

**Ломоносовские чтения-2023**

**11 апреля 2023 года**

# Догоняющее регулирование сферы больших данных

**Анастасия Моросанова, к.э.н.**

Кафедра конкурентной и промышленной политики  
Экономического факультета МГУ имени М.В. Ломоносова

Центр исследований конкуренции и экономического  
регулирования ИПЭИ РАНХиГС



## План доклада

1. Подходы к регулированию сферы больших данных (в том числе персональных)
2. Опыт Евросоюза – GDPR
3. Усиление регулирования в России
4. Секторальная специфика
5. Спрос на защиту: «парадокс конфиденциальности»

# Регулирование сферы больших данных

## Подходы к регулированию БД

	ЕС	США	Китай
<b>Защита данных</b>			
<b>Сбор и использование данных</b>	Полный охват	Частичный охват	Полный охват
<b>Пользовательские права</b>	Полный охват	Частичный охват	Частичный охват
<b>Обмен данными</b>			
<b>Подход</b>	Нормативный	Рыночный	Рыночный

Источник: Crisanto et al., *Big tech regulation: what is going on?*, 2021

## *General Data Protection Regulation (GDPR) - EU, 2018*

- Снизились венчурные инвестиции: количество сделок B2C упало на 17,6%, B2B – на 10,8%  
*(Jia et al., 2021)*
- Доходы электронной коммерции снизились на 13,3%  
*(Goldberg et al., 2020)*
- Количество новых приложений, выходящих на Google's app market упало на 45%  
*(Janßen et al., 2021)*
- Инновации стали больше «адаптивными», меньше стало «прорывных».  
*(Blind et al., 2022)*

## **В России:**

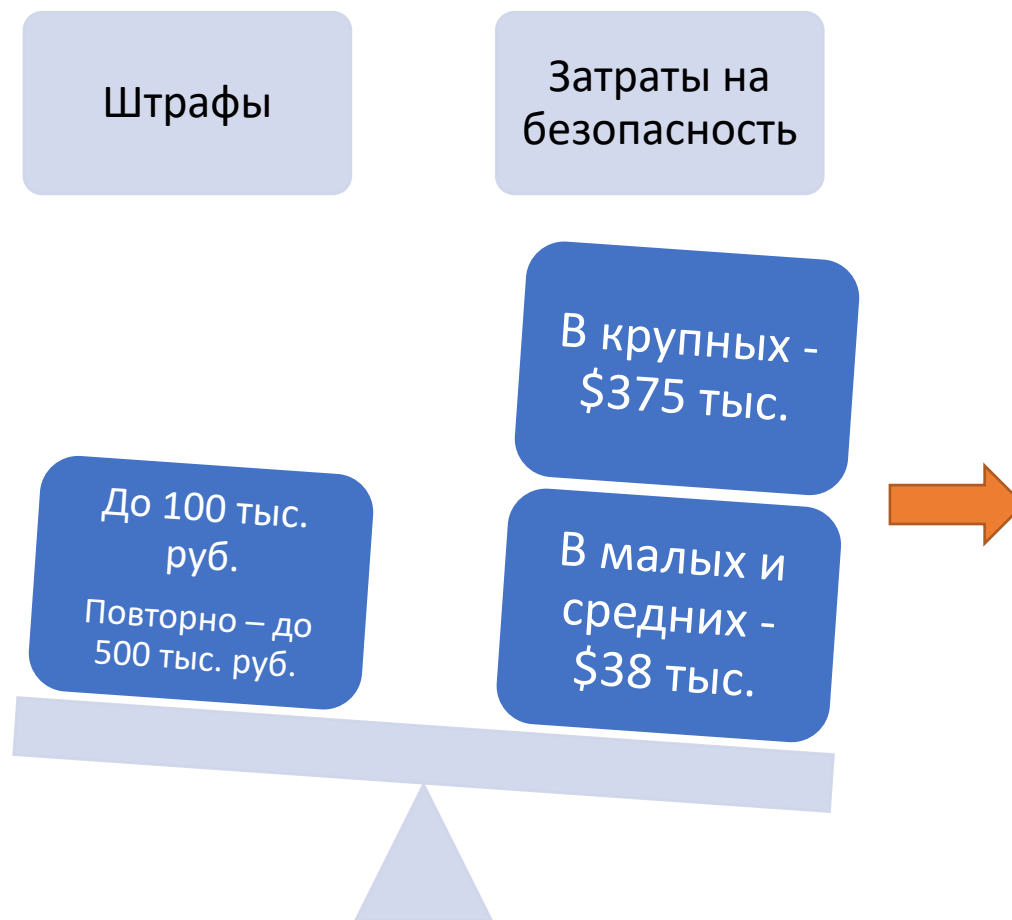
- Нет четкого следования какому-то из направлений регулирования\*
- Предпосылки к усилению контроля:
  - Усиление значения кибербезопасности
  - Тенденции регулирования в смежных сферах (*Закон о маркировке интернет-рекламы*)
  - ? Государственный оператор больших данных (*Минцифры, 2021*)

# Регулирование сферы больших данных в РФ: Федеральный закон «О персональных данных»

27.07.2006 N 152-ФЗ

- Нет четкого определения «персональных данных» (как оценивать cookies?)
- Размыто определение «Оператор персональных данных»: есть различия в стимулах между (1) первичным сбором данных, (2) продажей данных, (3) покупкой для аналитики
- Нет проработанного механизма возмещения ущерба пострадавшим пользователям

## Смещённые стимулы у операторов персональных данных



В 2022 году опубликовано:

- 2 126 095 255 записей данных
- 290 933 531 пользовательских данных
- 47 663 767 паролей (Лаборатория Касперского, 2023)

# Опыт ЕС: конфиденциальность vs. конкуренция

## Влияние регулирования

Затраты на внедрение защиты создают дополнительные барьеры для входа (или даже приводят к выходу)

Доверие на рынке возрастает к крупным платформам

Крупным платформам проще получить согласие пользователя

Крупные платформы неохотно делятся данными из-за рисков

*(Geradin, 2020; Campbell, 2015)*


## Особенности ЕС

Система «одного окна» (*нагрузка на регулятора Ирландии*)


Доминирующие платформы – не компании ЕС

- **Увеличение общей рыночной концентрации** на рынках веб-технологий, включая рынок больших данных (*Johnson, 2022; Marthews and Tucker, 2019; Peukert, 2022*)
- Особо уязвимые сферы: реклама, хостинг, анализ аудитории и социальные сети (*Johnson, 2022*)
- **Пострадали мелкие и средние фирмы** (*Geradin, 2021*), крупные платформы несут меньшие издержки (*Peukert, 2022*) или выигрывают (*Johnson, 2022*)
- Влияние на рынки **вне ЕС** (в том числе на США) (*Johnson, 2022; Peukert, 2022*)

# А у нас?



От усиления  
регулирования в  
сфере больших  
данных  
потенциально могут  
выиграть российские  
**крупные платформы**

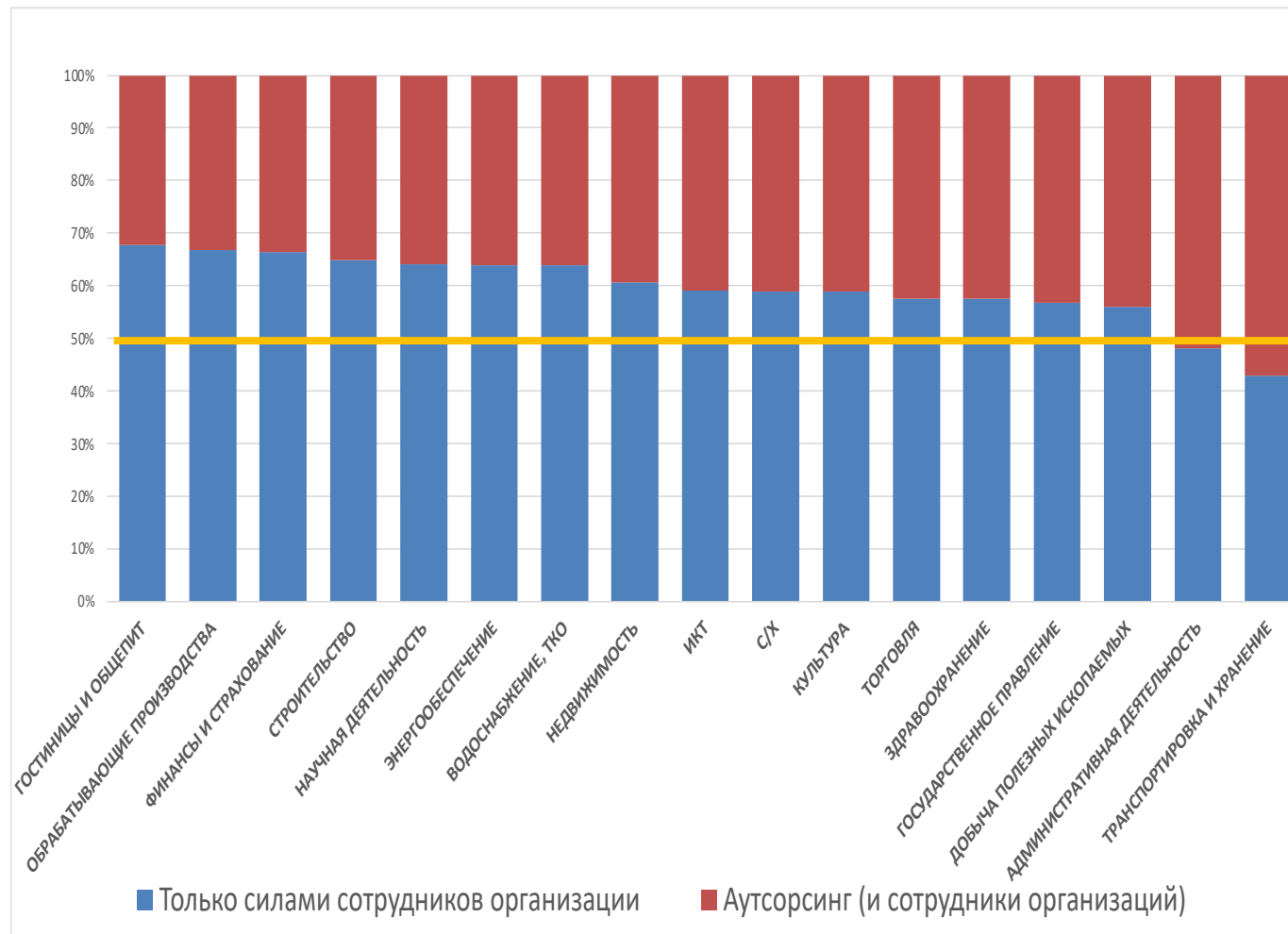
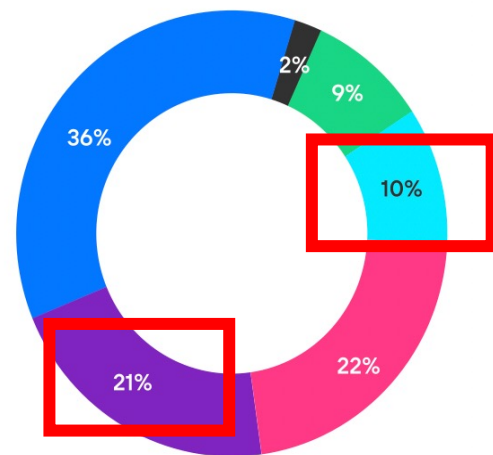


От усиления  
регулирования в  
сфере больших  
данных  
потенциально могут  
выиграть **российские**  
крупные платформы

# Компетенции персонала и анализ БД

Привлекаете ли вы внешних специалистов для внедрения проектов

- Да, привлекаем интеграторов для разработки проекта «под ключ»
- Нет, развиваем компетенции внутри компании или берем новых сотрудников в штат
- Да, привлекаем агентства для решения локальных задач
- Нет, у команды есть все компетенции
- Да, работаем с консалтинговыми компаниями
- Другое



Источник : Исследование VK Cloud и Arenadata «Технологии для работы с Big Data: готовность к использованию и основные барьеры», 2022

Источник: построено автором на основе данных Росстат (2021 г.)

# Регулирование сферы больших данных в РФ: ГОСТ

Год	РФ	ISO	Примечание
2021	ГОСТ Р «Информационные технологии. Большие данные. Обзор и словарь»	ISO/IEC 20546:2019	
2022	ГОСТ Р «Информационные технологии. Эталонная архитектура больших данных. Часть 2. Варианты использования и производные требования»	ISO/IEC TR 20547-2.	
2022	ГОСТ Р «Информационные технологии. Большие данные. Техническое задание. Требования к содержанию и оформлению»	---	<b>Отдельный раздел посвящен госзакупкам</b>
2023	ГОСТ Р «Информационные технологии. Искусственный интеллект. Структура управления процессами аналитики больших данных»	ISO/IEC 24668:2022	<b>Рассматриваются технологии до 2015 года Избыточен</b>
2023	ГОСТ Р «Информационные технологии (ИТ). Эталонная архитектура больших данных. Часть 1. Структура и процесс применения»	ISO/IEC TR 20547-1.	<b>Широкое трактование Не хватает отраслевой специфики</b>



# Есть ли различия в подходах к использованию БД?

	Экономические сферы	Значимые цели использования	Значимые источники БД	Роль безопасности
1.	<b>Финансы</b>	Продажи и маркетинг Производство	Сайт Учетные системы	Велика (внутреннее оборудование)
2.	<ul style="list-style-type: none"> <li>• <b>Государственное управление</b></li> <li>• <b>Здравоохранение</b></li> <li>• <b>Культура</b></li> </ul>	Производство Иное	Сайт Соц.сети Иное	Небольшое
3.	<ul style="list-style-type: none"> <li>• <b>Торговля</b></li> <li>• <b>Гостиницы</b></li> </ul>	Продажи и маркетинг	Учетные системы Сайт	Совершенно незначимо
4.	<ul style="list-style-type: none"> <li>• <b>Добыча</b></li> <li>• <b>Энергообеспечение</b></li> </ul>	Производство	Геолокация	Совершенно незначимо
5.	<ul style="list-style-type: none"> <li>• <b>Обрабатывающие производства</b></li> <li>• <b>Информация и связь</b></li> </ul>	Продажи и маркетинг Производство	Учетные системы	Незначимо
6*.	<p>Остальное:</p> <p>Водоснабжение, ТКО</p> <p>Административная деятельность</p> <p>Недвижимость</p> <p>Наука</p> <p>С/Х</p> <p>Строительство</p> <p>Транспортировка</p>	<p>Производство</p> <p>На втором месте – продажи и маркетинг</p>	Сайт	Незначимо

# Соотношение источников и цели использования БД: Безопасность

	Финансы	Государственное управление	Здравоохранение	Культура	Торговля	Гостиницы	Добыча	Электроснабжение	Производство	Информация и связь	Транспорт	Строительство	Водоснабжение	Недвижимость	Наука	Административное	С/Х
Данные между оборудованием	89,82%	25,81%	29,62%	27,30%	35,73%	40,28%	38,71%	29,03%	37,42%	19,89%	54,57%	41,44%	29,69%	38,26%	26,91%	33,87%	24,44%
Учетные системы предприятия	0,90%	9,73%	9,00%	10,48%	5,34%	12,50%	3,23%	12,90%	5,89%	32,39%	4,78%	11,60%	9,38%	9,40%	7,84%	7,53%	7,41%
Геолокация	3,53%	14,73%	21,80%	15,87%	43,09%	12,50%	35,48%	17,20%	30,57%	26,14%	15,00%	20,44%	20,31%	15,44%	23,31%	23,66%	24,44%
Сайт	0,52%	11,89%	12,09%	15,24%	2,76%	5,56%	3,23%	11,83%	3,82%	5,49%	6,30%	6,63%	10,94%	6,04%	10,17%	6,99%	8,15%
Операторы сотовой связи	0,85%	11,76%	9,48%	9,52%	6,72%	13,89%	5,38%	8,60%	12,42%	7,01%	9,13%	11,60%	10,94%	13,42%	8,05%	11,83%	17,04%
Социальные сети	0,99%	11,22%	8,53%	13,65%	2,39%	8,33%	10,75%	6,45%	4,46%	3,03%	4,57%	4,97%	6,25%	9,40%	16,95%	7,53%	8,89%
ДЗЗ	0,05%	5,14%	2,84%	2,22%	1,38%	2,78%	1,08%	6,45%	1,43%	1,52%	1,74%	2,21%	3,13%	2,01%	2,12%	3,23%	2,96%
Иное	3,35%	9,73%	6,64%	5,71%	2,58%	4,17%	2,15%	7,53%	3,98%	4,55%	3,91%	1,10%	9,38%	6,04%	4,66%	5,38%	6,67%

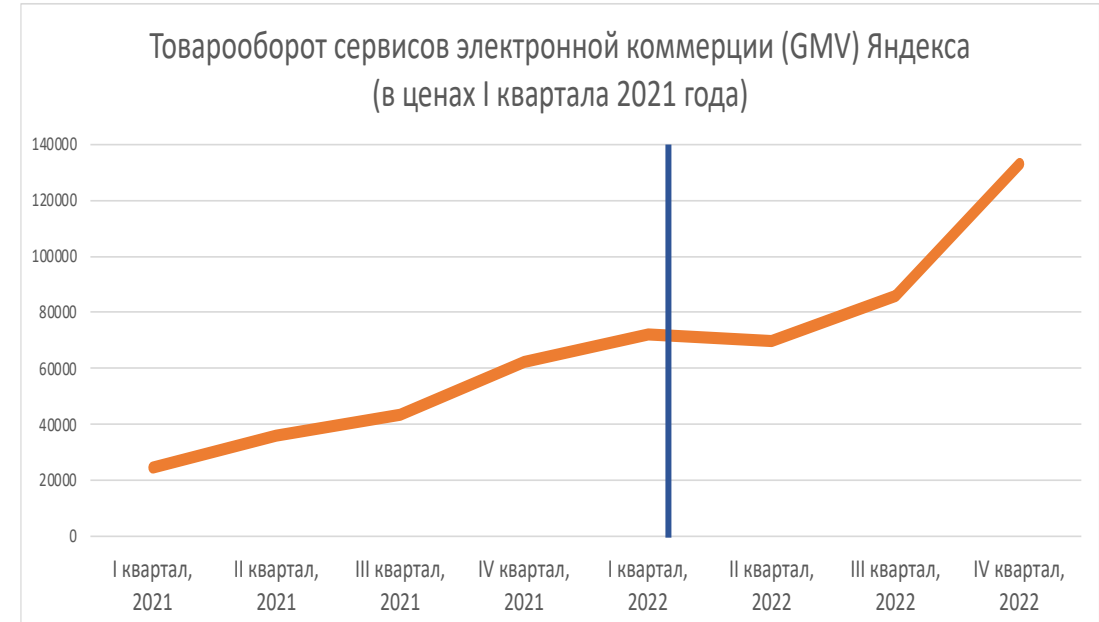
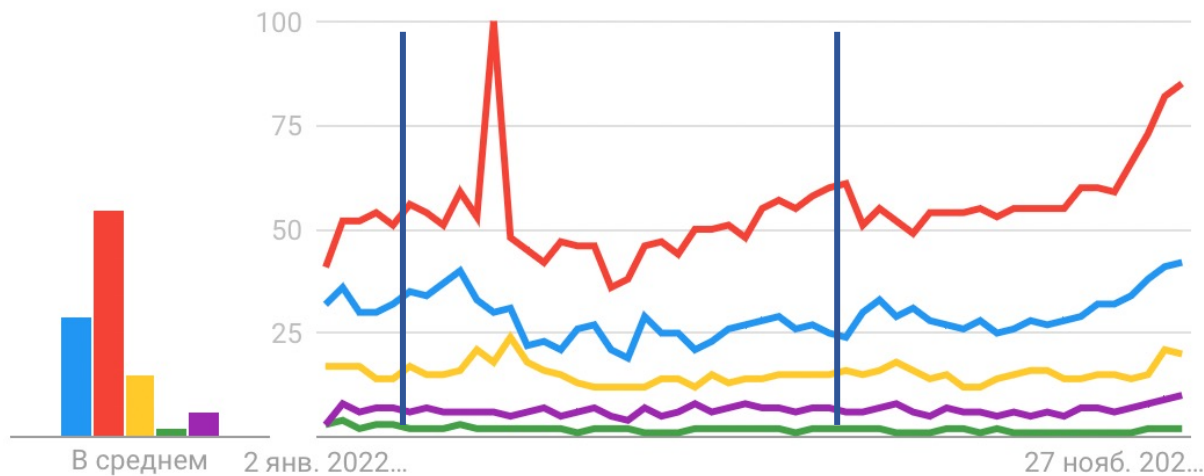
Источник: построено автором на основе данных Росстат (2021 г.)

# «Парадокс конфиденциальности»

Потребители выражают обеспокоенность насчет конфиденциальности своих данных, но это расходится с выявленными реальными предпочтениями

(Marthews, Tucker, 2019)

- сдэк отслеживание
- почта России отслеживание
- boxberry
- pony express
- достависта



# «Парадокс конфиденциальности»

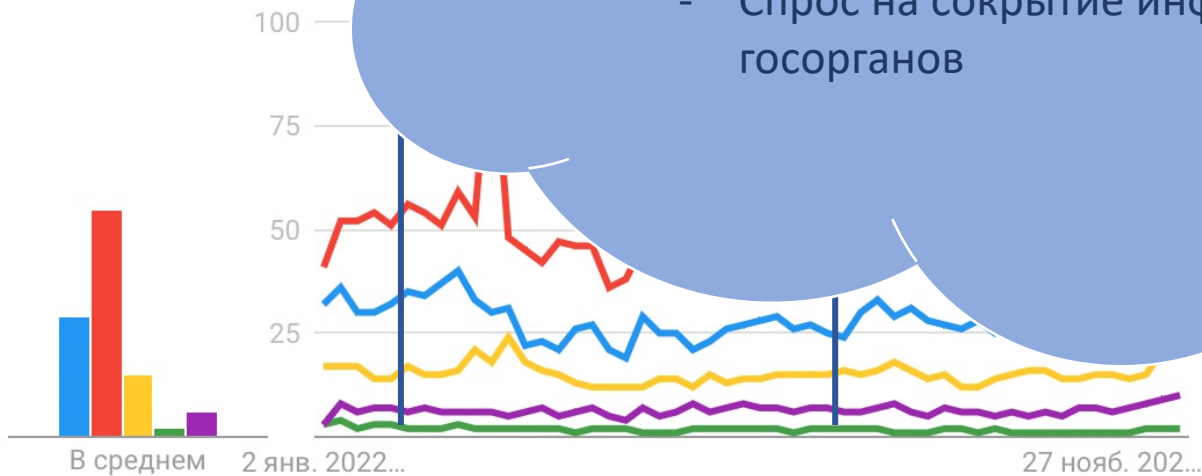
Потребители выражают беспокойство по поводу конфиденциальности, но это расходится с реальными предпочтениями.

## VPN

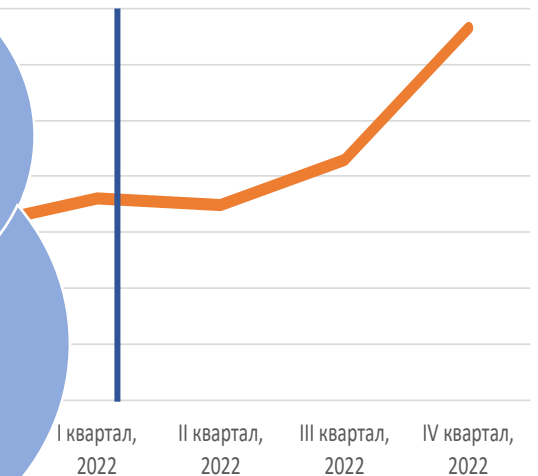
Причины:

- Доступ к ушедшим сервисам
- Спрос на сокрытие информации от госорганов

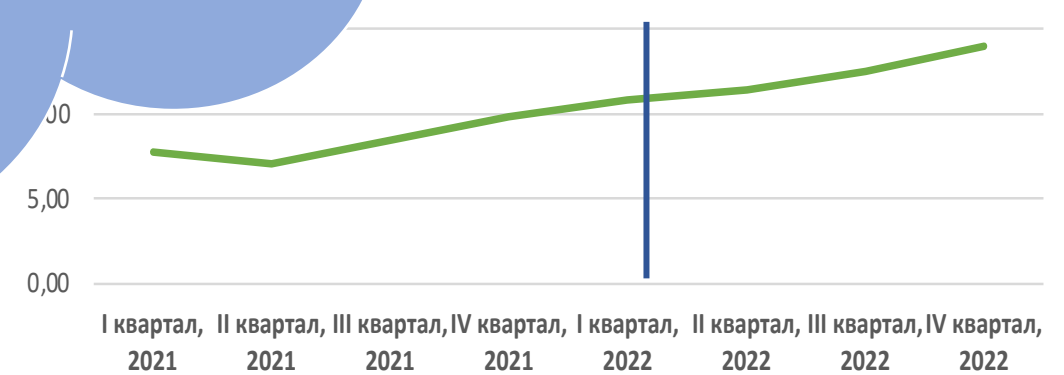
● сдэк отслеживан  
● pony express ● Д



Объем электронной коммерции (GMV) Яндекса (млрд руб.) в I кварталах 2021 и 2022 годов



Объем покупок на Яндекс Маркете (млрд руб.) в I кварталах 2021 и 2022 годов



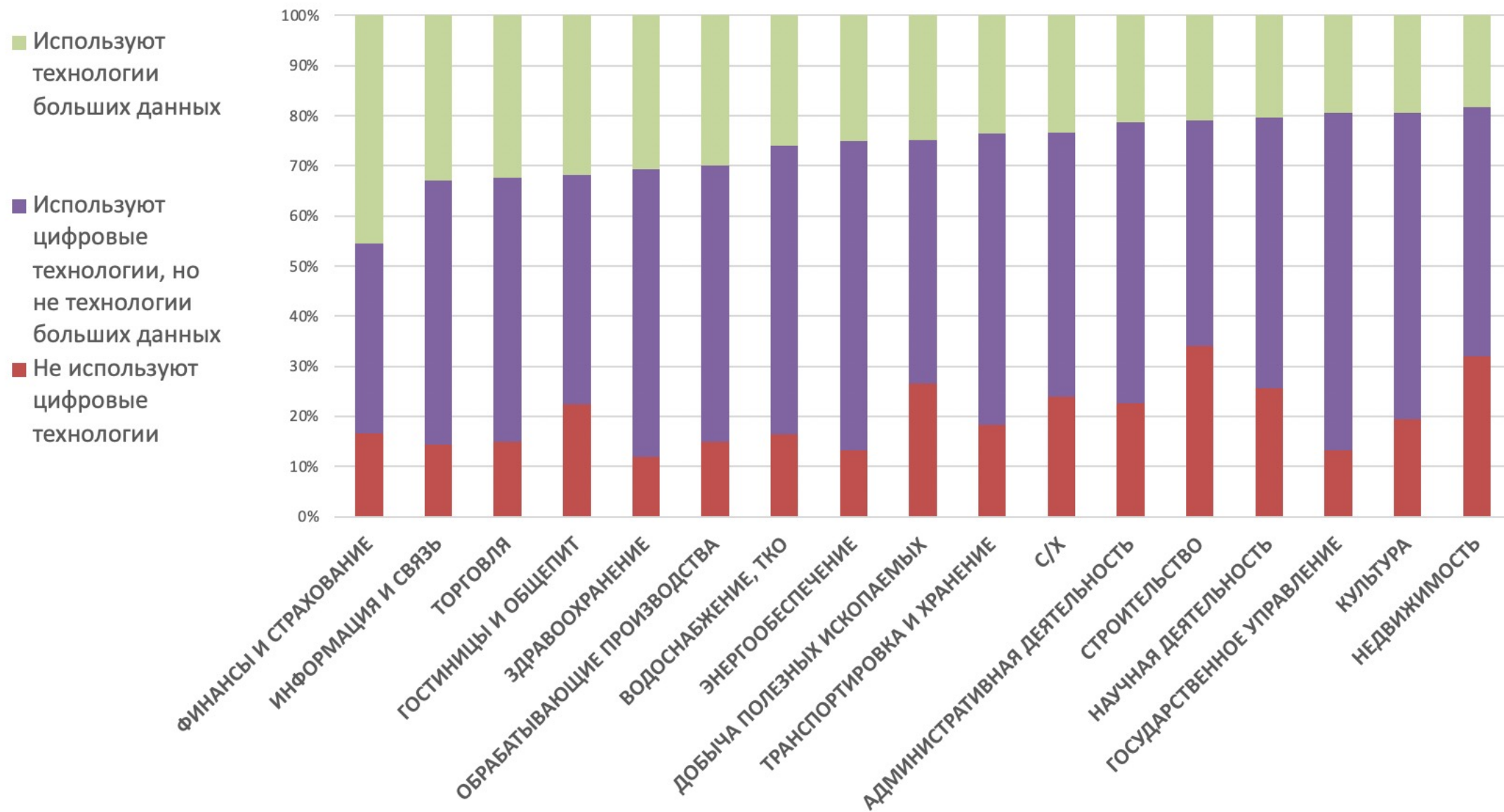


## Главные выводы

- Тенденция к **усилению ответственности** операторов персональных данных и увеличению значимости кибербезопасности
- Усиление контроля приведет к снижению инновационной активности и возможным экономическим потерям (опыт ЕС): на фоне санкций эффект может быть существеннее
- Возможность установления рамочных правил с учетом **отраслевой и страновой специфики**
- «Парадокс конфиденциальности» в России имеет свою специфику

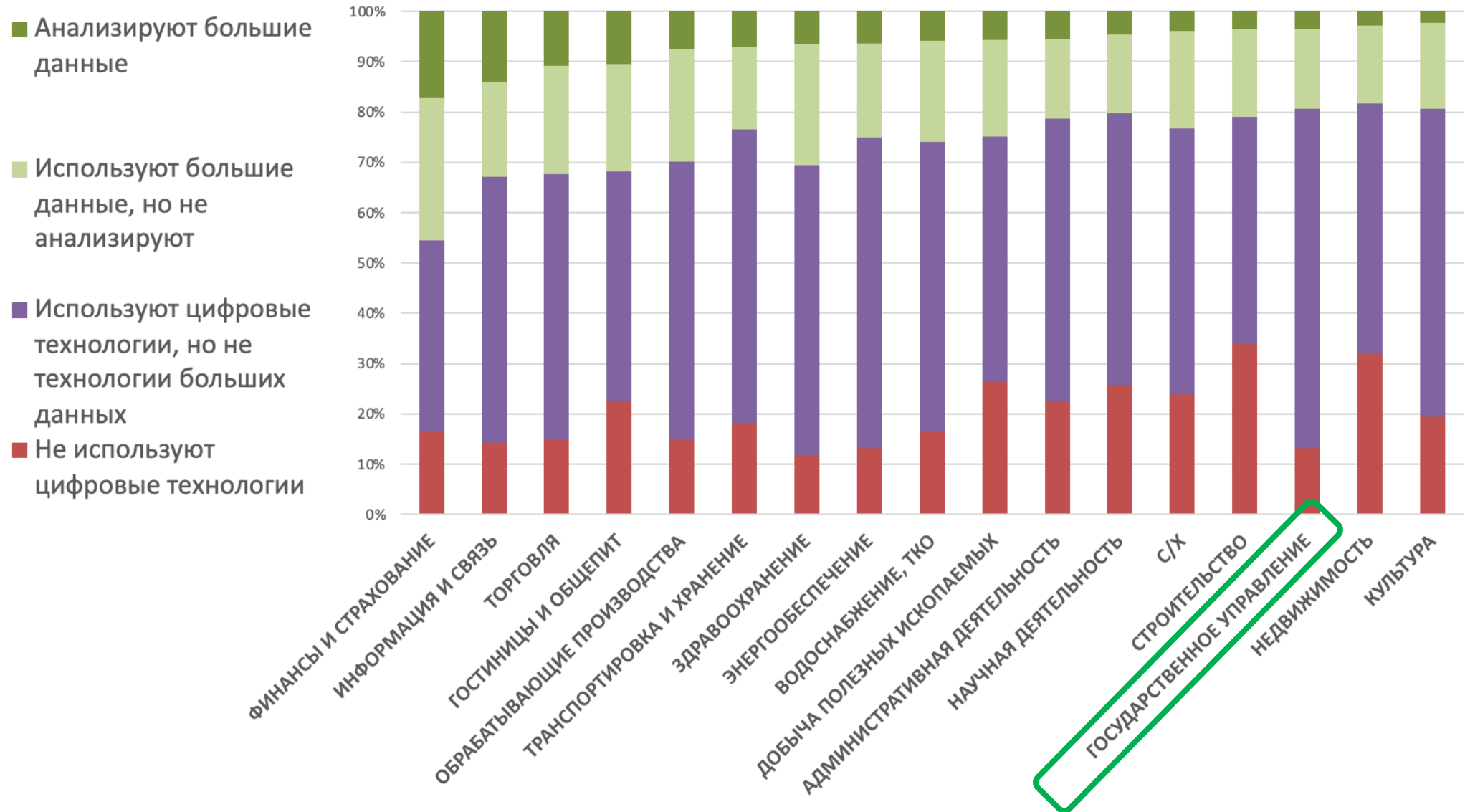
Спасибо за  
внимание!

# КТО ИСПОЛЬЗУЕТ БОЛЬШИЕ ДАННЫЕ?



Источник: построено автором на основе данных Росстат (2021 г.)

# А КТО АНАЛИЗИРУЕТ?



Источник: построено автором на основе данных Росстат (2021 г.)



# Новые вызовы: программное обеспечение

- Разработка отечественных аналогов

Категория ПО	Количество «ушедших» программ	Количество российских аналогов
ERP	34	9
CRM	66	11
Аналитические системы (BI)	68	17
Геоинформационные системы (ГИС)	71	24
СУБД для аналитики	2	1
СУБД и средства работы с данными	27	7

*Источник: Каталог Совместимости Российского ПО, [https://catalog.arppsoft.ru/replacement\\_list](https://catalog.arppsoft.ru/replacement_list)*

Лишь 20% категорий ПО, используемого компаниями и промпредприятиями, имеют «адекватные зрелые российские аналоги», **нужно 2-3 года на его доработку и внедрение** (источник: Минцифры, 2022)

- Open Source

*Но как же кибербезопасность?*

# Главные вопросы регулирования

