

Метою дослідження є:

Аналіз існуючих систем, що мають персоніфікований контент для створення власного рішення з технології створення веб-сайтів з персоніфікованим контентом, який буде найбільш актуальним для користувачів.

Визначення основних типів даних для персоніфікації сторінок веб-сайтів.

Аналіз існуючих систем

На сьогоднішній день вже багато веб-сайтів пропонують персоніфікований контент. Це, наприклад, Spotify, Facebook, Youtube Music, Instagram, Youtube. З усіх цих систем, найбільш точно персоніфікований контент надає Youtube, в якому всі наступні відео пропонуються базуючись на попередніх побажаннях користувача або на поточному відео.

У 2016 році команда Youtube представила велику наукову статтю [2] про те, як працює нейронна мережа додатку створення персоналізованої інформаційної підбірки та які дані використовуються для цього. Надання користувачу декількох «найкращих» рекомендацій зі списку кандидатів вимагає точного уявлення уподобань користувача.

На сьогодні валідність підбірок, які створюються стала значно вище завдяки використанню алгоритмом персоналізації не тільки даних про користувача, а і інформацію про доступні відео: рейтинг, кількість переглядів, ключові слова у відео, кількість реакцій на відео, активність користувачів з певним відео, час перегляду відео іншими користувачами, кількість вподобань або дизлайків на відео й, також, пов'язані відео.

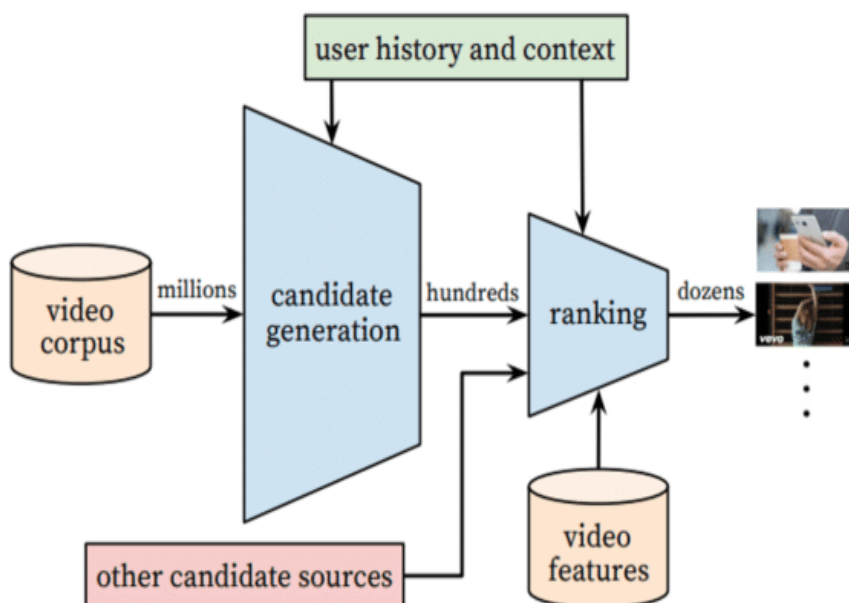


Рис. 1. Архітектура рекомендаційної системи - схема, що була представлена Youtube [2]

Для порівняння, розглянемо іншу дуже популярну соціальну мережу Instagram. Останнім часом розробники Instagram надають більше інформації про роботу своїх рекомендаційних алгоритмів. В Instagram використовується не один алгоритм, а багато. Алгоритми рекомендаційної системи між собою частково пов'язані та співпрацюють. В цих алгоритмах використовується наступна вхідна інформація:

- Наявність у підписах користувача, для якого будується підбірка, особи з подібними інтересами;
- який контент більш цікавий для користувача;

- що користувач дивиться найчастіше;
- як багато користувачів взаємодіють з аккаунтом даного користувача, чи залишають коментарі під його постами та відео.

Найбільш важливий показник – це ймовірність. Тобто: яка можливість того, що людина прокоментує саме цей пост або відео; що вона перейде на аккаунт з цієї публікації; що вона вподобає цей пост або відео; що цей користувач захоче зберегти даний запис; якому типу контенту цей користувач віддає перевагу [3].

Накопичений авторський досвід роботи в мережі Instagram показує, що рекомендаційний алгоритм сторінки «explore» працює задовільно. З'являється враження, що алгоритм не правильно використовує метадані користувача для того, щоб будувати підбірки. В результаті, контент, який користувачу потрібен найменше – алгоритм пропонує найбільше. Тобто намагаючись не взаємодіяти з аккаунтом та відео певної тематики, користувач отримує ще більше контенту такого роду. Навіть якщо поскаржитись, що цей контент не актуальний, результат не зміниться.

У даній статті було представлено два приклади веб-сайтів з персоніфікованим контентом. У підготовці було проаналізовано значно більше систем: Spotify, Facebook, Youtube Music, Instagram, Youtube. Виходячи з усіх систем, що було проаналізовано – Youtube надає найбільш точний персоніфікований контент, в якому всі наступні відео пропонуються базуючись на попередніх побажаннях користувача або на поточному відео.

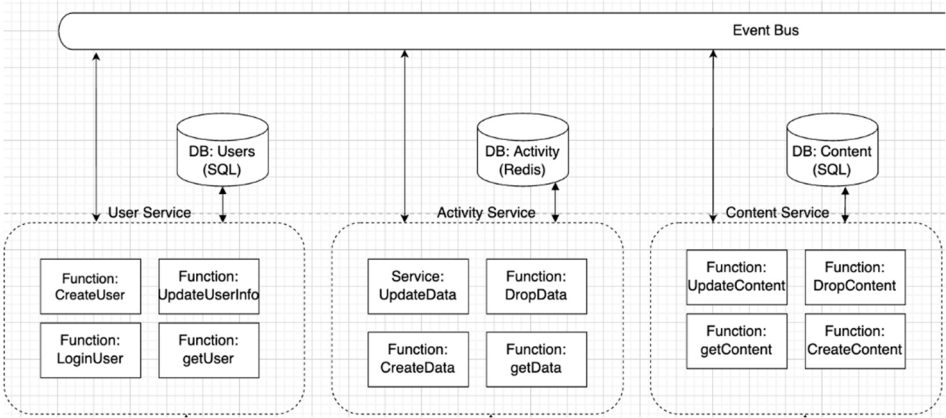
Вимоги до власного веб-додатку з персоніфікованим контентом

Проаналізувавши приклади систем з персоніфікованим контентом та порівнявши їх з боку користувача, можна дійти до висновків, які саме дані потрібно враховувати для того, щоб користувачеві контент був цікавим та не набридав.

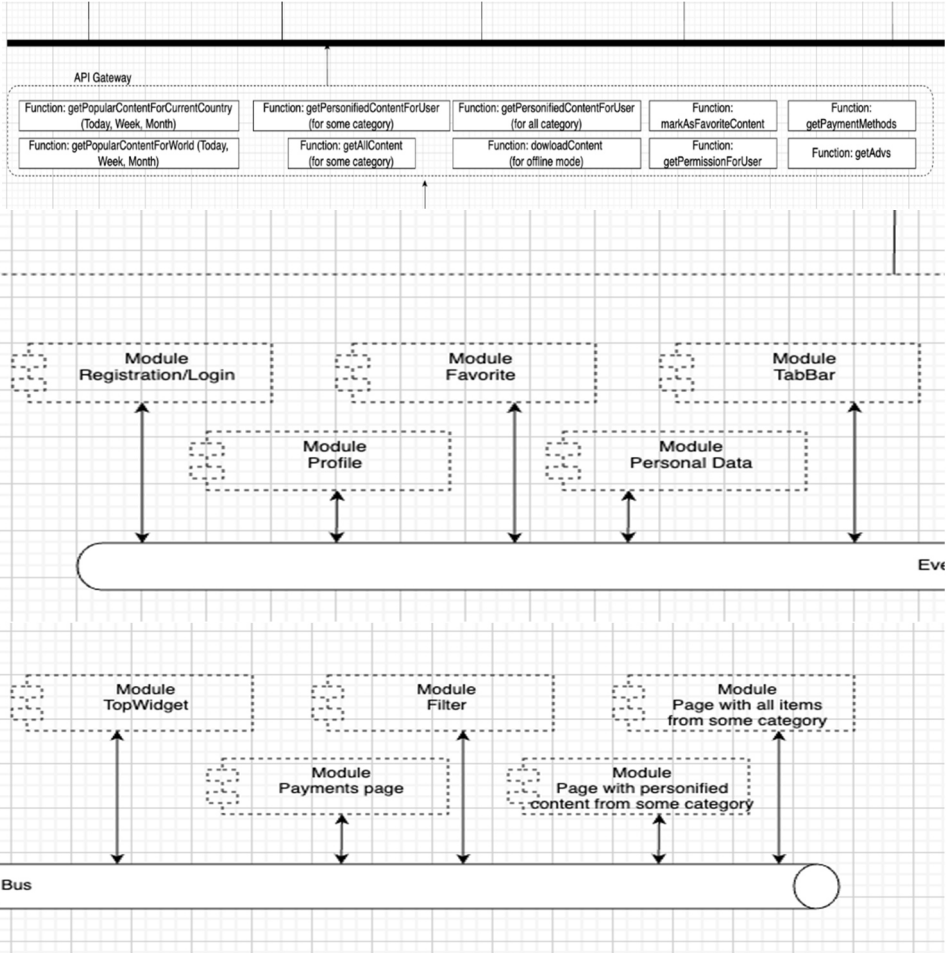
Будь-який проект монетизується для того, щоб він приносив прибуток. У більшості веб-сайтів прибуток отримується завдяки реклами. Персоніфікація реклами дозволяє сформувати контент, який цікавить людину у поточний проміжок часу.

Крім того, необхідно збирати інформацію по кожній категорії контенту: перегляди, взаємодії, збереження, поширення, ключові особливості - будувати рейтинг кожної категорії та групувати схожі за рейтингом категорії, щоб пропонувати їх разом. Рекомендаційна підсистема має накопичувати і використовувати інформацію від користувача:

- які категорії більш за все його цікавлять,
- на які особливості він звертає увагу (це можна реалізувати за допомогою фільтрів),
- скільки разів він переглядає одну одиницю контенту,
- який час він проводить у певній категорії чи у певній одиниці категорії,
- яку інформацію він проглядає і як часто переглядає подібну інформацію.



а)



б)

в)

г)

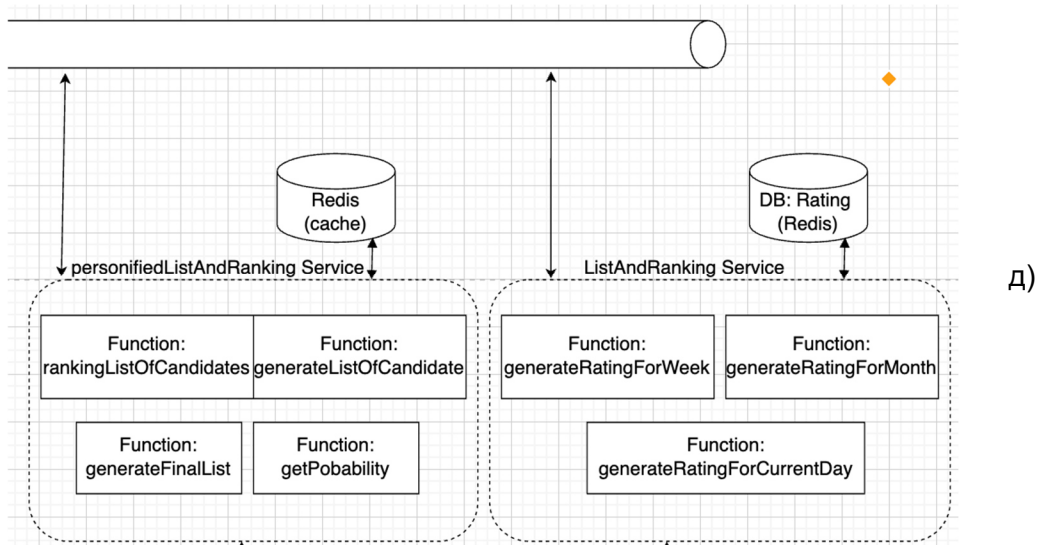


Рис. 2. Архітектура веб-додатку з персонізованим контентом: а) та б) Back-end; в) API Gateway; г) та д) Front-end

Веб-додаток побудований у відповідності до перелічених вимог має бути конкурентоспроможним на ринку сучасних систем з персоніфікованим контентом. Вказаним вимогам відповідає архітектура наведена на рисунку 2.

Front-end - розділений за принципом мікрофронтенду. Кожний логічний елемент (сторінка або блок контенту на сторінці) виділений в окремий модуль, який не залежить напряду від інших модулів. Передача даних відбувається по шині.

Таких принцип обрано саме для того, щоб додаток швидко підтягував контент на певний сторінці, та не підтягував нічого зайвого на тих сторінках, де певна частина інформації не потрібна. Продуктивність додатку – один із важливіших його показників і більшість користувачів не бажають працювати з додатком саме через цю причину.

Back-end, який розділений за принципом DDD (domain-driven design). Приведено головні сервіси (користувацький, контентний, сервіс зі статистичними даними про користувачів та сервіс з підбірками одиниць контенту) та сервіси бази даних, без яких існування додатку неможливе. Сервіси спілкуються за допомогою шини.

Архітектура веб-додатку, яка пропонується, логічно розділена на дві частини.

Висновки

Запропонований підхід до створення архітектури додатку допоможе спростити розробникам або тестувальникам, що будуть співпрацювати з цим продуктом, розуміння доменної області.

Спілкування між клієнтською та серверною частинами відбувається за допомогою API Gateway. Як і в більшості веб-додатків, сервіси повинні мати копії, щоб балансувати навантаження. Тобто, таких моделей, як на схемі, може бути багато. В реальній системі схема взаємодії «живих» сервісів буде виглядати значно складніше.

References

1. How Does the YouTube Algorithm Work in 2021. The Complete Guide. URL: <https://blog.hootsuite.com/how-the-youtube-algorithm-works/> (Date of application: 10.09.2021)
2. Deep Neural Networks for YouTube Recommendation. URL: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45530.pdf> (Date of application: 10.09.2021)
3. This is How the Instagram Algorithm Works in 2021 [Electronic resource] - access mode: <https://later.com/blog/how-instagram-algorithm-works> (Date of application: 10.09.2021)