



# Big Data | использовать нельзя регулировать

---

**Пеникас Генрих, к.э.н.**

Доцент, старший научный сотрудник, ВШЭ

Директор проектов, Казначейство, Сбербанк

14 марта 2017 г.

г. Москва, ул. Мясницкая, д. 9/11  
Высшая школа экономики



### Цель доклада

- Ответить на вопрос – есть ли потребность в регулировании Big Data?
- Чтобы каждый участник узнал новое о явлении Big Data.

### Структура доклада

*№ слайда*

3. Публикация регулятора о регулировании;
4. Что напоминает предложение?
5. В чем эффективность Big Data?
6. Чего стоят данные?
7. Кому полезно регулирование?
8. Как быть с невозможным?
9. При чем здесь профстандарты?



## В дек 2016 в Европе решили, не регулировать ли Big Data?

---

1. Европейский регулятор банков задался вопросом, регулировать ли Big Data [1]
2. Ввели определение Big Data, чтобы знать, кого потом регулировать;
3. Указали, что есть выгоды, но много недостатков;
4. Недостатки включают:
  1. Неэтичное использование частных данных;
  2. Предложение не тех продуктов;
  3. Непредложение продуктов (отказ в кредите);
5. Попросили комментарии на **23 вопроса** до **17 марта 2017**;
6. **Появление вопроса подобно тому, что было накануне кризиса 2007 г., а именно...**



## Что делать, если никто не понимает? Переложить ответственность

	Копула		Big Data
1956	Первое упоминание термина в статье математика - Шкляр	1943	Первое упоминание о нейронных сетях - У. Маккалок и У. Питтс
2000	Рост интереса к копулам – статья: <b>копула + корреляция дефолтов</b> - Дэвид Ли	1980- - 1990	Рост интереса к нейронным сетям и приложениям - Карты Кохонена
2000 - 2009	Использование гауссовской копулы в CDO - Халдейн (2009): 1 млн страниц приходилось на 1 инвестора в CDO <sup>2</sup>	2011- 2015	Стремительный рост запросов в Google на тему Big Data
2000- 2009	Ответственность за корректность возложена на рейтинговые агентства	2016	Предлагается переложить ответственность на регулятора
2007- 2009	Мировой финансовый кризис	???	<b>кризис</b>
2009	Копула – Financial Times [2]: <a href="#">формула, которая убила Уолл-Стрит</a>	???	<b>Нейронные сети – «формула, которая убила бизнес»</b>

Если не возложить ответственность на третью сторону, то как умерить мощь такого инструмента?

**А действительно ли инструмент такой мощный? Эффективный?**



## Big Data | действительно ли так эффективна?

---

### ПРИМЕР ОФЛАЙН РИТЕЙЛА

- Tesco – экономия от использования Big data – **150 млн дол / год** экономия расходов
- Чистая прибыль компании – 100 млн дол => 150% - Эффективно использовать, НО
- Операционные расходы – 53 млрд дол => **0.3% - Статистическая погрешность**

Источник - Д. Репин (2017) [3]

### ПРАГМАТИЗМ

- Говорим, что технология эффективна, поскольку покупают до 80% предложений.
- Но если бы клиентам **случайно предлагали товары**, они бы тоже покупали.
- НЕТ критериев эффективной сегментации
  - *методы кластер-анализа долго считались статистикой, не математикой.*
- Эффективность НЕ доказуема!
- **Если Вы не согласны, значит можем пойти от обратного, т.е....**
- **Значит, данные должны дорого стоить, если эффективность от них высока, так?**



## Уникальные случаи – \$ 3–5 k / данные человека; иначе – до \$8.

Дата	Цена	ед.изм.	Quotes	Комментарии
12.06.13 <a href="#">FT</a> , [4]	<b>\$0.26</b>	\$ / данные 1 чел.	For \$0.26 per person, buyers can access lists of people with specific health conditions or taking certain prescriptions.	Интересный калькулятор стоимости данных одного человека
22.04.14 <a href="#">Guardian</a> , [5]	<b>\$1.00</b>	\$ / месяц	I ran through various scenarios (pretending I was a rich CEO or a teacher) and it is hard to find a job and lifestyle that is worth more than one dollar.	Стоимость данных варьируется от 1 до 8 долларов в месяц в зависимости от частоты совершаемых действий.
	<b>\$8.00</b>	\$ / месяц	Datacoup will only pay you \$8 a month and while Handshake allows you to negotiate with brands wanting to buy your data, it is still in beta so it remains to be seen what money there is to be made there;	
	<b>\$432.00</b>	\$ / данные 1 чел.	Shawn Buckles, the Dutch student who sold his own personal data for £288	
	<b>\$2 733.00</b>	\$ / данные 1 чел. / мес	He [ Federico Zannier ] managed to make \$2,733 but had to share his keystrokes, mouse movements and even provide a screenshot every 30 seconds (за 1 месяц)	1 действие каждые 30 секунд; 2 - в минуту; 120 - в час; ~3 тыс. в день.
	<b>\$0.03</b>	\$ / действие		За месяц было 86.4 тыс. действий в месяц
23.11.15 <a href="#">Telegraph</a> , [6]	<b>\$4 861.50</b>	\$ / данные 1 чел. / мес	the average consumer values their personal data at £3,241	оценка цены предложения данных

Если данные дорого не стоят, то компании стремительно не обогащаются на Big Data.  
**Зачем же нужно регулирование?**



## Регулирование выгодно конкурентам

---

- **Telecom / IT || 2009: The Economist [7]**
  - Регулирование помогло США обогнать Японию в инновациях.
  - Монополисты были замещены **малыми, специализированными компаниями.**

*THE most important factor that led to America's stunning success in information technology was not the free market but **government regulation**. ... they changed the industry's structure, replacing monoliths with smaller, specialised companies...*

- Наверное, имели в виду Amazon, Alphabet (Google), Apple etc.

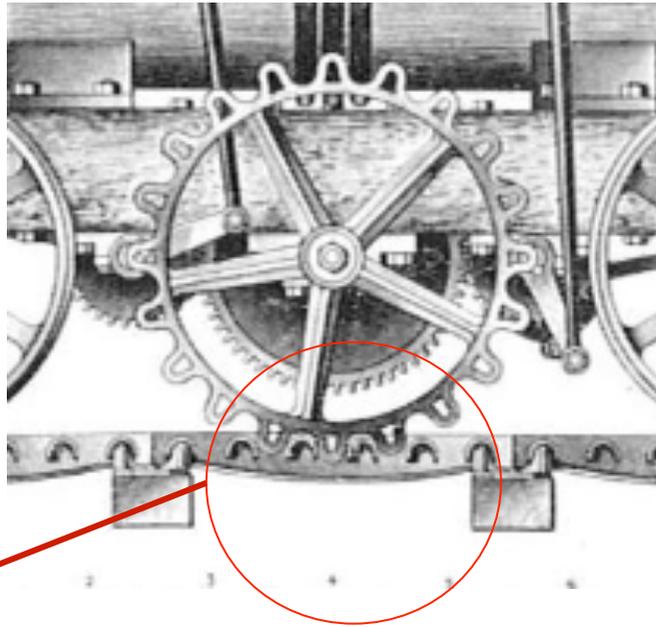
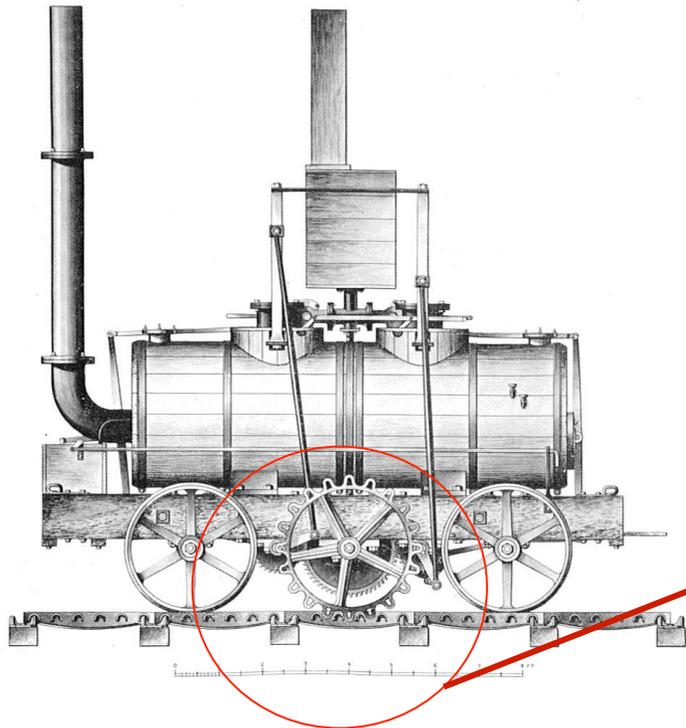
- **Banks || 2016: The Economist [8]**
  - Ужесточение регулирования (Базель III+) **выгодно американским** банкам, поскольку у европейских меньше риски и капитал нужно наращивать в ЕС;

*But **European bankers and officials had complained** for months that the proposals would **penalise banks** that have lots of (low-risk) corporate loans or mortgages (eg, in Germany or Sweden). They sniffed an American plot: **American banks, holding fewer such assets, would be untouched.***

- **Как отказаться от интуитивно необходимого?**



## Как возникают шедевры и инновации? Отсекают лишнее



Изначально ж/д поезда делали с зубчатыми колесами (1812) [9],

Поскольку беспокоились, что железо будет скользить по железу.

**Selgin, 1996:**

“...most people are much better at imagining catastrophic banking failures than they are at imagining how such failures might be avoided by an open, diverse, and non-hierarchical banking industry”.

- Из документа ЕВА следует, что Big Data невозможно использовать без регулирования.
- В ж/д поездах от зубьев отказались.
  - Кроме фуникулеров, где большой угол наклона.
  - В кризис регулирование может пригодится
    - Но стандарты ценнее нормативов.
- **Big Data тоже может использоваться без регулирования.**



## Выводы

---

1. Big Data **может** давать интересные идеи;
2. Но истинность таких идей не доказуема;
3. Эффективность от использования НЕ очевидна, НО
  - ❑ Де Голль: «**Не все, кто ищут находят; но те, кто находят, ищут**»
4. Регулирование лишь исказит условия конкуренции, создаст новые проблемы;
  - ❑ Всегда помните про **Эффект кобры** (достижение формальных, не реальных целей).
5. Те, кто считают, что нарушено личное пространство, не публикуйте частное.
  - ❑ Помните, что бесплатные сервисы должны окупаться (если Вы не знаете, как; знают другие).
6. Ответственность должна оставаться за теми, кто заинтересован:
  1. За компаниями – иначе уйдут клиенты
  2. За клиентами – что делать; что смотреть, где покупать, какую информацию оставлять и т.п.
  3. *Регулятор здесь лишний.*
    - Б. Педуро: «**Регулирование никогда не опередит инновации**»
7. **Востребовано НЕ регулирование Big Data, но знание/образования/отраслевые стандарты, вкл. проф. стандарты.**



## Использованные материалы

---

1. [https://www.esma.europa.eu/sites/default/files/library/jc-2016-86\\_discussion\\_paper\\_big\\_data.pdf](https://www.esma.europa.eu/sites/default/files/library/jc-2016-86_discussion_paper_big_data.pdf)
2. <https://www.wired.com/2009/02/wp-quant/>
3. <http://www.forbes.ru/tehnologii/337857-roznichnye-cifry-kak-tehnologii-bolshih-dannyh-prihodyat-v-oflayn-riteyl>
4. <http://www.ft.com/cms/s/2/927ca86e-d29b-11e2-88ed-00144feab7de.html#axzz43ZcF6llf>
5. <http://www.theguardian.com/news/datablog/2014/apr/22/how-much-is-personal-data-worth>
6. <http://www.telegraph.co.uk/technology/news/12012191/How-much-is-your-personal-data-worth.html>
7. <http://www.economist.com/node/13766329>
8. <http://www.economist.com/node/21711074/print>
9. [https://en.wikipedia.org/wiki/Salamanca\\_\(locomotive\)#/media/File:Blenkinsop%27s\\_rack\\_locomotive,\\_1812\\_\(British\\_Railway\\_Locomotives\\_1803-1853\).jpg](https://en.wikipedia.org/wiki/Salamanca_(locomotive)#/media/File:Blenkinsop%27s_rack_locomotive,_1812_(British_Railway_Locomotives_1803-1853).jpg)



# Благодарю за внимание!

**Пеникас Генрих, к.э.н.**

Доцент, старший научный сотрудник, ВШЭ  
Директор проектов, Казначейство, Сбербанк

E-mail: [penikas@hse.ru](mailto:penikas@hse.ru)

О банках: <http://basel.io/analytics/banx>

Семинар: <https://www.hse.ru/science/shadow/>

О себе: <https://www.hse.ru/staff/penikas>