



ФЕДЕРАЛЬНАЯ СЛУЖБА
ГОСУДАРСТВЕННОЙ
СТАТИСТИКИ

ПОДХОДЫ К ИИ-ТРАНСФОРМАЦИИ ГОСУДАРСТВЕННОЙ СТАТИСТИКИ

Новиков Л.В.

03.06.2025

Заместитель руководителя Росстата



Леонид Новиков

Заместитель руководителя Федеральной службы государственной статистики Российской Федерации

Руководитель проекта по созданию Государственной информационной системы «Цифровая аналитическая платформа предоставления статистических данных»

Старший директор по финансовым технологиям НИУ ВШЭ, отвечал за цифровизацию бэк-офиса Вышки

Генеральный директор крупного ИТ-интегратора

Заместитель руководителя по производству в Управлении специальной связи по г. Москве и Московской области

Более 14 лет опыта цифровизации в федеральных учреждениях РФ

Организационные

Технологические

Внешние



Растущее количество альтернативных источников информации



Повышение требований к скорости принятия решений и, как следствие, актуальной информации



Противоречивость данных, предоставляемых в Правительство ведомствами



Рост скептицизма населения к официальным данным, требующий повышения прозрачности



Растущее количество административных и больших данных, рост значимости этих данных



Развитие цифрового государства и платформенных решений (например, ГосТех)



Развитие искусственного интеллекта и машинного обучения

Внутренние



Сложности в привлечении, адаптации и удержания кадров (в том числе из-за необходимости изучения большого количества специфических ИС)



Растущая стоимость владения комплексом систем Росстата



Неизбежность вывода из эксплуатации исторических систем



Дублирование назначения исторических систем Росстата



Стратегия развития Росстата
и государственной статистики до 2030 года



Центры компетенций

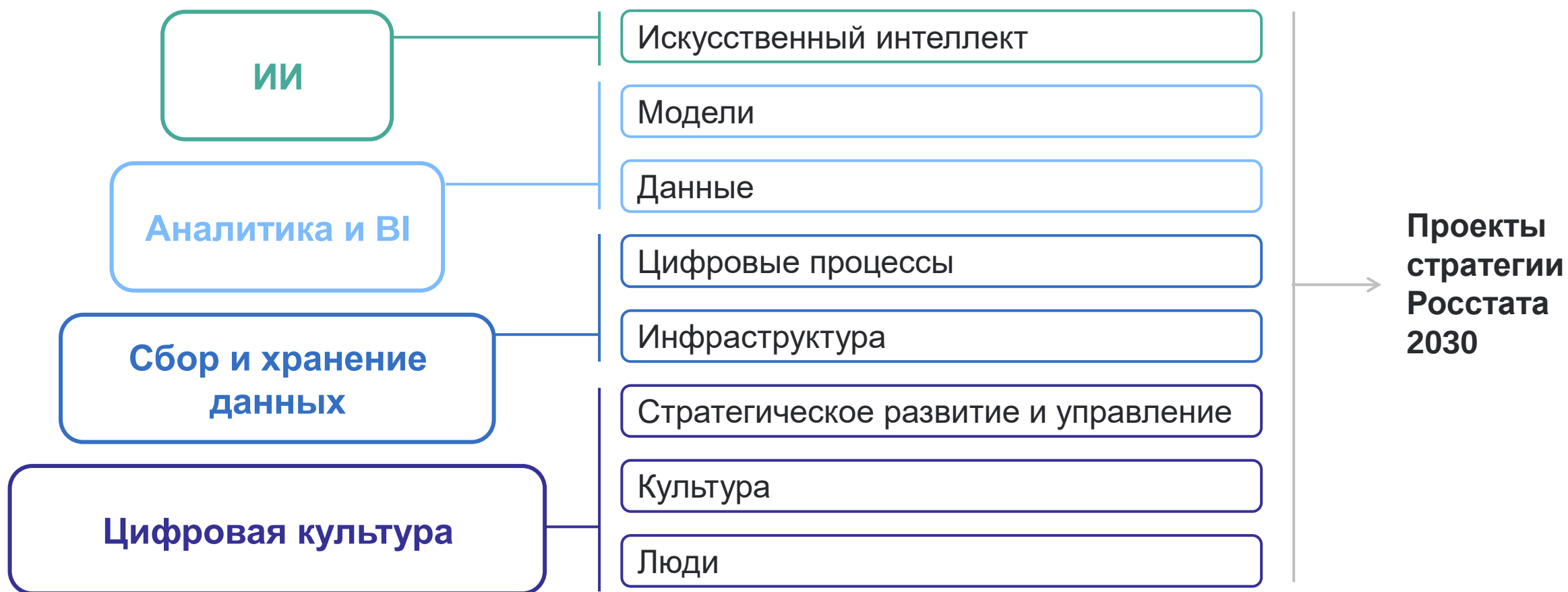
по опережающей статметодологии,
экспертизы НСИ и качества данных,
квалифицированного заказчика



ГИС ЦАП

как единый цифровой инструмент
Росстата на платформе Гостех

Пирамида ИИ-трансформации



Цель проведения диагностики – ответы на вопросы:

- Есть ли все необходимые условия для внедрения ИИ?
- Какие препятствия мешают масштабному внедрению ИИ?



Стратегическое развитие и управление

Является ли разработка и внедрение ИИ приоритетом?



Процессы

Определены ли процессы с максимальным эффектом от внедрения ИИ и готовы ли к такому внедрению?



Модели

Осуществляется ли разработка / поддержка / тиражирование собственных или используются/дорабатываются внешние модели?



Данные

Эффективна ли система сбора и обработки данных для использования ИИ?



Инфраструктура

Доступна ли необходимая ИТ-инфраструктура для создания. внедрения ИИ-решений?



Люди

Какова оценка уровня собственных компетенций для разработки и внедрения ИИ?



Культура

Как оценивается уровень готовности рабочей культуры для разработки и внедрения ИИ?

*Источники: анализ Sber AI

СТРУКТУРА ОЦЕНКИ ИИ-ЗРЕЛОСТИ*

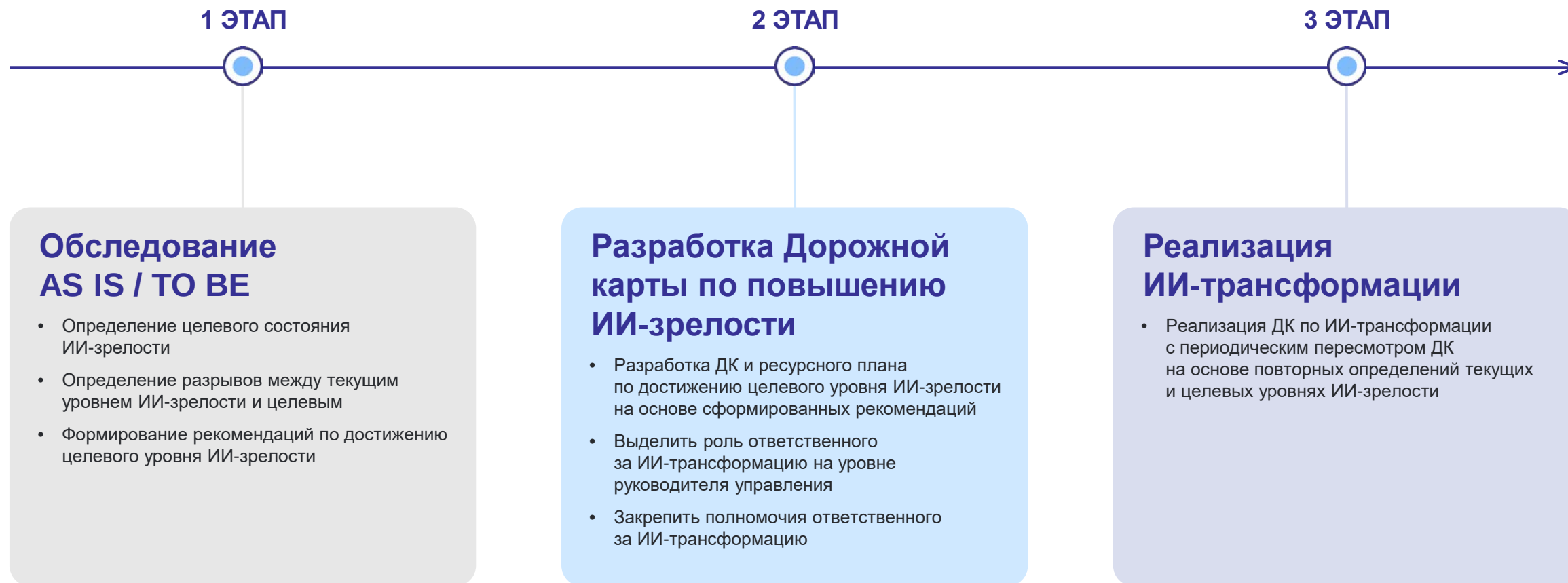
РОССТАТ

Направления

Параметры диагностики ИИ-зрелости

	Стратегическое развитие и управление	Стратегия ИИТ	Уровень целеполагания	Связь с ключевыми полномочиями ведомств	Обеспеченность финансовыми ресурсами
		Орган управления	Охват стратегии ИИТ	ДК внедрения лучших практик	Измерение эффекта моделей
	Процессы	Подход работы с ИИ-инициативами	Отбор процессов	Улучшение моделей	
		Автоматизация процессов	Охват процессов ИИ-решениями	Тиражирование моделей	
	Модели	Уровень собственной разработки	Контроль качества	Безопасность данных	
		Создание моделей	Внедрение моделей	Безопасность перс. данных	
	Данные	Стратегия по работе с данными	Управление качеством данных	Облачная инфраструктура	
		Экосистема данных	Партнерская экосистема	Доверие к сложным моделям	
	Инфраструктура	Вычислительные мощности при обучении моделей	Вычислительные мощности при исполнении моделей		
		Среды разработки при обучении моделей	Среды разработки при исполнении моделей		
	Люди	Общий уровень компетенций в ИИ	Система управления проектом		
		Качество технических специалистов	Структура команд DS		
	Культура	Принятие решений на основе данных	Метод проб и ошибок		
		Культура обмена опытом	Желание и умение экспериментировать		

*Источники: анализ Sber AI



ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Стратегическое
развитие
и управление

Ключевое наблюдение

- Отсутствие формального закрепления команды по ИИТ
- Отсутствие актуального стратегического документа по ИИ

Рекомендация

- Выделить роль ответственного за ИИ-трансформацию
- Закрепить полномочия ответственного за ИИ-трансформацию
- Разработать или включить в существующие стратегические документы ИИ-трансформации мероприятия по повышению ИИ-зрелости

Опыт Росстата

Стратегия до 2030 года

- Запущена **стратегия развития системы государственной статистики и Росстата до 2030 года**, предполагающую 18 направлений трансформации от кадрового обеспечения и культуры до системы экспериментальных расчетов (выявление потребностей и планирование, совершенствование методологии, сбора, обработки, распространения, цифровую трансформацию системы госстатистики)

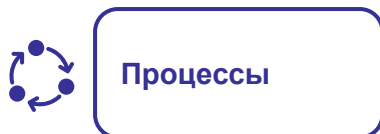
Проектный офис

- В Росстате создан отдельный **проектный офис по реализации стратегии**: закреплены роли команд и структура управления

ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Ключевое наблюдение

- Отсутствие отбора и приоритизации процессов для внедрения ИИ

Рекомендация

- Создать перечень процессов, требующих цифровизации
- Разработать методику приоритизации процессов для внедрения ИИ, исходя из их влияния на эффекты и реализуемости, включающей:
 - Выбор подрядчика / самостоятельной разработки моделей
 - Расчет стоимости и сроков
- Организовать ведение централизованного бэклога запросов и предложений по внедрению ИИ на уровне руководства

Опыт Росстата

Реестр процессов

- Создан **реестр процессов**, состоящий из 1085 процессов, включая 650 производственных, соединяющий федеральный план статработ, производственный план и публикуемые показатели статистики на портале официальной статинформации ЕМИСС

Реинжиниринг процессов

- Стратегия предусматривает **реинжиниринг всех производственных процессов** Росстат и субъектов официального статучета

Центр компетенций

- Создан **Центр компетенций** на базе подведомственного учреждения для **реинжиниринга процессов**

ТМПСИ

- Процессы характеризуются схожестью, поэтому моделируются через призму **ТМПСИ**, по группам процессов и отражением вариацией их имплементации структурными подразделениями

Новая технологическая основа статпроизводства

- Измеримость и прозрачность процессов заложены в новую технологическую основу статпроизводства – ГИС ЦАП

ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Модели

Ключевое наблюдение

- Множество систем для расчета разных статистических показателей, используется большое количество отдельных файлов вместо единой расчетной модели

Рекомендация

- Создать инструмент реализации единой расчетной модели на основе принципов международных моделей статистического производства (типовая модель производства статистической информации (GSBPM), типовая модель статистической информации (GSIM), общая архитектура статистического производства (CSPA), типовая модель работы статистических организаций (GAMSO))

Опыт Росстата

Единые модели

- В основу трансформации закладывается внедрение **единых моделей** для унификации, эффективности и скорости

Методология и экспериментальные расчеты

- Стратегия предусматривает направления по развитию **статистической методологии, системы экспериментальных расчетов**

Центр компетенций по методологии

- Для их реализации на базе подведомственного учреждения создан **Центр компетенций по опережающей статистической методологии**

DS-специалисты

- В Росстате действует **аналитическое управление с DS-специалистами** для построения моделей и реализацию методов машинного обучения

ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Данные

Ключевое наблюдение

- Разрозненность данных в разных системах и ведомствах
- Часть этих данных не имеет четкой методологии обработки и верификации
- Низкое использование ведомственных и больших данных
- Разная технологическая готовность субъектов официального статучета

Рекомендация

- Перейти на использование единых справочников и показателей для всех субъектов статучета на основе зрелых MDM-систем
- Предоставить готовый инструмент для ведомств с низкой технологической базой статпроизводства
- Внедрить инструменты управления данными для ИИ, включая обезличивание при доставке данных в среду разработки/исполнения

Опыт Росстата

Гармонизация НСИ

- Стратегия включает направление по **гармонизации НСИ** для перехода на эталонные справочники и показателей

Стандарт качества

- Стратегия включает направление по созданию **стандарта качества** данных для использования всеми субъектами статучета, включая учет требований для их использования ИИ-системами

АБД

- Стратегия включает направления по использованию **административных и больших данных (АБД)**

Центр компетенций по качеству данных

- Создан **Центр компетенций** по экспертизе нормативно-справочной информации и экспертизе данных

Единая информационная система НСИ

- Создается **единая система НСИ** и показателей для всех субъектов официального статучета (часть ведомств полностью переходит на решение Росстата по сбору и обработки данных, часть интегрируется со своими ГИС)

Пилоты по АБД

- **Пилоты по использованию АБД** вместо ручного сбора цен на лекарства и авиабилетов

ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Инфраструктура

Ключевое наблюдение

- Отсутствуют среды разработки и исполнения ИИ моделей

Рекомендация

- Выделить задачи требующие собственной разработки
- Для задач чувствительных к внешней инфраструктуре рекомендуется поэтапно рассмотреть:
 - сервисы типа «ML Space»
 - сервисы типа «ML Space Private»
 - собственную инфраструктуру для разработки и исполнения моделей

Опыт Росстата

ГИС ЦАП

- Росстат создает инфраструктуру статпроизводства, включая исполнение ИИ-моделей в виде федеральной системы **для всех субъектов официального статучета «Цифровая аналитическая платформа предоставления данных» (ГИС ЦАП)** на базе единой платформы Российской Федерации **ГосТех**

ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Люди

Ключевое наблюдение

- Не актуализируются знания сотрудников по ИИ

Рекомендация

- Организовать регулярное обучение в рамках федеральных учебных возможностей или создание собственных программ обучения на базе региональных ВУЗов для сотрудников, в т.ч. из других регионов
- Внедрить систему нематериальных поощрений сотрудников за самообразование и инициативность в сфере ИИ

Опыт Росстата

Кадровый потенциал

- Стратегия включает направления развития **кадрового потенциала системы государственной статистики**, в т.ч. включающего создание корпоративного университета с возможностью развития ИИ-компетенций

Организации высшего образования

- Росстат активно сотрудничает с **образовательными организациями высшего образования**. В настоящее время 142 «дорожные карты» по развитию взаимодействия Росстата с образовательными организациями высшего образования утверждены в 53 субъектах Российской Федерации

Бесплатное ДПО

- Сотрудникам доступны **бесплатные программы ДПО** с развитием компетенций в области ИИ

ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Культура

Ключевое наблюдение

- Отсутствует процесс обмена опытом и знаниями между территориальными органами и ведомствами (субъектами официального статучета)
- Высокая разница в опыте между руководящим составом

Рекомендация

- Выстроить регулярный процесс обмена опытом и знаниями между командами
- Организовать регулярные двусторонние визиты в другие ведомства и территориальные органы
- Организовать профильную конференции по ИИ в центральном аппарате
- «Ручное сопровождение» со стороны руководства территориальных органов и субъектов официального статучета, с опытом внедрения ИИ, при внедрении первых ИИ-решений для «новичков»

Опыт Росстата

Для обмена опытом в Росстате:

Коллегия Росстата

- Проводится ежегодная **коллегия Росстата** с участием всех территориальных органов

Командировки в ТОГС

- Ежегодный план **командировок для поездок в территориальные органы** по обмену опытом

ПРИМЕРЫ РЕКОМЕНДАЦИЙ ПО ИИ-ТРАНСФОРМАЦИИ НА ОСНОВЕ НАБЛЮДЕНИЙ

РОССТАТ

Направления



Культура

Ключевое наблюдение

- Нетерпимость к ошибкам: барьер в проведении экспериментальных исследований – риск недостижения результатов при использовании выделенных бюджетных средств

Рекомендация

- Выделить «венчурный» бюджет для пилотов
- Ориентировать региональные ВУЗы на организацию процесса проведения исследований и разработки решений в направлении потребностей региона
- Проработать возможность стимулирования привлечения в регион команд и компаний, ведущих прикладные исследования в области применения ИИ в госуправлении и приоритетных для региона отраслях

Опыт Росстата

Лидеры Росстата

- Ежегодно проходит конкурс «**Лидеры Росстата**» для поиска новых решений при реализации стратегии и поиска сотрудников с высоким уровнем лидерских качеств

Мастерская ценностей

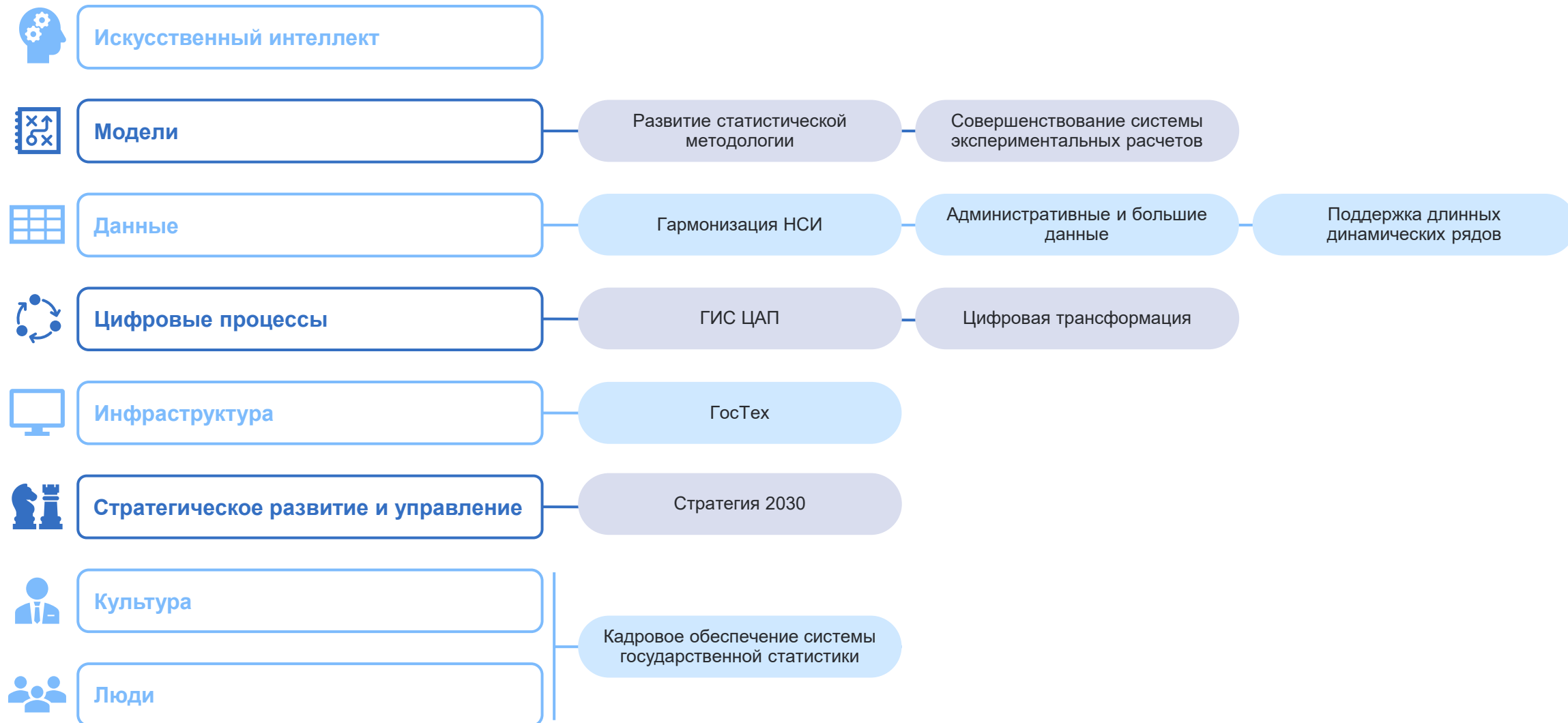
- Проводится ежегодный проект «**Мастерская ценностей**», направленный на укрепление корпоративной культуры и повышение вовлеченности. Это сочетание конкурсного формата, обучения и практической работе в командах.

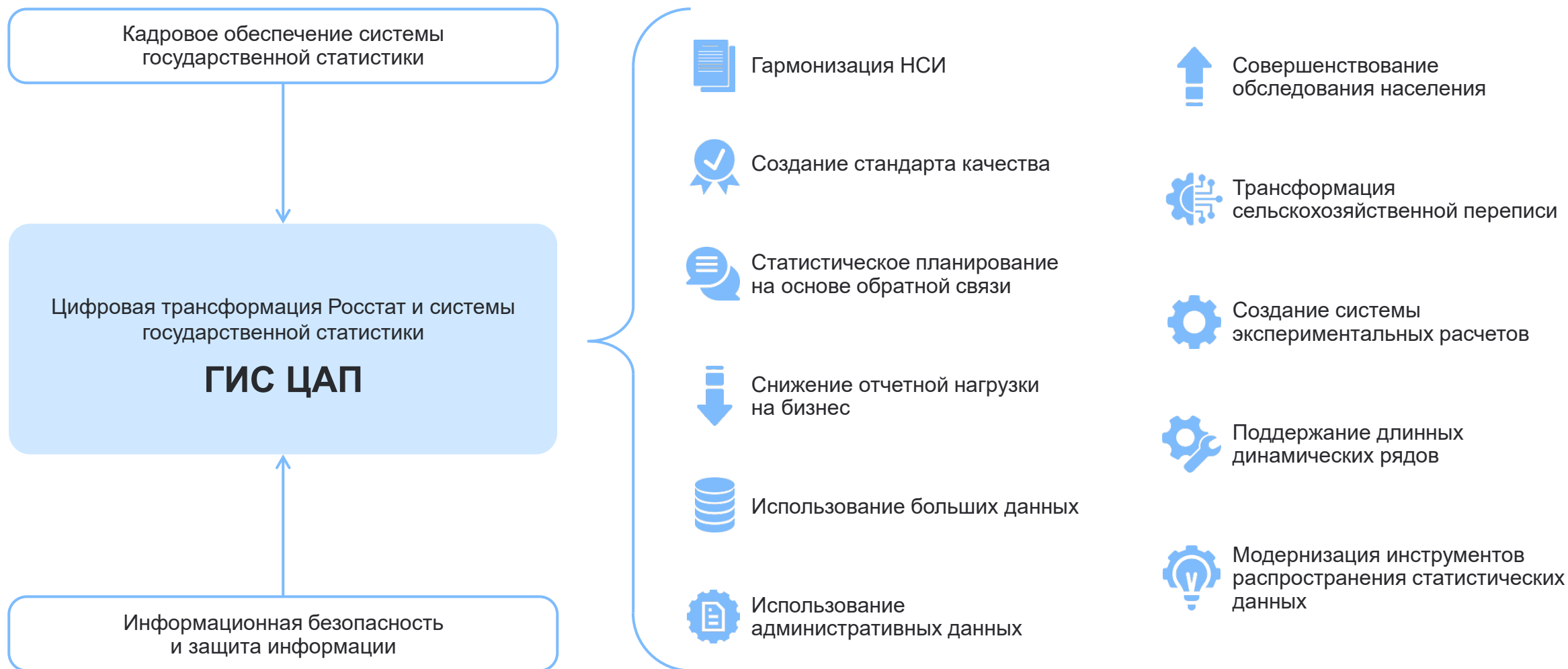
Кубок Росстата

- Проводится ежегодный **кубок Росстата** – чемпионат по стратегии и управлению

СТРАТЕГИЯ СТАВИТ ЗАДАЧИ ПО РЕАЛИЗАЦИИ ВСЕХ КОМПОНЕНТОВ МОДЕЛИ ИИ-ТРАНСФОРМАЦИИ

РОССТАТ





- 1 Государственные услуги**
Прикладные сервисы поддержки взаимодействия с внешними пользователями
- 2 Государственные данные**
Прикладные сервисы обработки госданных
- 3 Ведомственные процессы**
Прикладные сервисы управления бизнес-процессами и обеспечения деятельности
- 4 Технологические сервисы**
Инструменты разработки, интеграции приложений, работа с СУБД
- 5 Виртуализация**
Единая среда динамического управления инфраструктурой
- 6 ИТ-инфраструктура**
Оборудование, каналы связи, средства защиты информации

- 7 Информационная безопасность**
Безопасная разработка.
Компоненты платформы аттестованы ФСТЭК и ФСБ России
- 8 Госмаркет**
Маркетплейс цифровых продуктов для госорганов
- 9 Управление платформой**
Инструменты управления разработкой, тестированием и разворачиванием ГИС



ГИС ЦАП на ГосТех

единая платформа статпроизводства для всех субъектов официального статучета, а также сбора и обработки административных («ведомственных») данных

Цель

переход к новой структурной и функциональной модели статпроизводства с принципами:

-  Единого информационного пространства данных, НСИ и обеспечения сопоставимости
-  Принципиального нового качества данных, соответствие стандарту качества
-  Прослеживаемости и взаимосвязи понятий и объектов статучета
-  Однократного предоставления первичных данных при многократном переиспользовании
-  Прослеживаемости сроков и процессов статпроизводства
-  Использования административных и больших данных вместо сбора

Назначение

автоматизация всего **жизненного цикла статпроизводства** согласно ТМПСИ, **постепенная замена** всех **57 систем ИВС** Росстата и замена/дополнение систем субъектов ОСУ

ГИС ЦАП

ключевой инструмент реализации Стратегии развития Росстата и государственной статистики до 2030 года



Политическая цель –
«единая точка правды»

Почему Росстат? – «Цифровой перекресток» государственных данных

- 1 Самый широкий охват гос. данных
- 2 Наличие большого объема специализированных человеческих и технических ресурсов для работы с данными
- 3 Наличие зрелой методологии обработки данных
- 4 Наличие выстроенных контактов со всеми ФОИВ

РОССТАТ

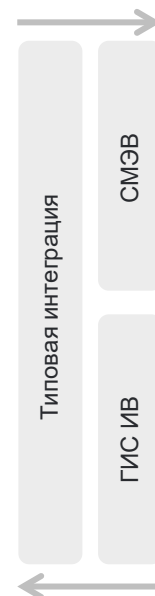
Центр компетенций по единой НСИ, стандарту качества данных

Оператор ГИС ЦАП

ГИС ЦАП

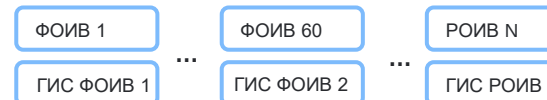


Обмен данными

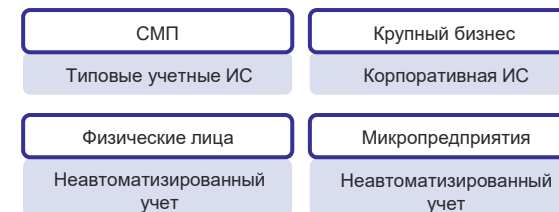


ОИВ

Субъекты ведомственного учета и операторы адм. данных



Субъекты официального статистического учета



Для государства:

- 1 Снижение стоимости производства официальной и ведомственной статистики (снижение расхода бюджетных средств)
- 2 Оперативность и качество данных
- 3 Улучшение качества решений за счет возможности опоры на оперативные данные

Для ОИВ:

- 1 Готовые инструменты и процессы статпроизводства
- 2 Снижение стоимости процессов статпроизводства
- 3 Сокращение расходов на поддержку собственных систем сбора отчетности

Для бизнеса:

- 1 Снижение отчетной нагрузки за счет переиспользования данных ведомствами
- 2 Расширение возможности анализа и построение собственной аналитики за счет доступа к микроданным
- 3 Прямая загрузка данных в корпоративные системы поддержки принятия решений

Для Росстата:

- 1 Возможность реализации задач стратегии
- 2 Удобство работы в едином информационном пространстве
- 3 Получение более совершенных средств статпроизводства
- 4 Снижение трудоемкости процессов
- 5 Сокращение расходов на поддержку систем

КОНЦЕПЦИЯ ИСПОЛЬЗОВАНИЯ БОЛЬШИХ ДАННЫХ В СТАТИСТИКЕ

РОССТАТ



ПРЕИМУЩЕСТВА

- Меньше времени на подготовку данных
- Гораздо более подробная статистика
- Повышенное качество данных
- Снижение нагрузки на отчетность



ИСТОЧНИКИ ДАННЫХ

- Сотовые операторы
- Экосистемы (Яндекс, Mail.ru, СБЕР)
- Открытые библиотеки (OpenStreetMap, АИС «Реформа ЖКХ» и др.)
- ККТ
- Данные спутниковых снимков
- Данные РЖД и авиакомпаний

ПРИОРИТЕТНЫЕ ОБЛАСТИ



Статистика потребительских цен

- Статистическая практика
- Обследования
- Большие данные
- Административные данные



Статистика туризма

- Численность туристов
- Количество поездок
- Продолжительность пребывания
- Количество посетителей



Статистика розничной торговли и услуг

- Данные ККТ (ФНС России)
- Данные системы маркировки



Статистика населения

- Постоянное население
- Временное население
- Потоки маятниковых трудовых мигрантов



Статистика сельского хозяйства



ЦЕЛЬ

- Автоматизация процесса сбора цен
- Оперативность формирования показателей



ПРОБЛЕМЫ

- Полная автоматизация мониторинга цен
- Отсутствие единой маркировки товаров и услуг



РЕАЛИЗОВАНО

- Разработка новой методологии ИПЦ
- Разработка системы приема и обработки «Больших данных» в рамках ИПЦ
- Веб-скрейпинг
- Нейросеть для идентификации товаров
- Экспериментальный расчет по новой методологии



ПРЕИМУЩЕСТВА

- Точность наблюдений
- Репрезентативность показателей
- Актуализация системы весов для расчета ИПЦ
- Увеличение количества товаров (услуг) представителей и ценовых котировок



ЦЕЛЬ

Отслеживание изменения стоимости небольшой группы схожих товаров вместо отслеживания отдельного конкретного товара



ИСТОЧНИКИ ДАННЫХ

Данные чеков
контрольно-кассовой
техники

~ **230 Гб**
ежедневно



ДЕЙСТВИЯ:

Предобработка данных

с целью ускорения и упрощения дальнейших операций

Формирование подвыборок

по ключевым словам

Обработка подвыборки

используется модель CountVectorizer, по результатам анализа которой получают вектора из названий позиций в чеках. Далее вектора кластеризуются.

Модель CountVectorizer

Тест

‘Картофель свежий’

‘Картофель мытый’

‘Морковь свежая’

Формирование векторов

Хеш-функция f

$f(\text{'картофель'})$:1
 $f(\text{'морковь'})$:2
 $f(\text{'свежий'})$:3
 $f(\text{'мытый'})$:4

1	2	3	4
1	0	1	0
1	0	0	1
0	1	1	0



**Спасибо
за внимание**