

Сложность — норма жизни

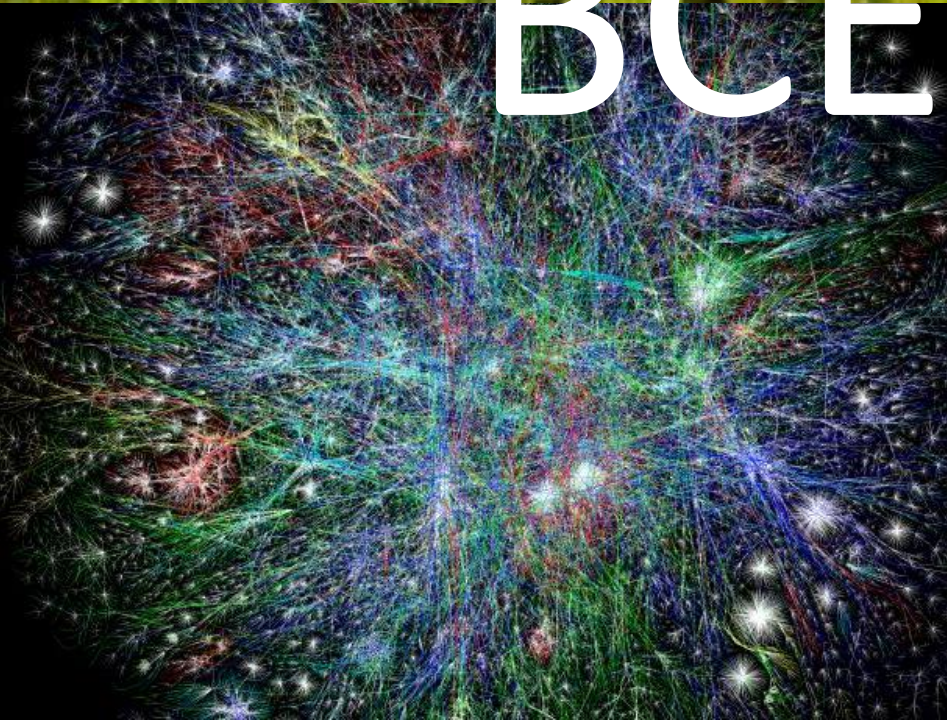
Станислав Макаров

КОНФЕРЕНЦИЯ

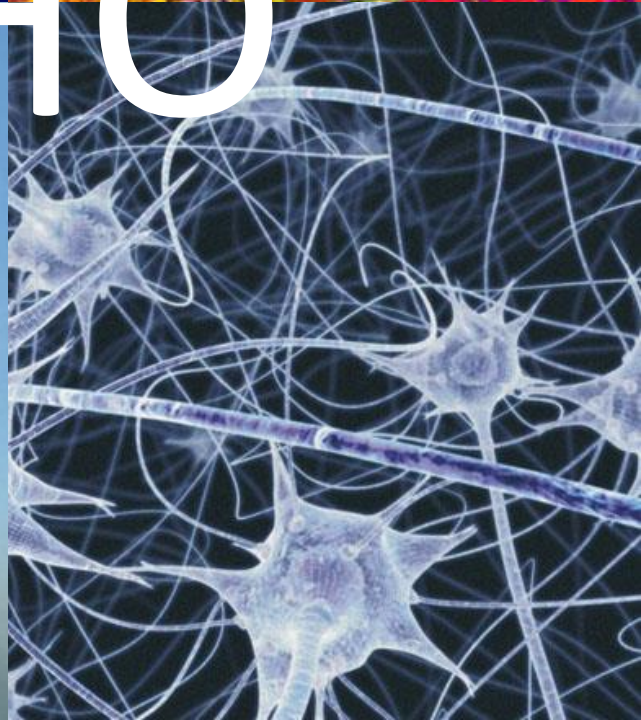
ВІ в России: от простого к сложному

2 июня 2016

МОСКВА

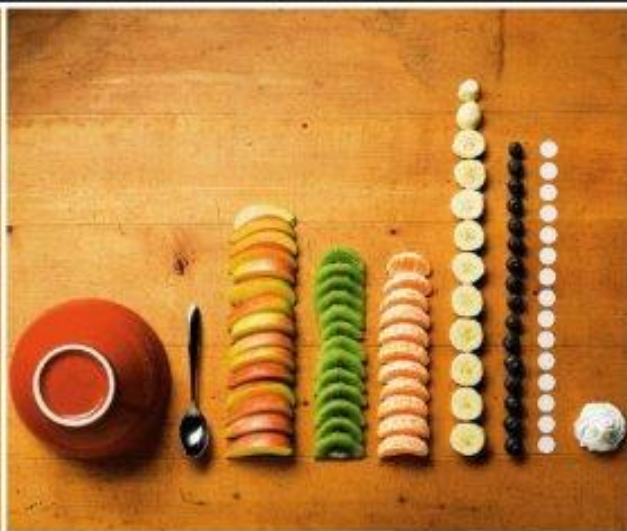
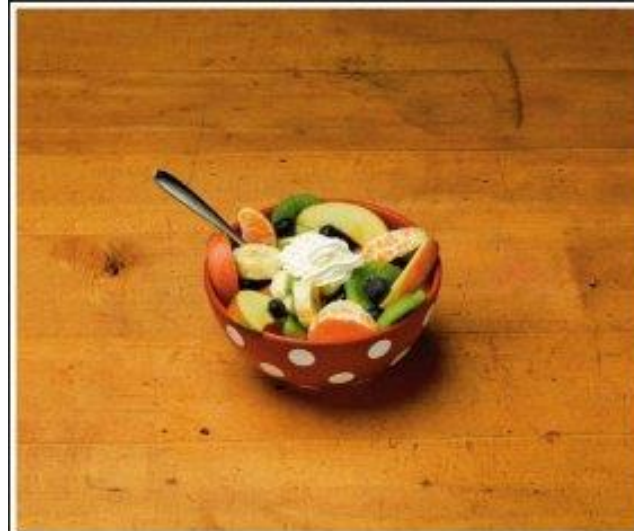
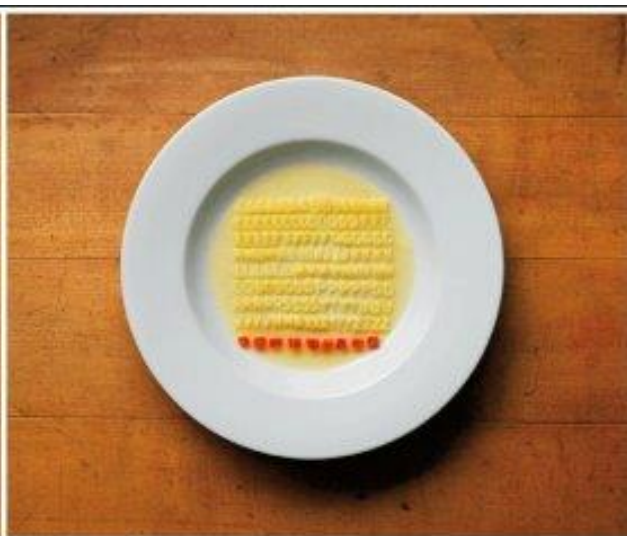


ВСЕ СЛОЖНО



А хочется простоты!

Разложить все по полочкам — не аналитика





Welcome, Visitor | [Login](#)

[Search](#)

About us

SenseMaker®

Training

Resources

Despatches

Blogs

Network

Contact

Register

Dave Snowden

Making
sense of
complexity in
order to act

Cynefin Framework

Learn more about Dave Snowden's widely-cited, award-winning framework for decision making

[LEARN MORE](#)

SenseMaker®

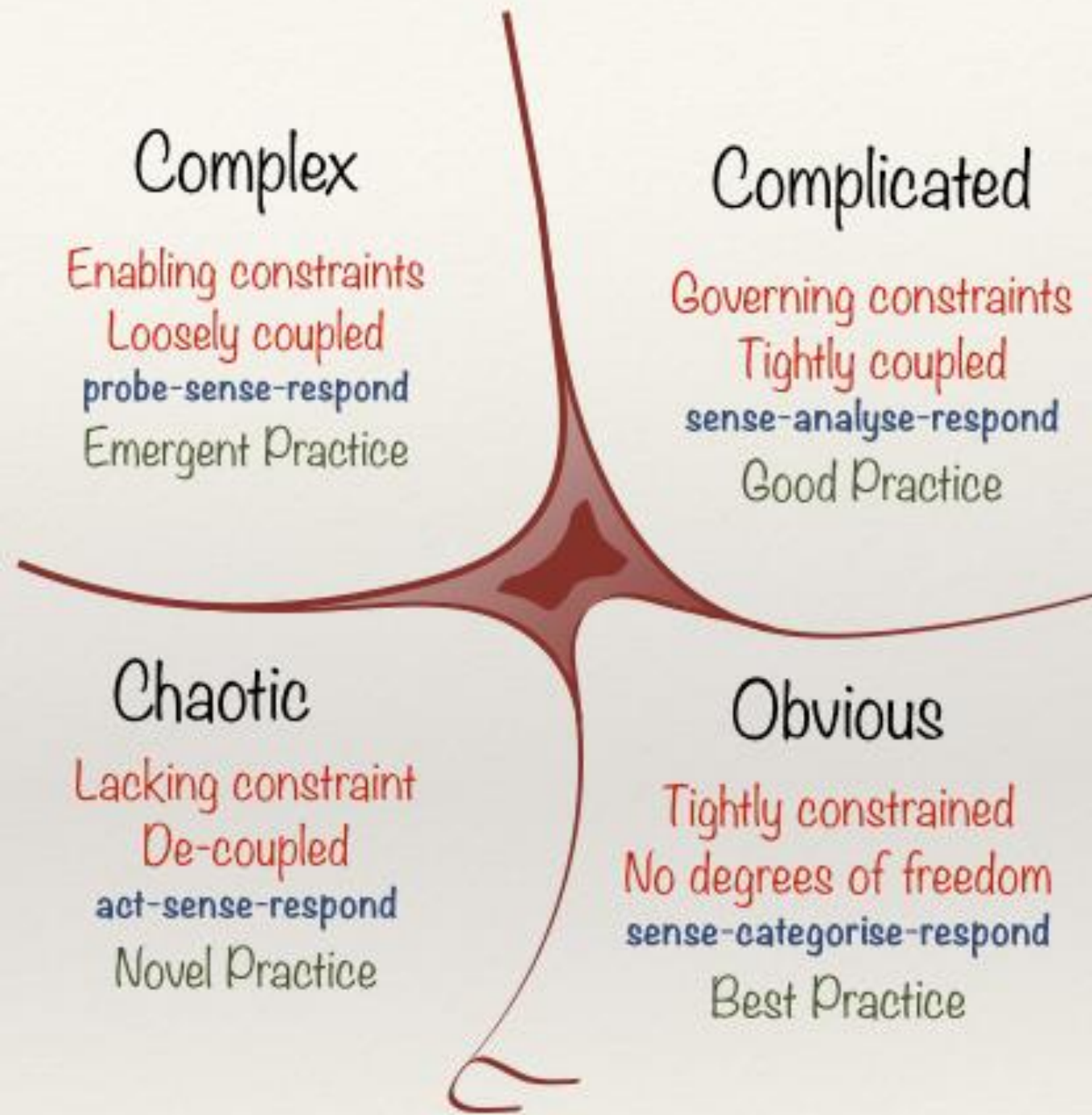
Links micro-narratives with human sense-making to create advanced decision support, research and monitoring capability.

[LEARN MORE](#)

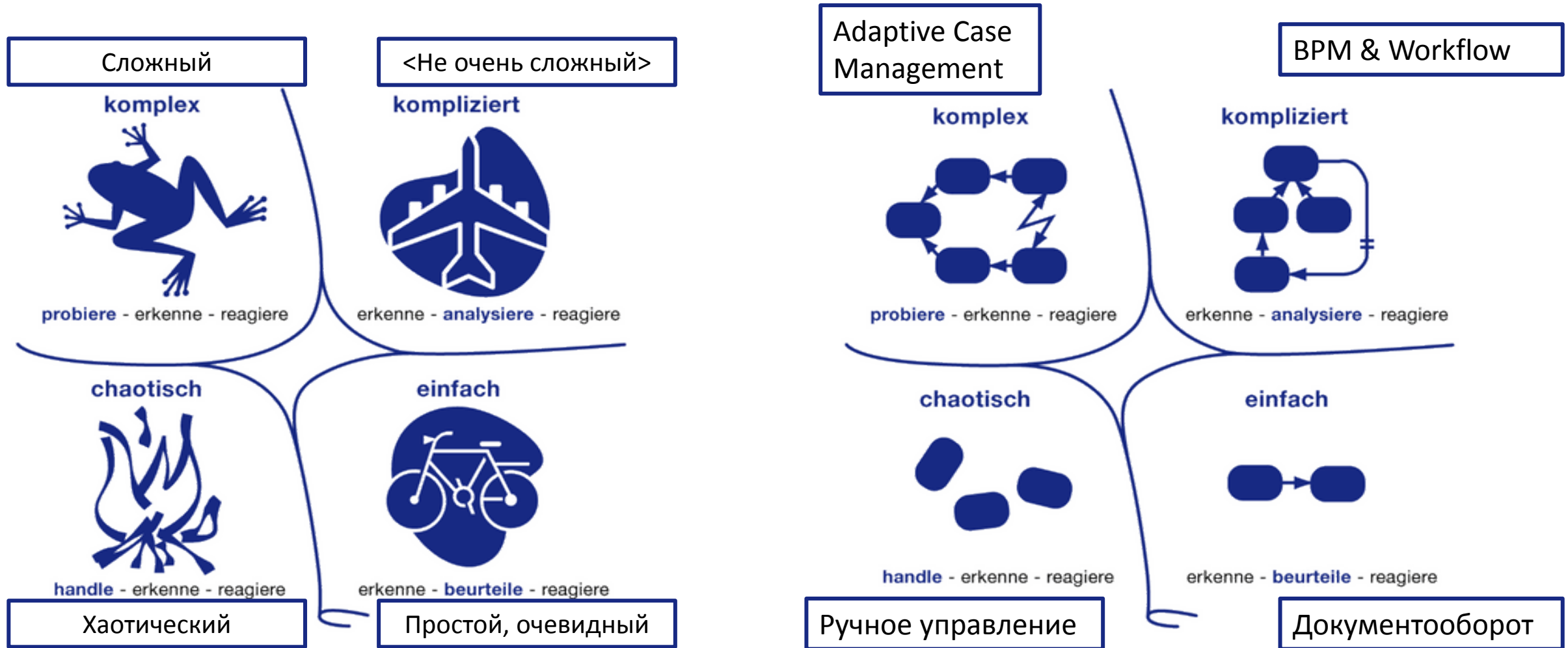
Найти в сложности смысл —
чтобы действовать

Фреймворк Cynefin

- Простое
- Сложное (1)
- Сложное (2)
- Хаос



Synefin – наглядно о сложном



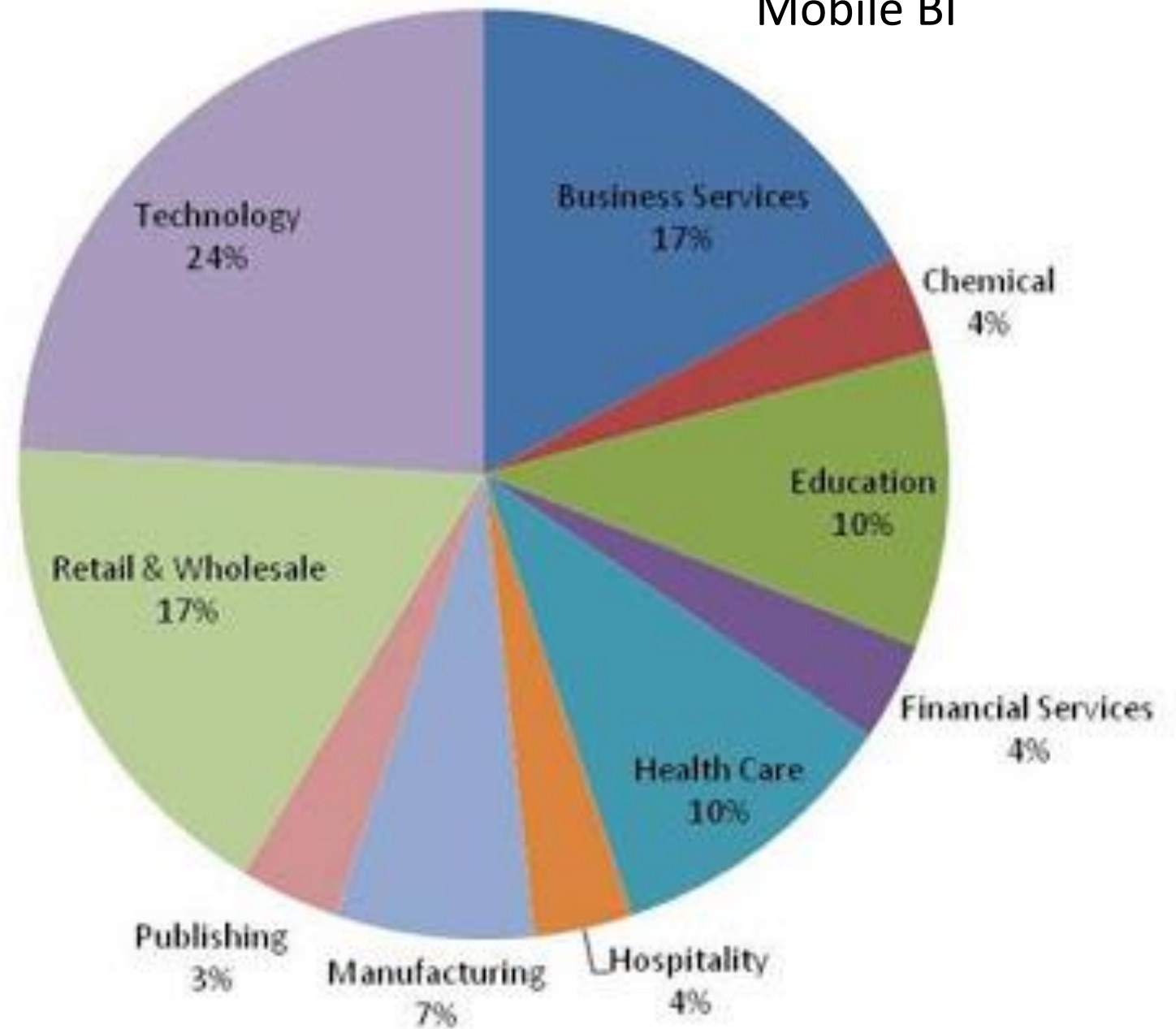
Задача аналитики



Из кучи данных собрать красивую картинку,
причем, достоверную и осмысленную

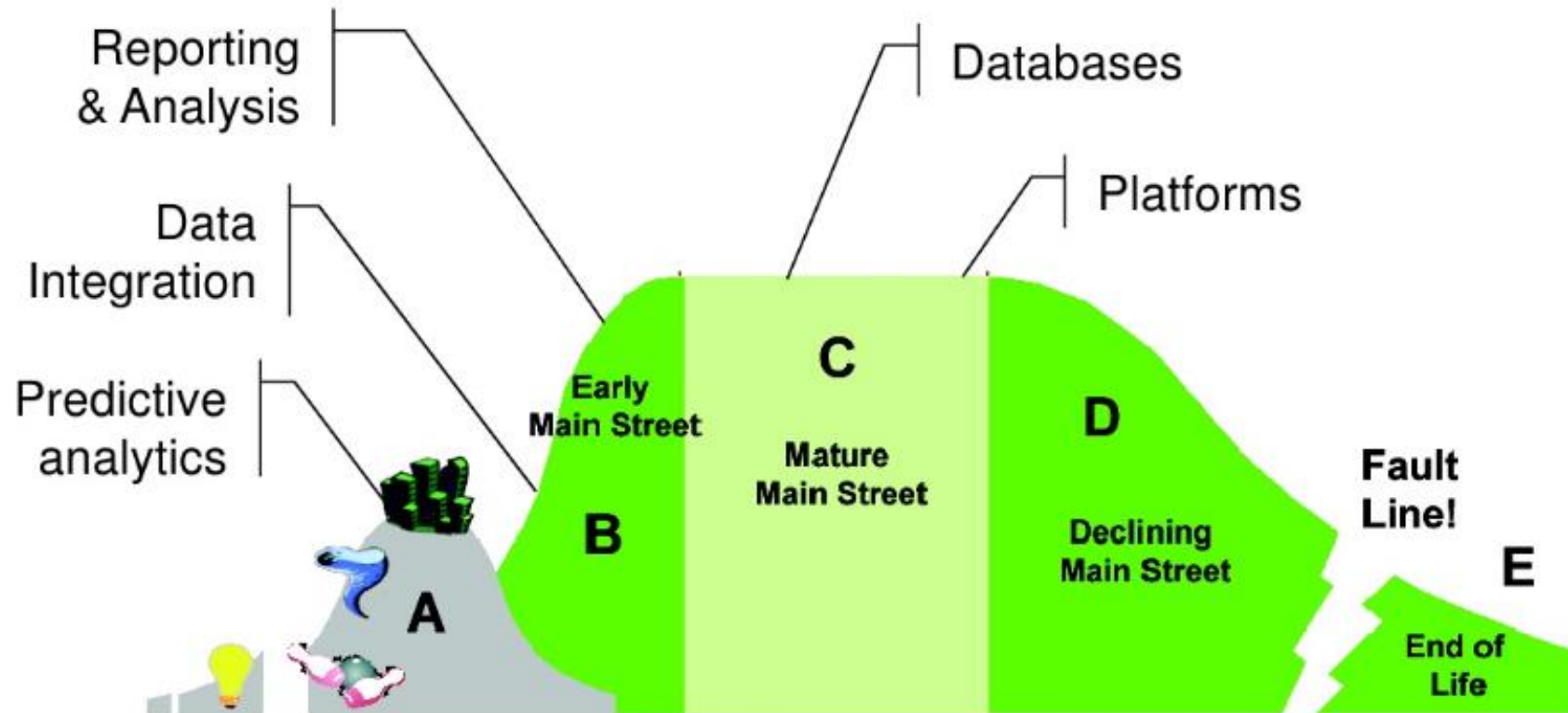


BI нужен всем



ВІ становиться мейнстримом

Спрос и предложение растут, цена и прибыль падают



Коммодитизация ВІ неизбежна



Риски:

- Упрощенный подход
- Неверная интерпретация данных
- Слепое доверие системе
- Шарлатанство

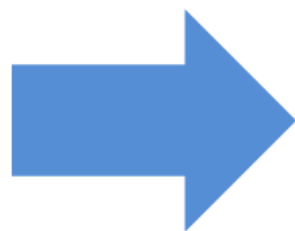
Выгоды:

(о хорошем вам и так расскажут)

Каждый суслик ~~агроном~~ аналитик

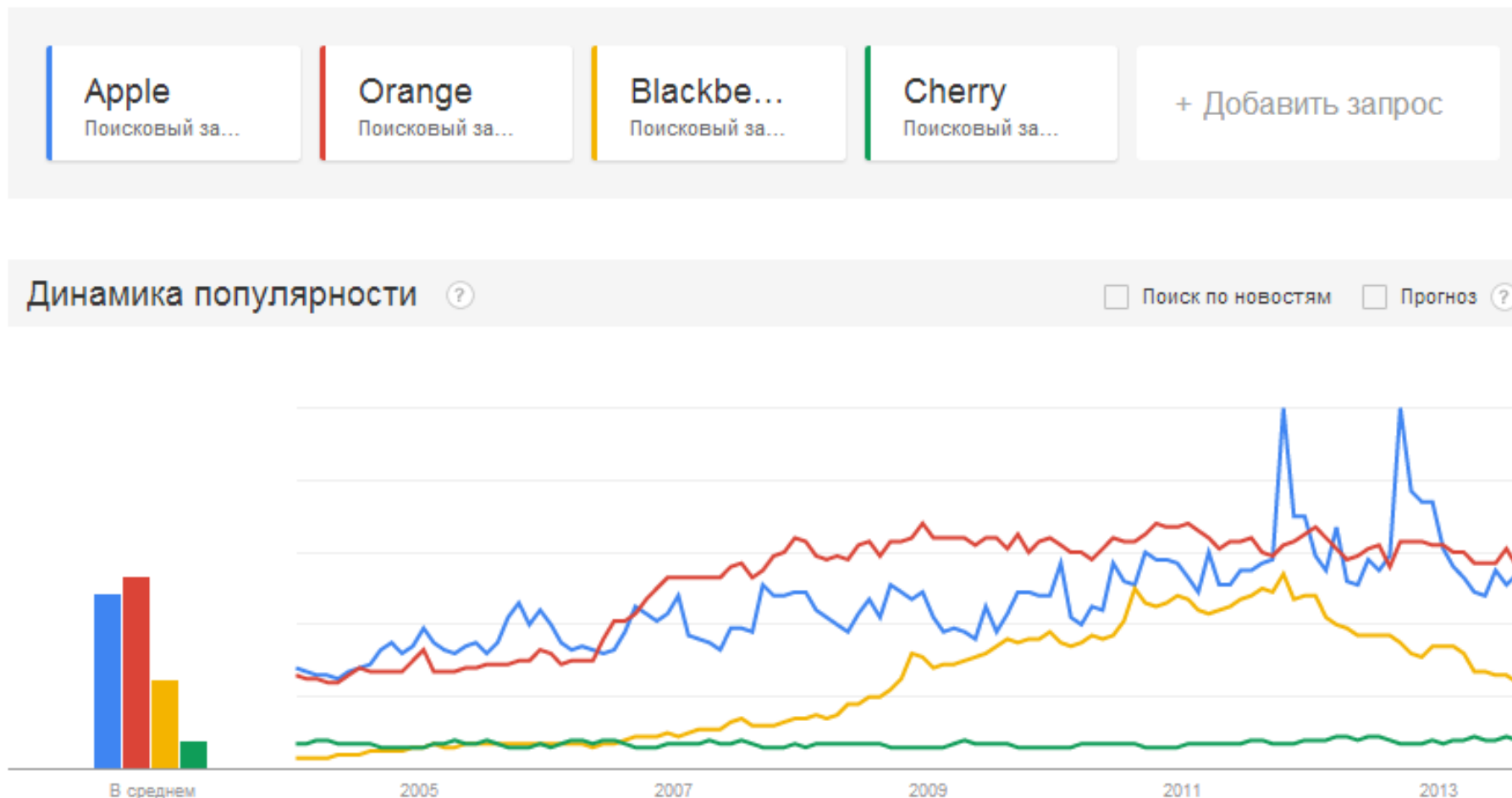
Простота – это иллюзия

If



Then

Любой вопрос – любой ответ

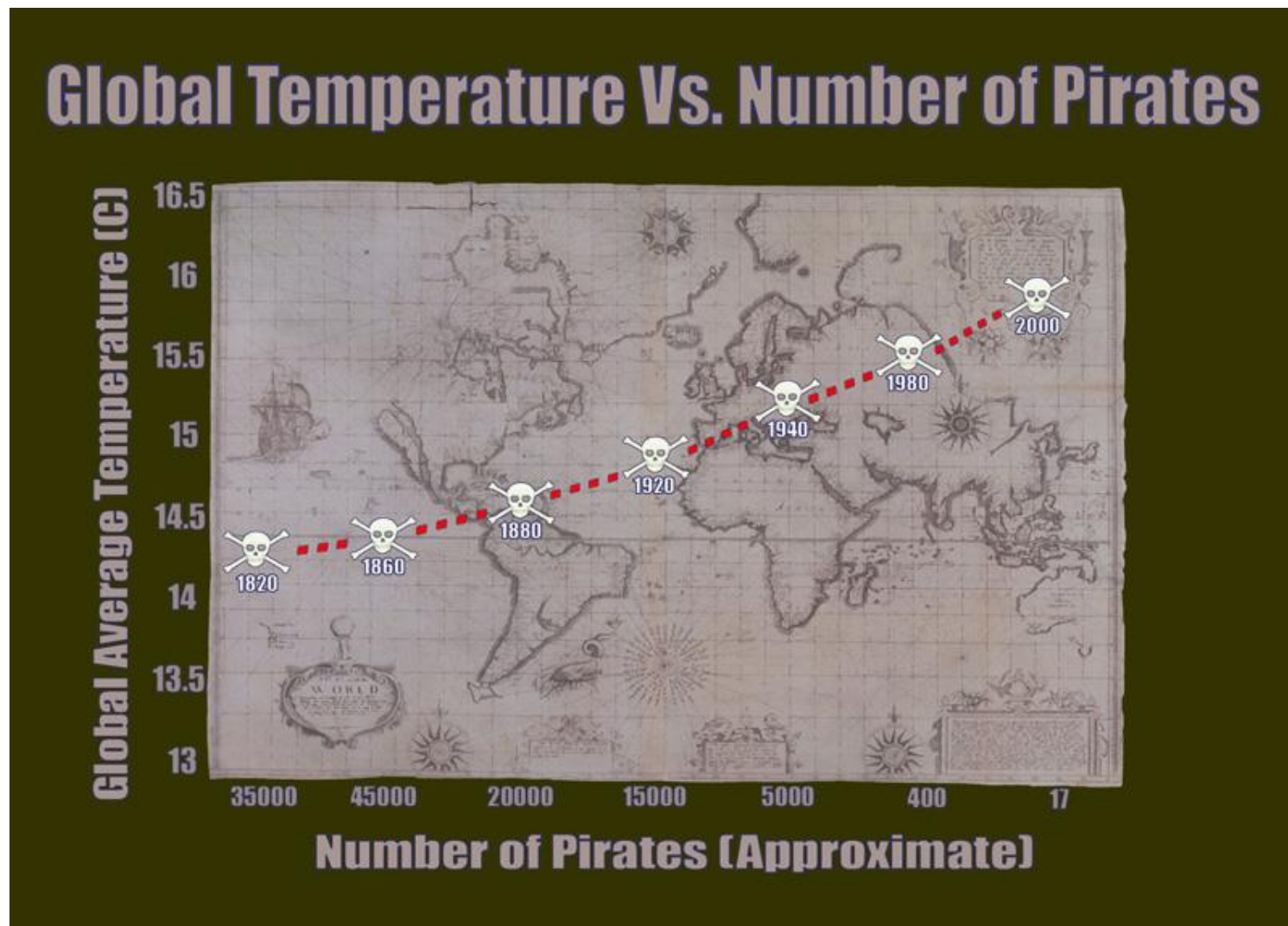


Вишня пользуется гораздо меньшим спросом, чем яблоки, апельсины и ежевика.

– британские ученые (с)

Пираты и глобальное потепление

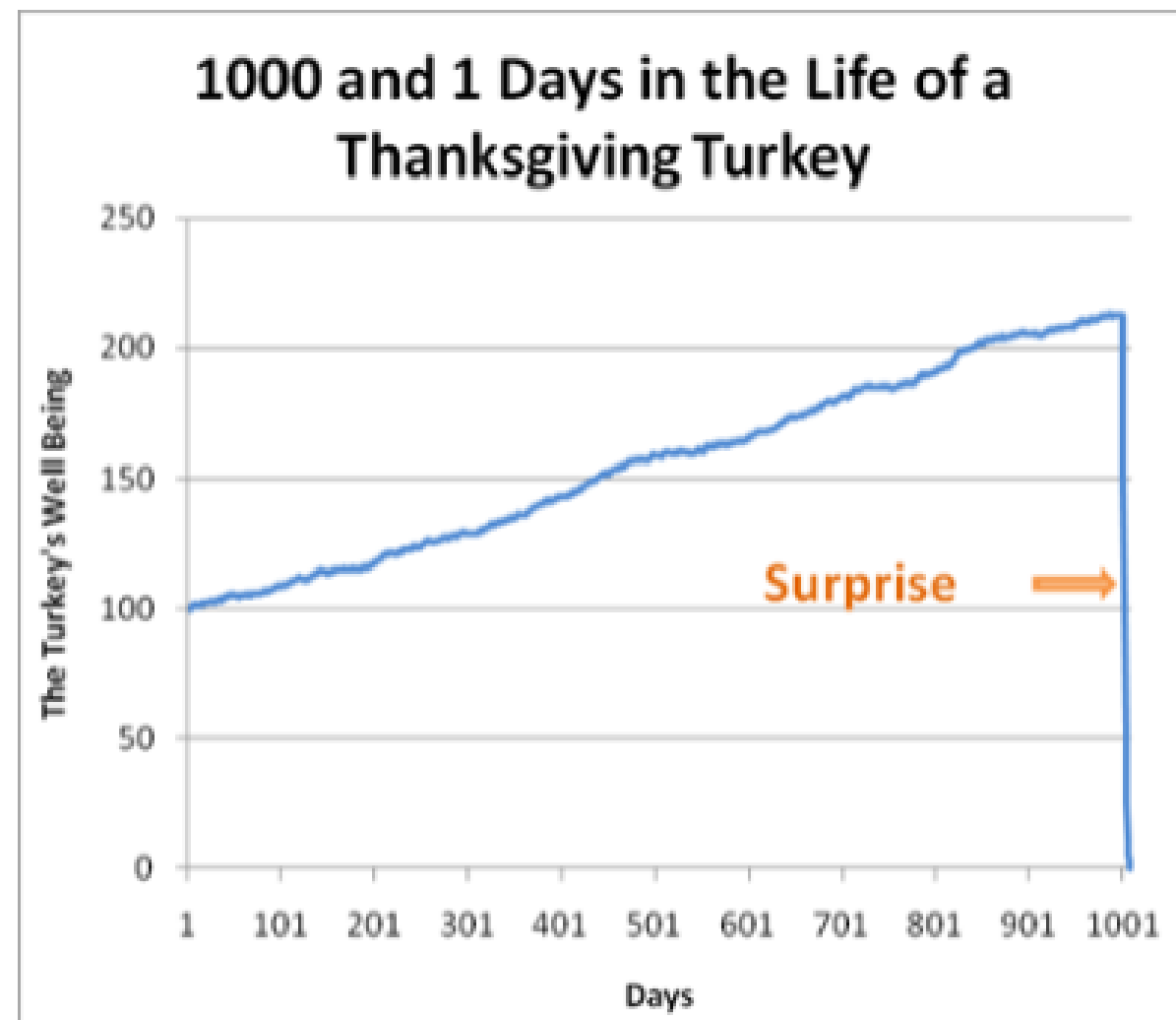
- Число пиратов сократилось
- Температура выросла
- Есть связь?



Ошибка индейки, Нассим Талеб

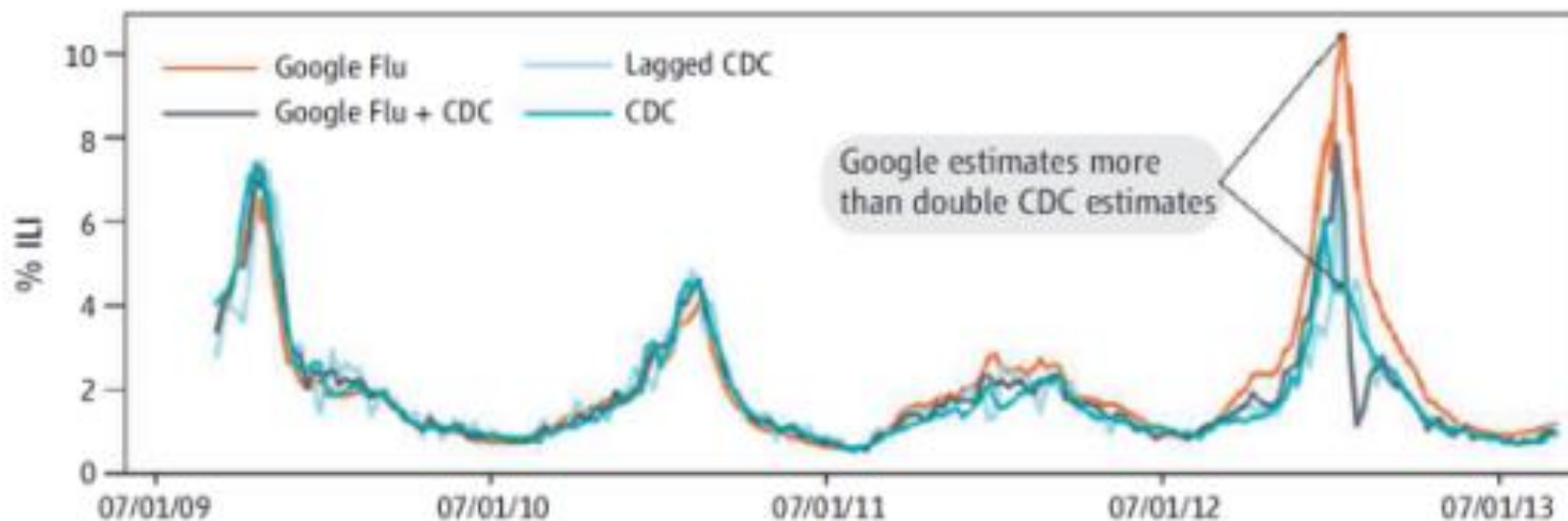
"Мясник откармливает индюшку тысячу дней; с каждым днем аналитики все больше убеждаются в том, что мясники любят индюшек «с возрастающей статистической достоверностью». Мясник продолжает откармливать индюшку, пока до Дня благодарения не останется несколько суток.

Тут мясник преподносит индюшке сюрприз, и она вынуждена пересмотреть свои теории – именно тогда, когда уверенность в том, что мясник любит индюшку, достигла апогея и жизнь индюшки вроде бы стала спокойной и удивительно предсказуемой."



Провал Google Flu Trends

- Если люди ищут статьи о гриппе и лекарствах, значит началась эпидемия
- На самом деле это показывает уровень тревожности по поводу гриппа, но не уровень заболеваемости



Верите ли вы?



ВЕГЕТАРИАНЦЫ РЕЖЕ ОПАЗДЫВАЮТ НА САМОЛЕТ.

Это еще одно открытие стартапа Kaggle, который обнаружил, что люди, сообщившие авиакомпаниям, что хотят получить во время полета вегетарианскую пищу, менее склонны опаздывать на рейс. Объяснение двоякое: во-первых, такой заказ персонализирует услугу и делает ее более значимой в глазах потребителя; во-вторых, пассажиры, которые заранее заботятся о завтраке, в целом лучше планируют свою жизнь.

Разве мы не знали это всегда?



ОЦЕНКИ В ВАШЕМ ДИПЛОМЕ НИЧЕГО НЕ ЗНАЧАТ.

Google в последние несколько лет активно переводит свой HR на научную основу, старается все больше решений принимать на основании твердых фактов и статистических исследований. Одно из открытий, [сделанных компанией](#), — средний балл в аттестате или дипломе никак не коррелирует с успехами сотрудника на работе.

Мистика!



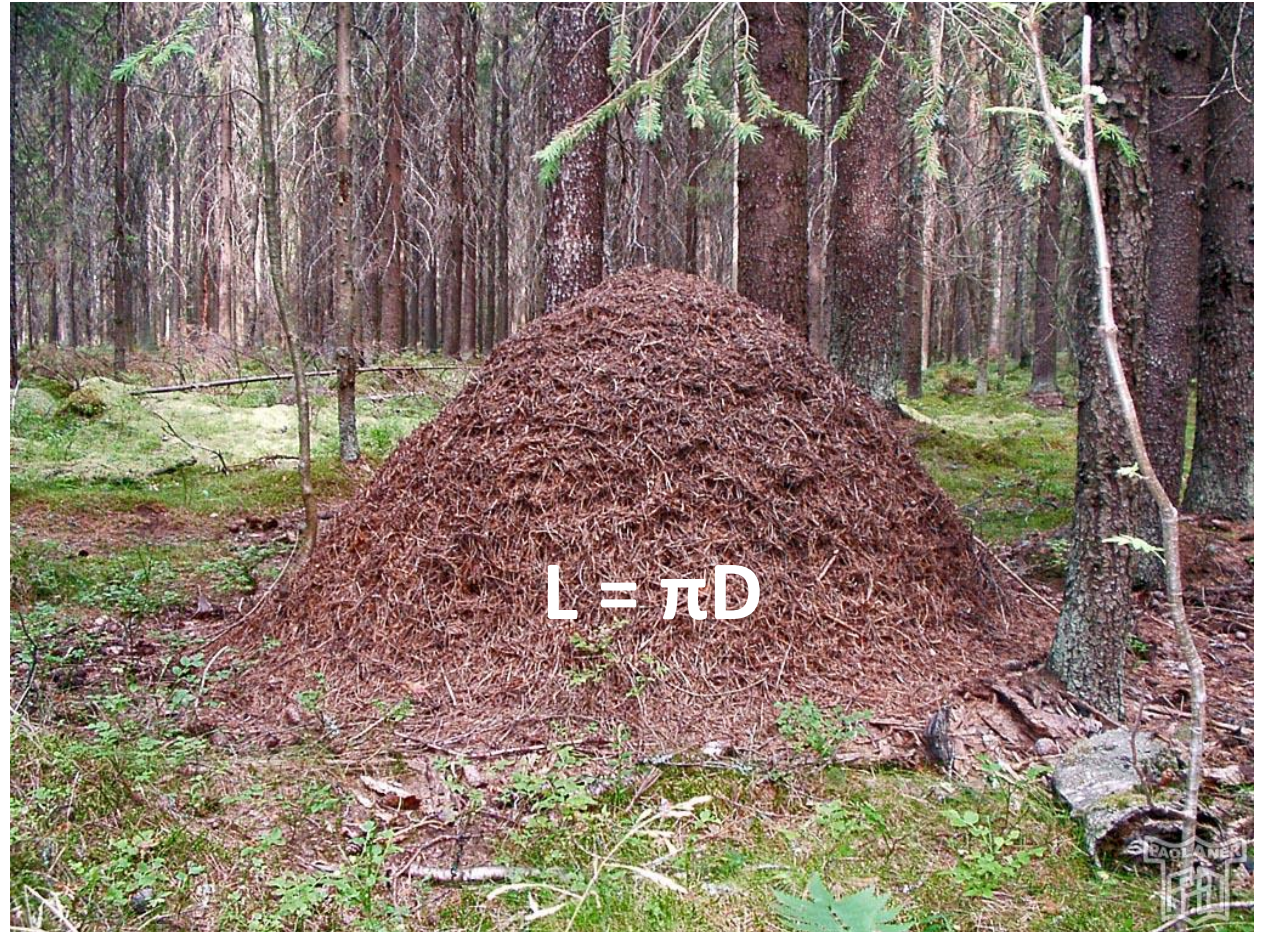
ПРОДАВАТЬ ЛУЧШЕ В НОВОЛУНИЕ

Софтверная компания InsideSales

[исследовала](#) сделки, совершенные группой продавцов за девять лет, и обнаружила, что средний объем сделок в новолуние в два раза выше, чем в полнолуние, и на 43% выше, чем в промежуточных фазах Луны. Убедительного объяснения не найдено, но исследователи предполагают, что ответ кроется в динамике эмоций потребителей, которая, в свою очередь, связана с природными циклами.

Не может быть!

На основании большого массива наблюдений, сделанных в разных климатических условиях, биологи обнаружили, что окружность муравейника примерно в три раза больше его диаметра.



Теория прежде всего



Если нет теоретической модели и гипотезы – это не аналитика, а гадание на кофейной гуще

Мнимые закономерности на основе статистики

В общественных и социальных науках, политологии, маркетинге, психологии, экономике, антропологии на самом деле нет сколько-нибудь строгой теории — теории в том смысле, как это понимают физики.

"Вся наука — или физика, или коллекционирование марок" — Эрнест Резерфорд.

Раньше было три вида лжи:
ложь, наглая ложь
и статистика.

Теперь у нас есть
Big Data.



Atkritka.com facebook.com/AntiBigData

Что такое VI?

«Возможность понимания связей между представленными фактами.»



Hans Peter Luhn, IBM, 1958

Без математики ничего не получится

$$\frac{\partial}{\partial a} \ln f_{a, \sigma^2}(\xi_1) = \frac{(\xi_1 - a)}{\sigma^2} f_{a, \sigma^2}(\xi_1) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(\xi_1 - a)^2}{2\sigma^2}\right)$$
$$\int_{\mathbb{R}_+} T(x) \cdot \frac{\partial}{\partial \theta} f(x, \theta) dx = M\left(T(\xi) \cdot \frac{\partial}{\partial \theta} \ln L(\xi, \theta)\right) = \int_{\mathbb{R}_+} \frac{\partial}{\partial \theta} \ln L(x, \theta) \cdot f(x, \theta) dx$$
$$\int_{\mathbb{R}_+} T(x) \cdot \left(\frac{\partial}{\partial \theta} \ln L(x, \theta)\right) \cdot f(x, \theta) dx = \int_{\mathbb{R}_+} T(x) \cdot \left(\frac{\frac{\partial}{\partial \theta} f(x, \theta)}{f(x, \theta)}\right) \cdot f(x, \theta) dx$$

Кирилл Корнильев, IBM:

«Мы говорили о нехватке программистов, но неожиданно сейчас мы столкнулись с нехваткой математиков.

То есть, управление бизнесом сейчас, из-за внедрения аналитики, потребовало высшей математики в гораздо большем объеме. И дефицит профессии, которую условно можно назвать «аналитик» (или data scientist), сейчас налицо.

А математика это вам не экраны пальчиком на планшете двигать, математику учить надо. Тут время понадобится – поэтому эффект от внедрения аналитики будет нарастать довольно долго.»



Больше данных – не лучше

К 2018 году 90% внедренных озер данных будут бесполезны потому что они будут переполнены информацией, собранной неизвестно с какой целью. (Gartner, Strategic Planning Assumption, Gartner BI Summit, 2015).

Данные в озере могут быть неконсистентны и не иметь метаданных, поэтому реально только очень опытные аналитики, хорошо знающие контекст, смогут сливать и согласовывать данные из разных источников.



Готовы ли вы к большим данным?

Неконтролируемый доступ к различным источникам вызывает риск того, что человек не может отделить существенную информацию от несущественной.

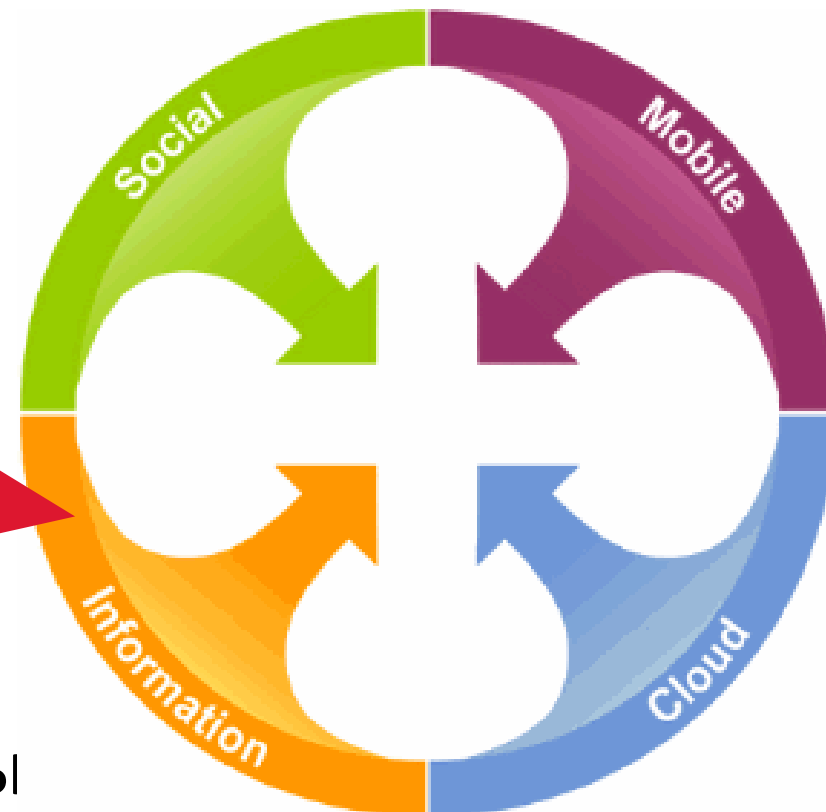
Умберто ЭКО



3-я платформа IDC, Nexus of Forces Gartner

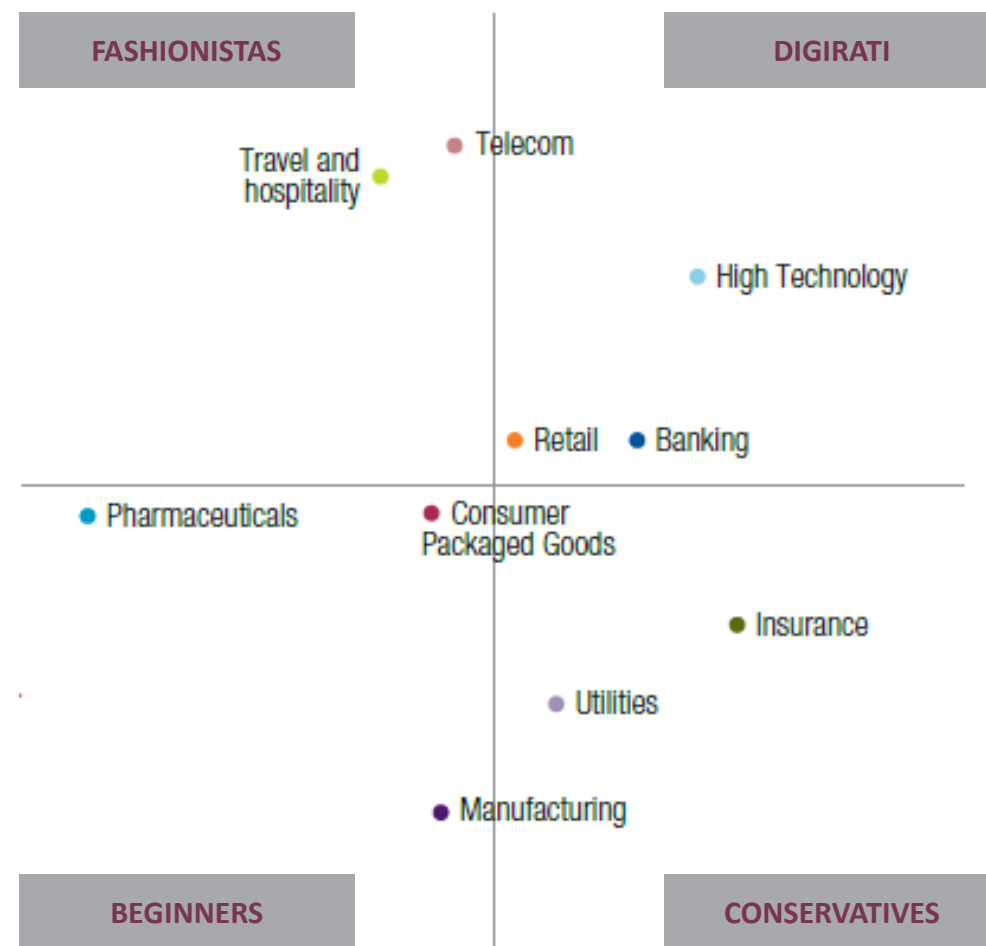
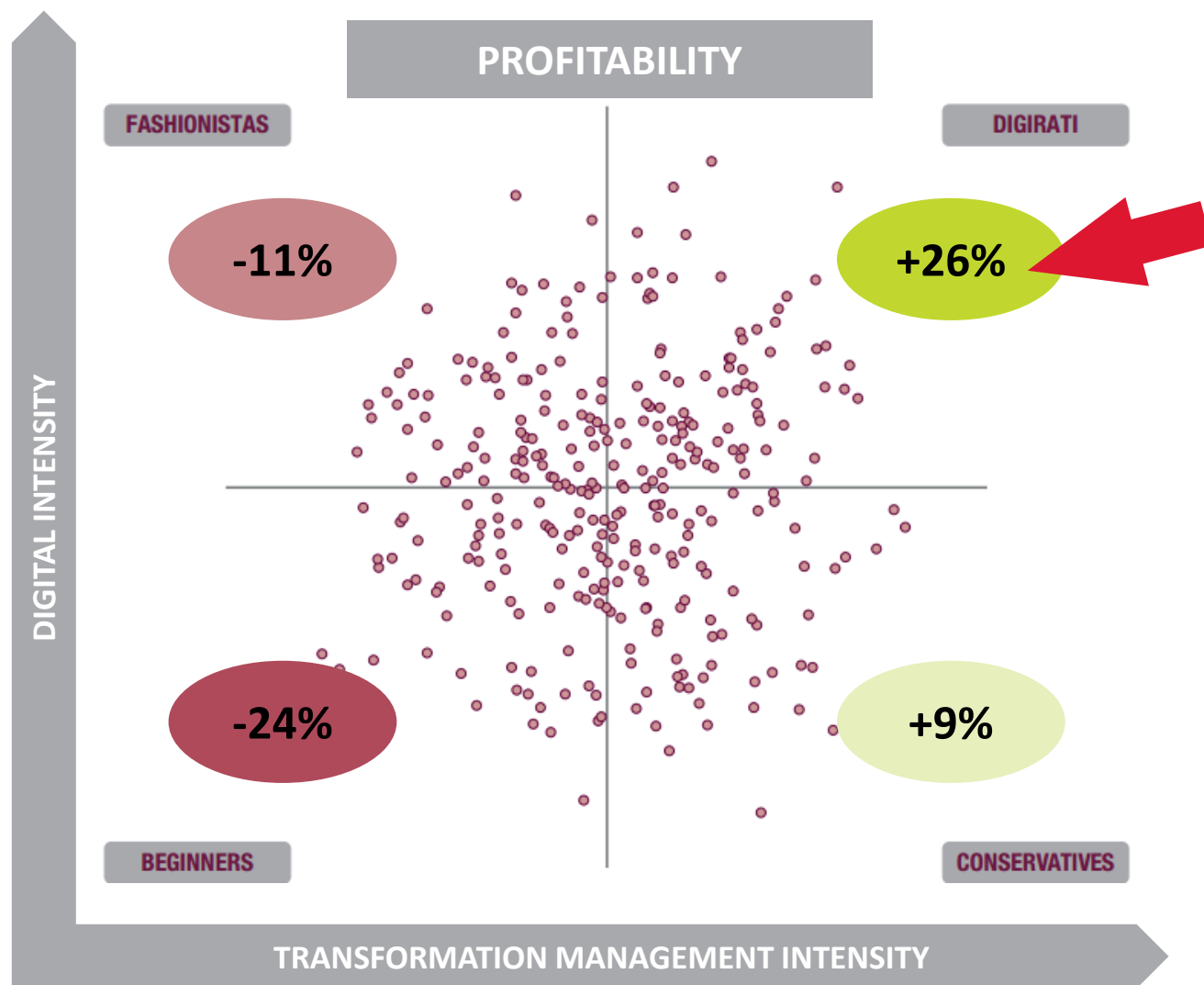


Аналитика –
ключевой элемент
новой ИТ-парадигмы



Мобильность, Облака, Аналитика, Социальность

Цифровое предприятие: прибыль выше



Так брать ли мне зонтик?

The screenshot shows the weather.com website. At the top, there's a navigation bar with links like Home, Forecasts, Severe Weather, Maps, Weather Apps, Video, Witness Weather, and Travel. Below this is a search bar with the text "Enter Zip, City or Place (e.g. Disney World)" and a "FIND WEATHER" button. To the right of the search bar are links for "My Saved Locations", "Save a Location", and "SEND TO MY PHONE". The main content area is titled "National & Regional Forecasts" and is updated as of November 7, 2010, at 12:43 pm ET. It features a section for "Dual storms: West and Northeast" with a map of the Northeast showing a low pressure system (L) over Boston and a high pressure system (H) over the Atlantic. The map also labels NYC and Philly. Below the map, there's a list of links: "New England: First snow for some ahead", "Slideshow: Early-season snow", "Latest on Tomas", "Tropical Update", "NFL Gametime Weather Index forecasts", "Indonesian volcano sparks travel chaos", and "National Forecast".

Людям нужны конкретные ответы.
А не «многобукв», или,
еще хуже - «многоцифр».

The graphic is a light blue rectangle. At the top left is a black silhouette of an open umbrella. To its right, the text "do i need an umbrella?" is written in a lowercase, sans-serif font. Below this text is a small input field with the placeholder text "enter zip code" and a button labeled "Well, do I?". Below the input field, the location "Newton Highlands, MA" is displayed. In the bottom left corner, a large white box contains the word "NO" in a large, bold, sans-serif font. To the right of this box, the temperature "(40° F)" is shown, followed by the text "no chance of rain" in a bold, sans-serif font.

О себе: Станислав Макаров



20+ лет в работы в ИТ

Программистом, продавцом, дистрибутором, инженером,
тренером, журналистом, бизнес-аналитиком

Член экспертного совета Докфлоу

facebook.com/s.makarov15
twitter.com/Stas_Makarov
Skype: Stanislav.Makarov
s.makarov15@gmail.com



www.facebook.com/AntiBigData