



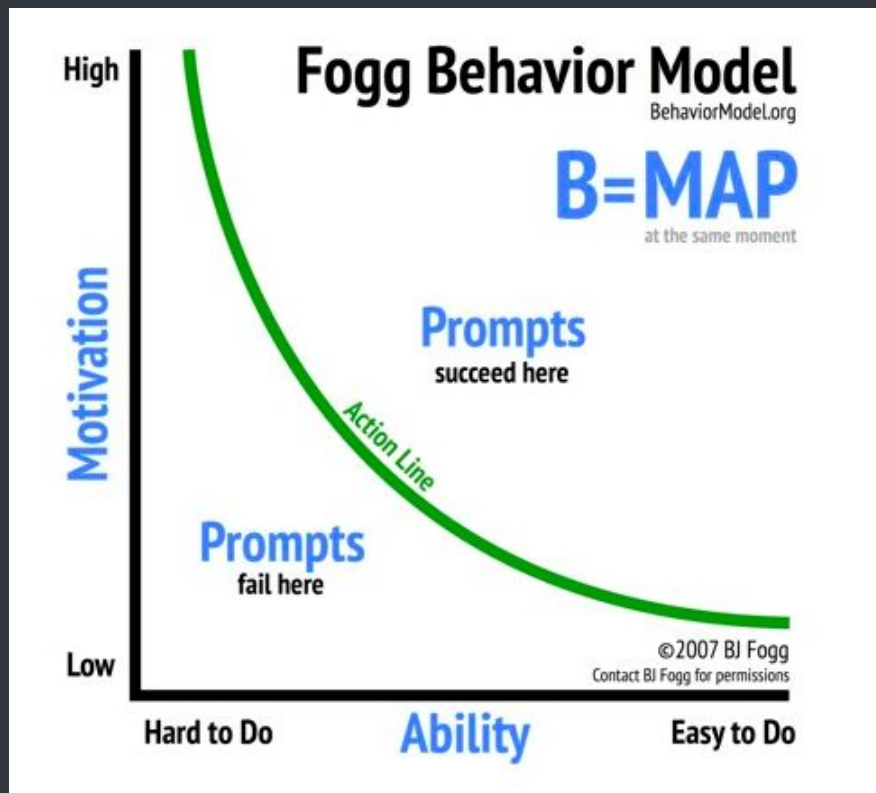
Предсказание и анализ поведения человека в интернете

Содержание



- 1) Поведенческая модель
- 2) Реклама (сегментация аудитории)
- 3) Социальные сети (рекомендации друзей)
- 4) Инфраструктура

3 Поведенческая модель



4 Доступность



5 Мотивация



Sensation is a Core Motivator



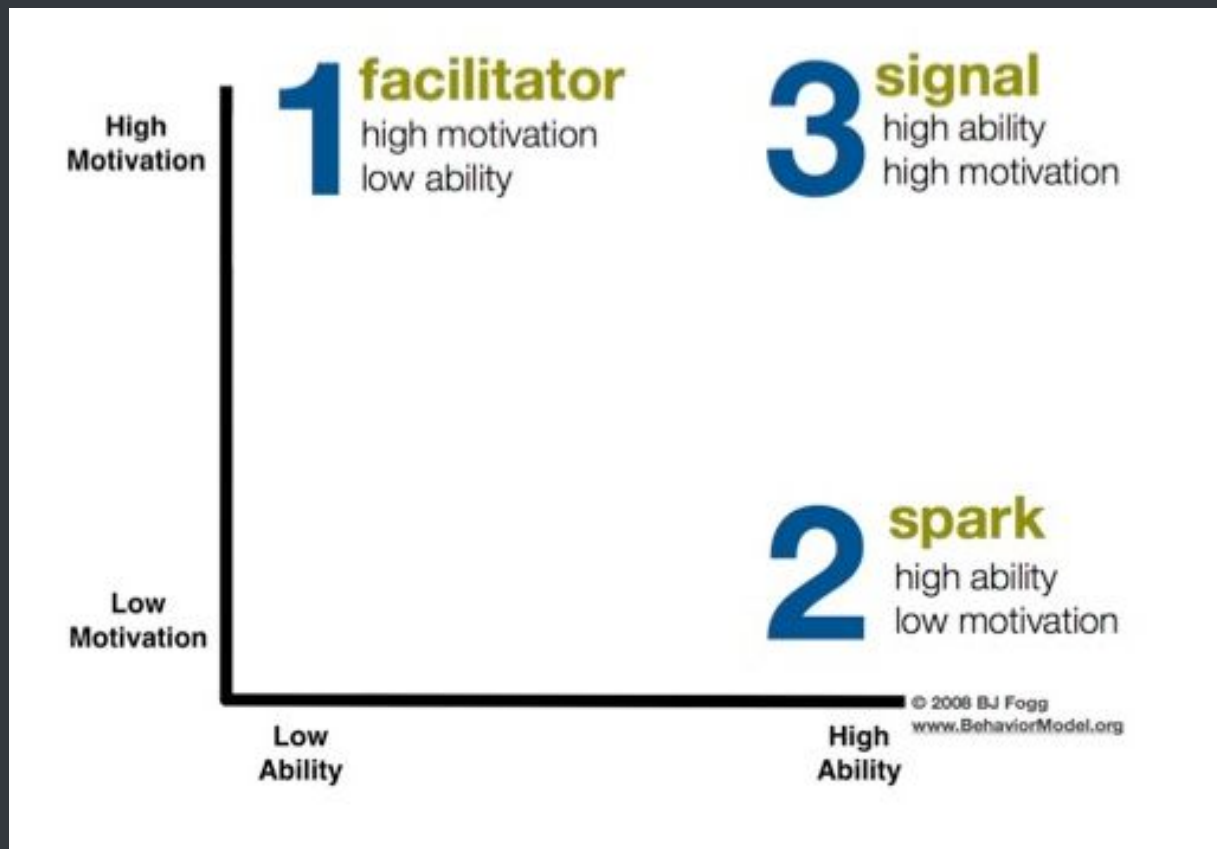
Anticipation is a Core Motivator



Belonging is a Core Motivator





6 Триггеры



7 Facilitator



  Tip of the Day

The **Escape** key in any tool window moves the focus to the editor.

Shift+Escape moves the focus to the editor and also hides the current (or last active) tool window.

The **F12** key moves the focus from the editor to the last focused tool window.

☒ Show tips on startup

[Previous Tip](#) [Next Tip](#) [Close](#)



Тест: Какая прическа подходит именно ВАМ?



Людам до 23 лет - не нажимать! В 2012 году три знака ЗОДИАКА ждет страшная потеря..



Загрузи ФОТО, НАЖМИ ПОИСК и найди знаменитость похожую на тебя!



Как забеременеть, если не получается? Бабушкин метод! ЖМИ!

9 Signal



СЪЕШЬ
ЧТО-НИБУДЬ



ДЕДЛАЙН
ЗАВТРА



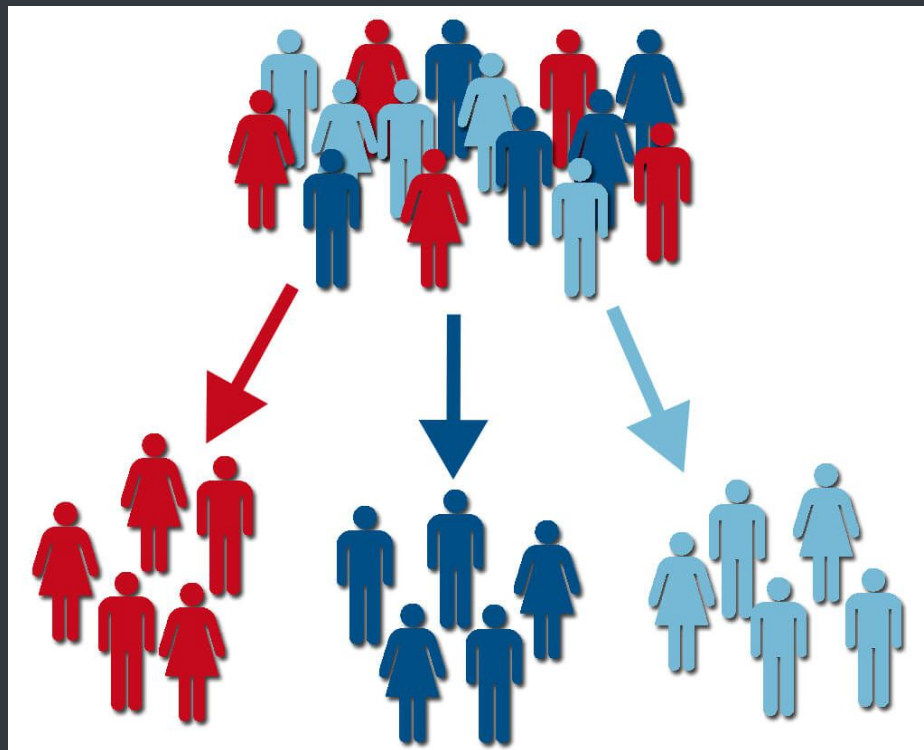
ХОРОШИЕ
ДНИ

¹⁰ Реклама



- 1) SOCIAL NETWORK
- 2) RTB

11 Сегментация аудитории

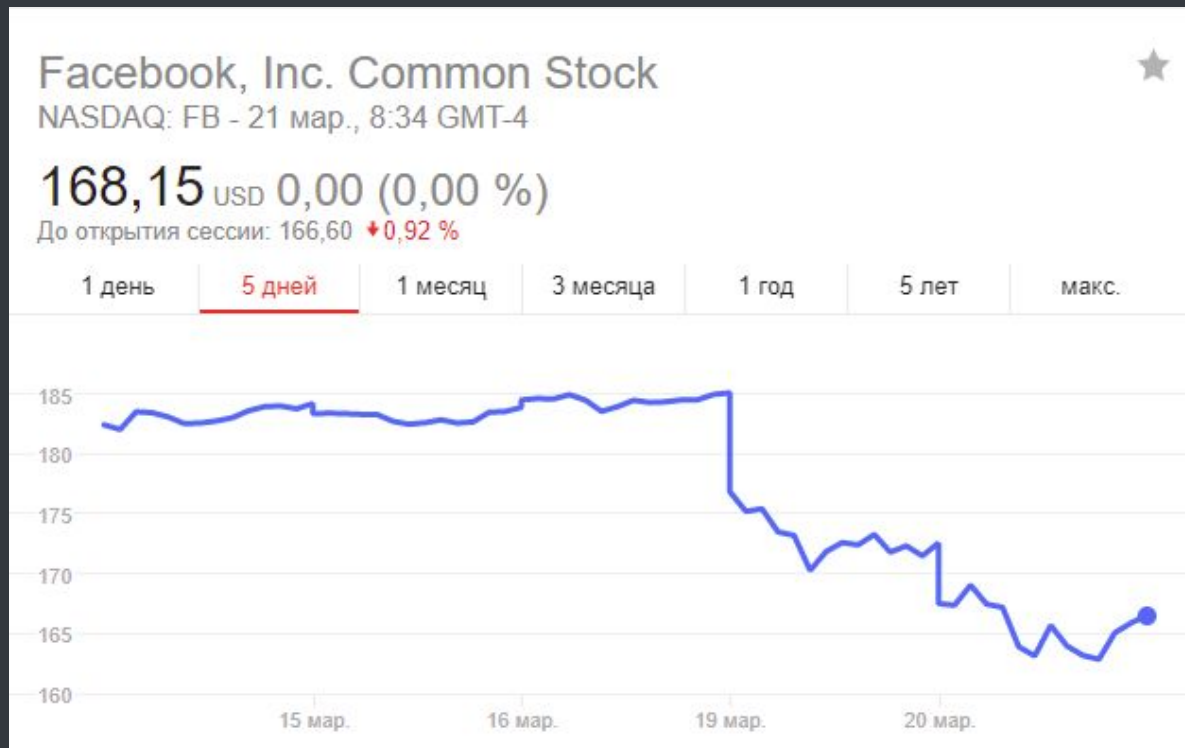


12 Критерии сегментации



- 1) Возраст
- 2) Пол
- 3) Уровень образования
- 4) Социальный статус
- 5) Локация
- 6) Интересы
- 7) Паттерны поведения
- 8) Психологические характеристики
- 9) Политические взгляды
- 10) Религия
- 11) Отношение к алкоголю,
наркотикам и сигаретам
- 12) Беременность
- 13) Сексуальные предпочтения
- 14)
- 15)

13 Предсказание психологических характеристик



14 Большая пятерка



15 Черты личности и предпочтения веб сайтов



Table 2. Websites and website categories with highest and lowest mean personality levels, estimated on the Likes dataset.

Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Liberal & Artistic	Well Organized	Outgoing & Active	Cooperative	Emotional
Arts.Animation	Reference.Education	Computers.Internet	Reference.Education	Recreation.Pets
Business.Marketing	Shopping.Electronics	Reference.Education	Computers.Internet	Recreation.Scouting
Business.Services	Shopping.Children	Science.Environment	Business.Logistics	Science.Physics
Arts.Photography	Reference.Dictionaries	Arts.Music	Health.Diseases	Sports.Hockey
modcloth.com	lww.com	clubzone.com	abebooks.com	cineplex.com
senate.gov	ecollege.com	ideeli.com	socialsecurity.gov	comparedby.us
boingboing.net	ecnext.com	thanksmucho.com	myrecipes.com	myprofilepimp.com
astrology-online.com	exct.net	discoveryeducation.com	bluemountain.com	barbie.com
gutenberg.org	education.com	list-manage.com	serialssolutions.com	yellowpages.ca
cafeastrology.com	kodak.com	trails.com	ecollege.com	biglots.com
...
gateway.com	candystand.com	lyricsty.com	localtribune.org	ncsu.edu
newegg.com	crunchyroll.com	fanfiction.net	funnyjunk.com	sheetmusicplus.com
fitnessmagazine.com	allthetests.com	behindthename.com	sciencebuddies.org	pitt.edu
ourtoolbar.com	bestuff.com	newworldencyclopedia.org	allthetests.com	highschoolsports.net
nhl.com	lyricsdepot.com	personalitypage.com	marvel.com	myrecipes.com
pier1.com	letmewatchthis.com	gaiaonline.com	supercheats.com	lww.com
Reference.Education	Health.Mental Health	Arts.Movies	Kids&Teens.Society	Arts.Photography
Arts.Television	Arts.Music	Shopping.Children	Health.Mental Health	Science.Maths
Sports.Soccer	Arts.Animation	Arts.Literature	Science.Physics	Business.Marketing
Shopping.Children	Arts.Literature	Arts.Comics	Recreation.Pets	Business.Logistics
Conservative	Spontaneous	Shy & Reserved	Competitive	Calm & Relaxed

16 Черты личности и поведение в социальной сети

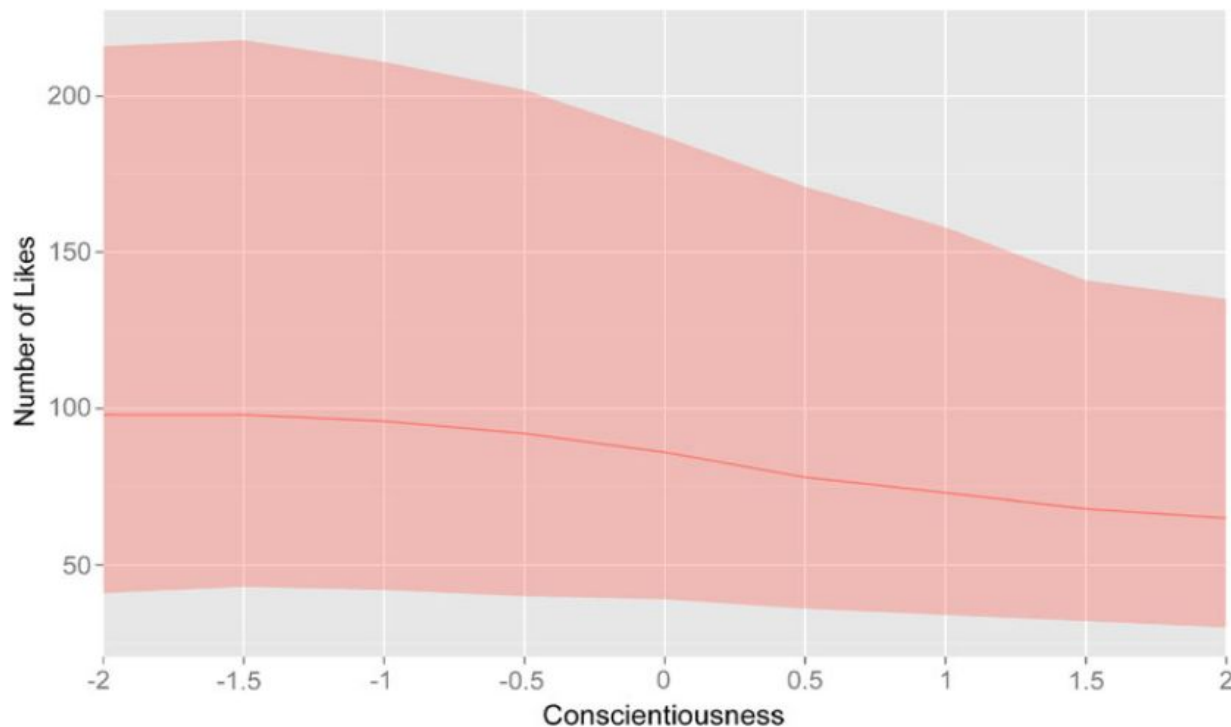


Fig. 3 Median number of Likes for users characterized by different levels of Conscientiousness. Ribbon represents the interquartile range, or the middle 50 percentiles of the number of users' Likes

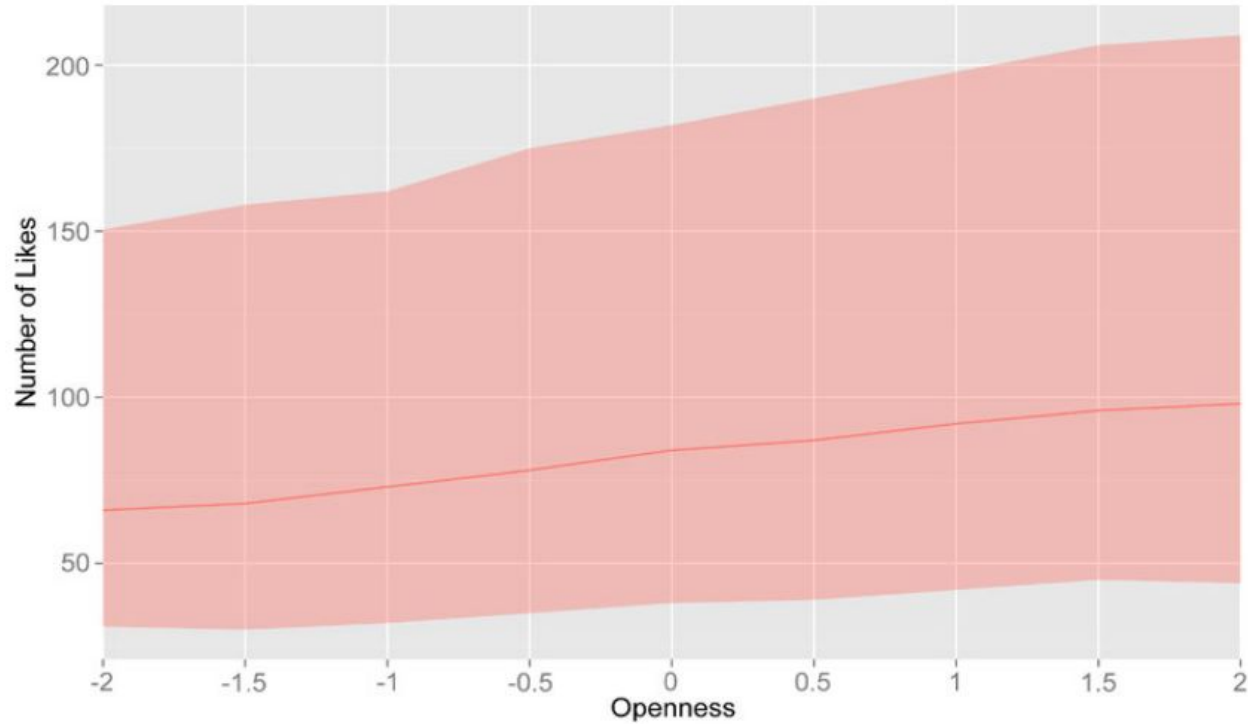


Fig. 2 Median number of Likes for users characterized by different levels of Openness. Ribbon represents the interquartile range, or the middle 50 percentiles of the number of users' Likes

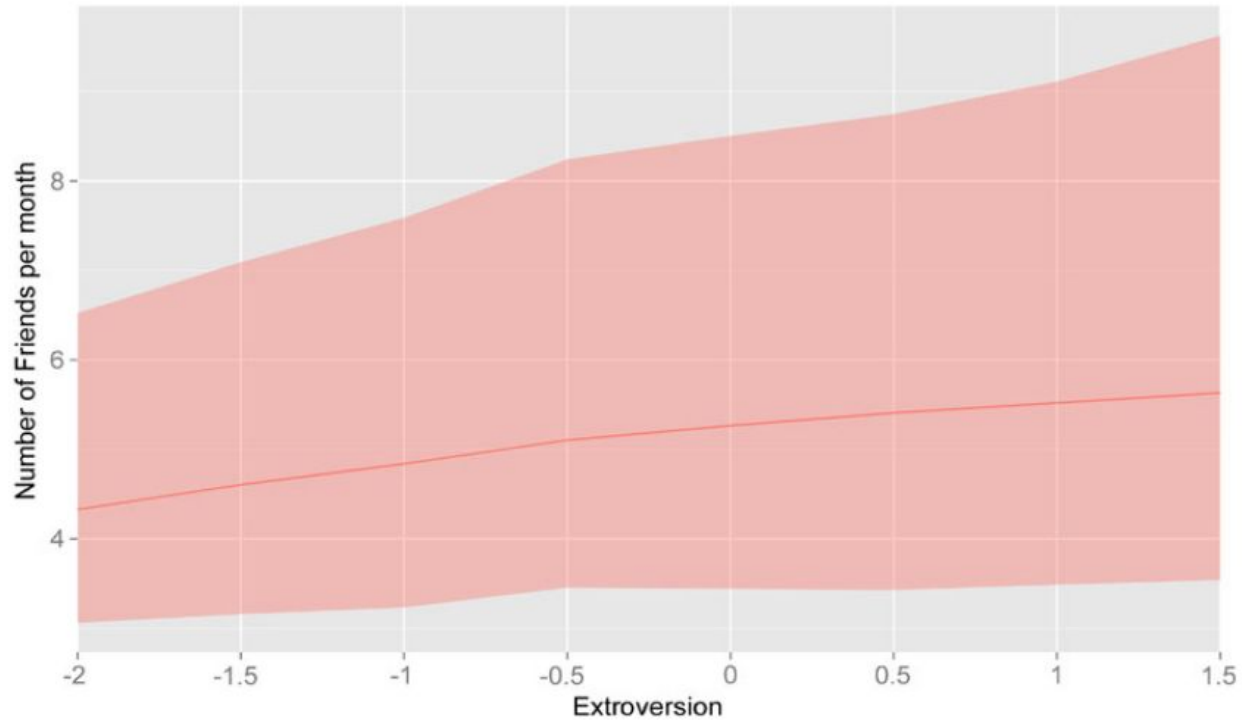


Fig. 6 Median of friends added per month by users characterized by different levels of Extroversion. Ribbon represents the interquartile range, or the middle 50 percentiles of the number of friends added per month

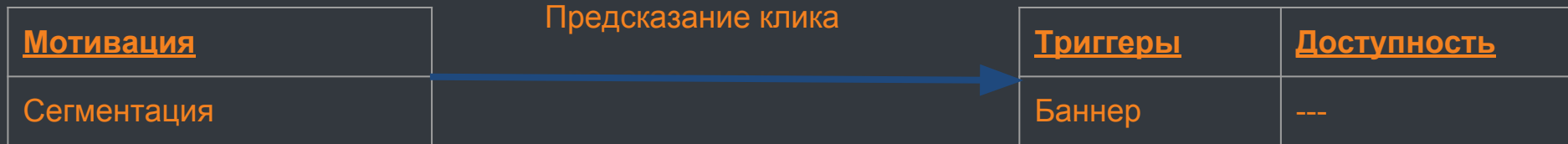
19 Результаты предсказания



Table 9 Predicting personality, Satisfaction with Life, Intelligence, and age based on multiple profile features using a multivariate linear regression with 10-fold cross validation. Table presents prediction accuracy expressed by the Pearson correlation coefficient, sample size, and Facebook features used in the prediction

Trait	Accuracy (r)	n	Features used in the prediction
Openness	.11	18,720	Friends, Groups, Likes, Network Density, Photo Tags, Photos
Conscientiousness	.16	18,720	Friends, Groups, Likes, Network Density, Photo Tags, Photos
Extroversion	.31	16,900	Friends, Groups, Likes, Network Density, Photo Tags, Statuses
Agreeableness	.05	45,565	Friends, Likes, Photo Tags
Neuroticism	.23	9,515	Friends, Likes, Photos, Statuses
Satisfaction with Life	.33	311	Events, Likes, Network Density, Photo Tags
Intelligence	.20	395	Events, Likes, Photo Tags, Photos, Statuses
Age	.50	3,826	Events, Friends, Groups, Likes, Network Density, Photo Tags, Photos

20 В контексте поведенческой модели



21 Социальные сети



22 Рекомендации друзей



Возможно, вы знаете этих людей

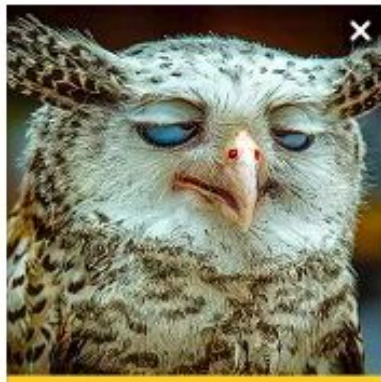
[Посмотреть всех](#)



ЛжеГраф 🐱 Генр 🐱 Газе...
2 общих друга



🐱 Javier 🐱 Hernandez
2 общих друга

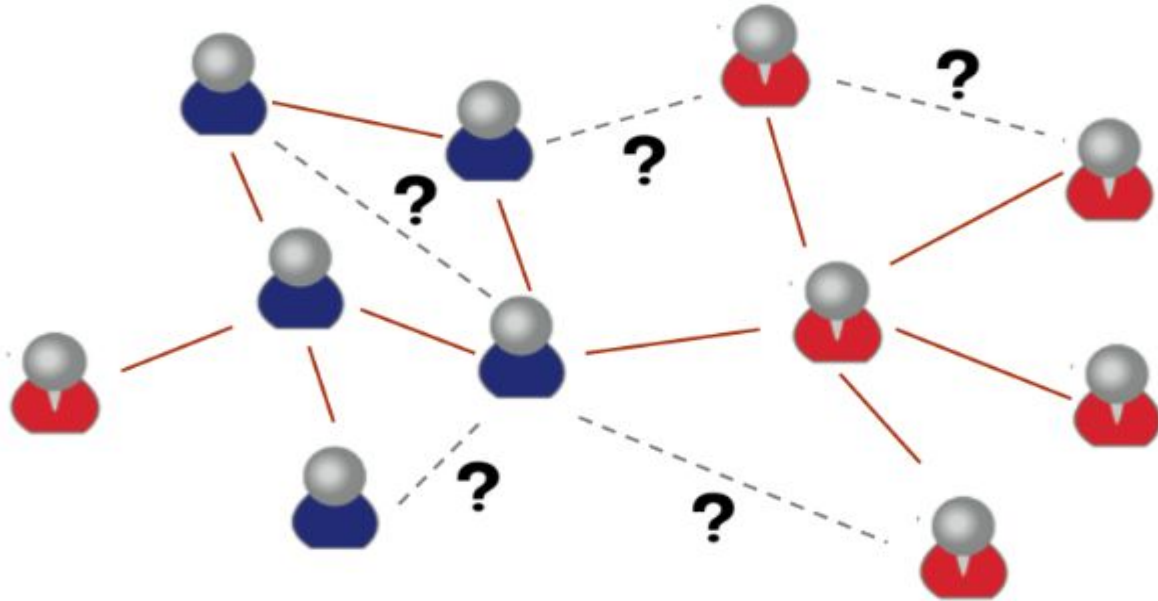


Mari Rouge
2 общих друга

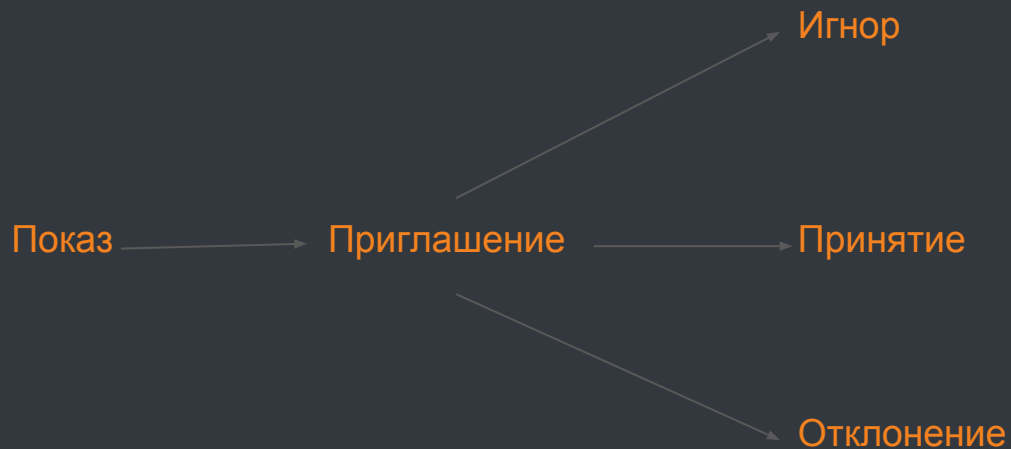


Анюта 🐱
3 общих

23 Link prediction



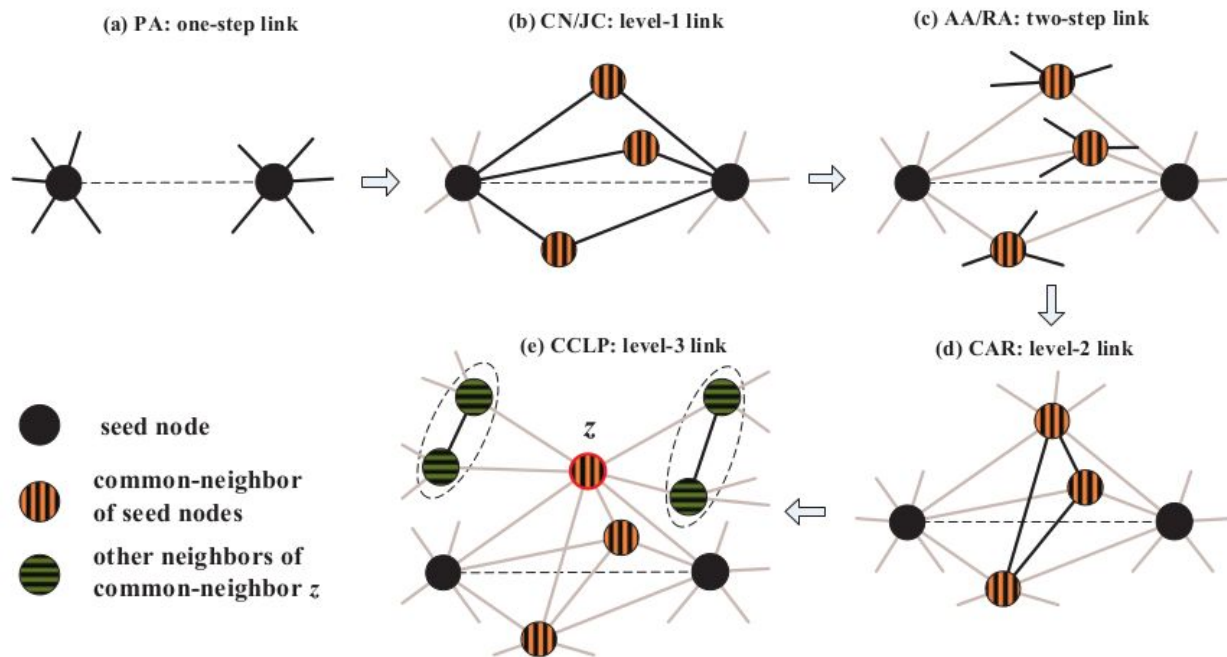
24 Наш выбор - классификация



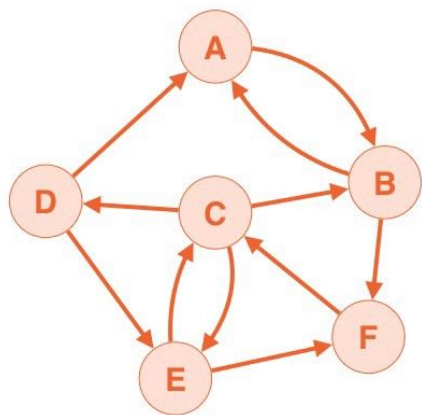


- профильные
- вариации на тему друзей и общих друзей
- комбинированные на основе фич текущего профиля и тех кого мы рекомендуем

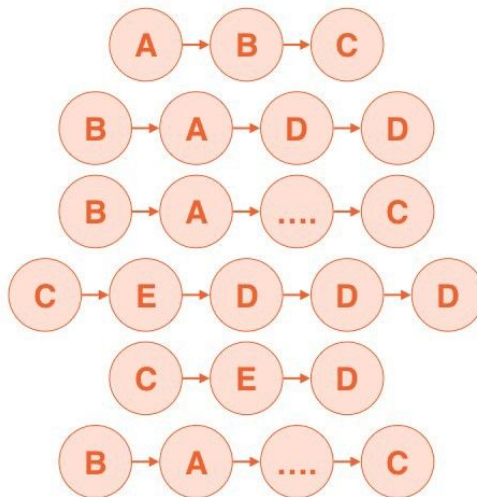
26 Иллюстрация графовых метрик схожести



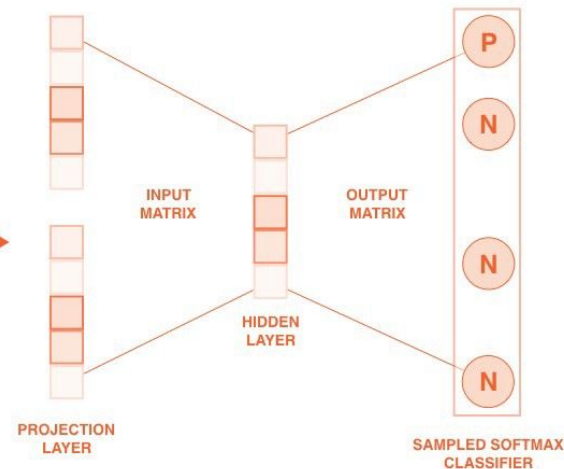
27 Современный подход



1.Swing I2I Graph



2.Random Walk Generation



3.SGNS Model

28 Как оценить качество?



Продукт:

- Принятия
- Отклонения
- Удаления
- Игноры

Модель:

- NDCG
- AUC
- Log Loss
- и т.д.



29 Специфика прода

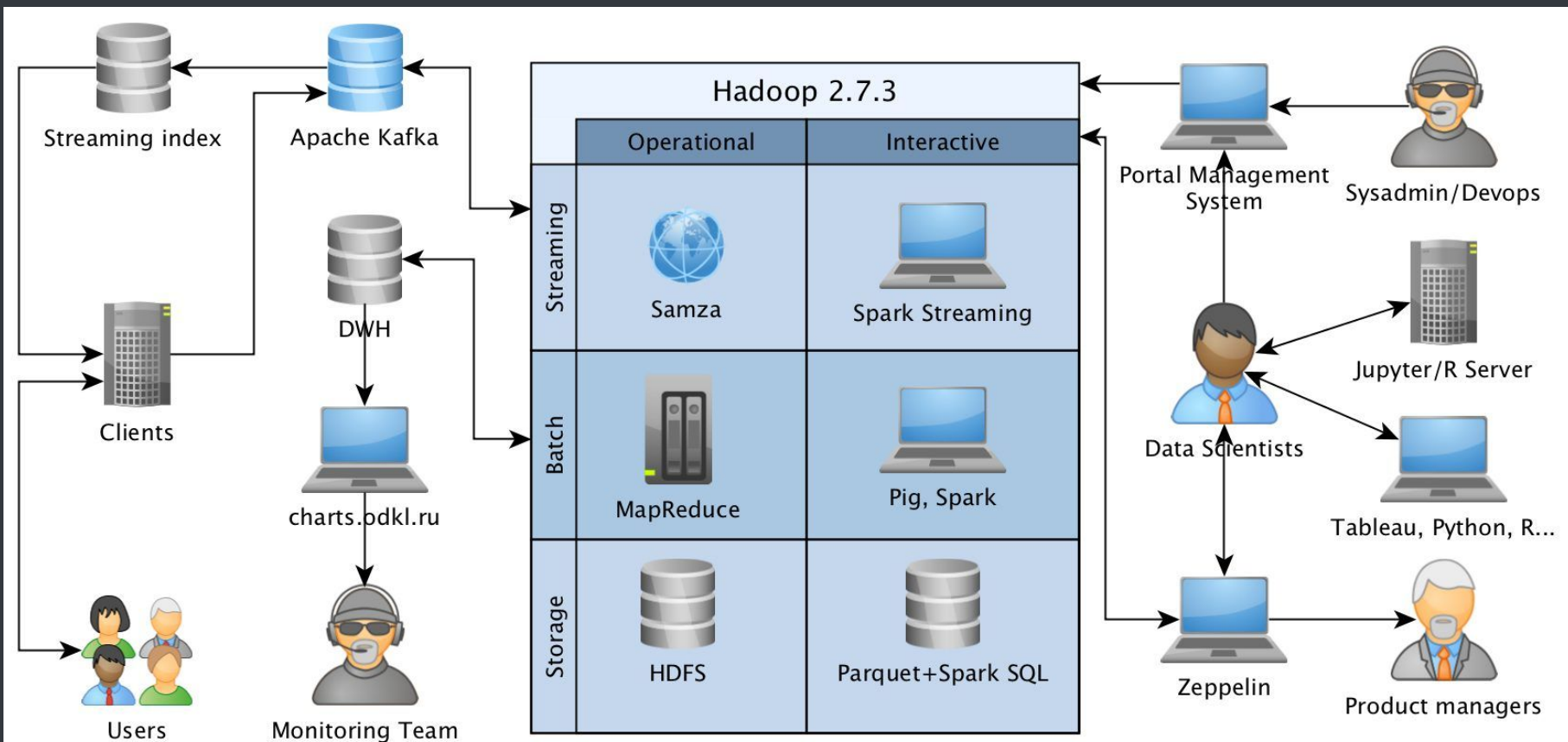
- 1) Людей надо стимулировать на действие и это отдельная задача
- 2) Мы не можем растить социальный граф до бесконечности и поэтому важно понять когда стоит уменьшить нагрузку триггеров на пользователя
- 3) Оффлайн метрики после определенного момента перестают напрямую коррелировать с онлайн
- 4) Модель создает bias

В контексте поведенческой модели



<u>Мотивация</u>	<u>Триггеры</u>	<u>Доступность</u>
Качественные рекомендации	<ul style="list-style-type: none">1) Пуши2) Нотификации3) Портлеты в ленте4) Новые лица	Минимум кликов до целевого действия

31 Инфраструктура





- 1) <http://www.michalkosinski.com/home/publications>
- 2) <https://arxiv.org/abs/1411.511>
- 3) https://en.wikipedia.org/wiki/Big_Five_personality_traits
- 4) https://www.mebook.se/images/page_file/38/Fogg%20Behavior%20Model.pdf
- 5) <https://blog.prototypr.io/zillions-of-homepages-for-zillions-of-customers-af85272d536b>



ОДНОКЛАССНИКИ