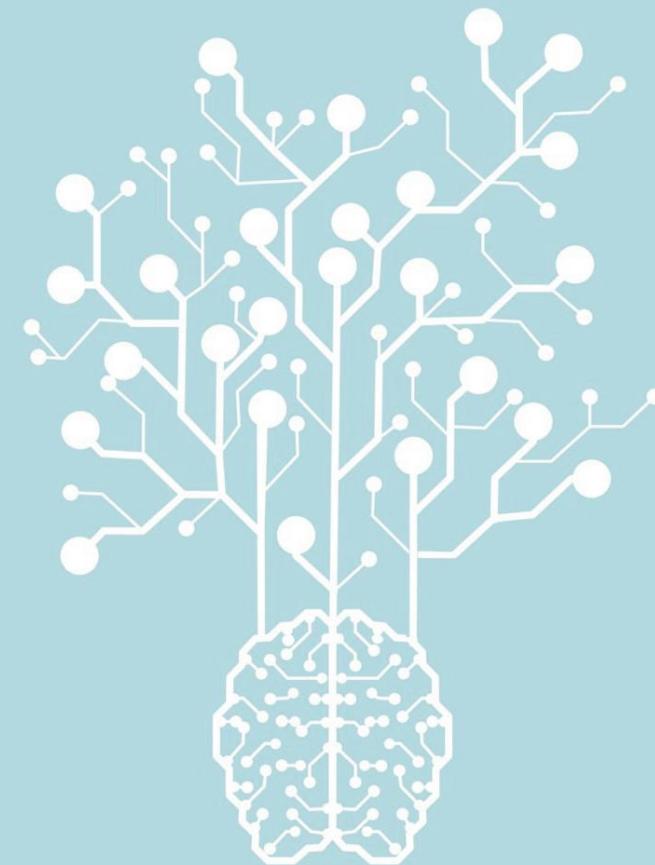


Разработка приложения диагностики цифровых компетенций

Ефимов Иван

Аспирант



ЦЕЛИ ИССЛЕДОВАНИЯ И РЕЗУЛЬТАТЫ

Цели исследования

- ▶ Разработать приложение для диагностики цифровых компетенций.
- ▶ Охарактеризовать уровень цифровой грамотности, используя анализ данных, собранных на основе тестирования.

Задачи исследования

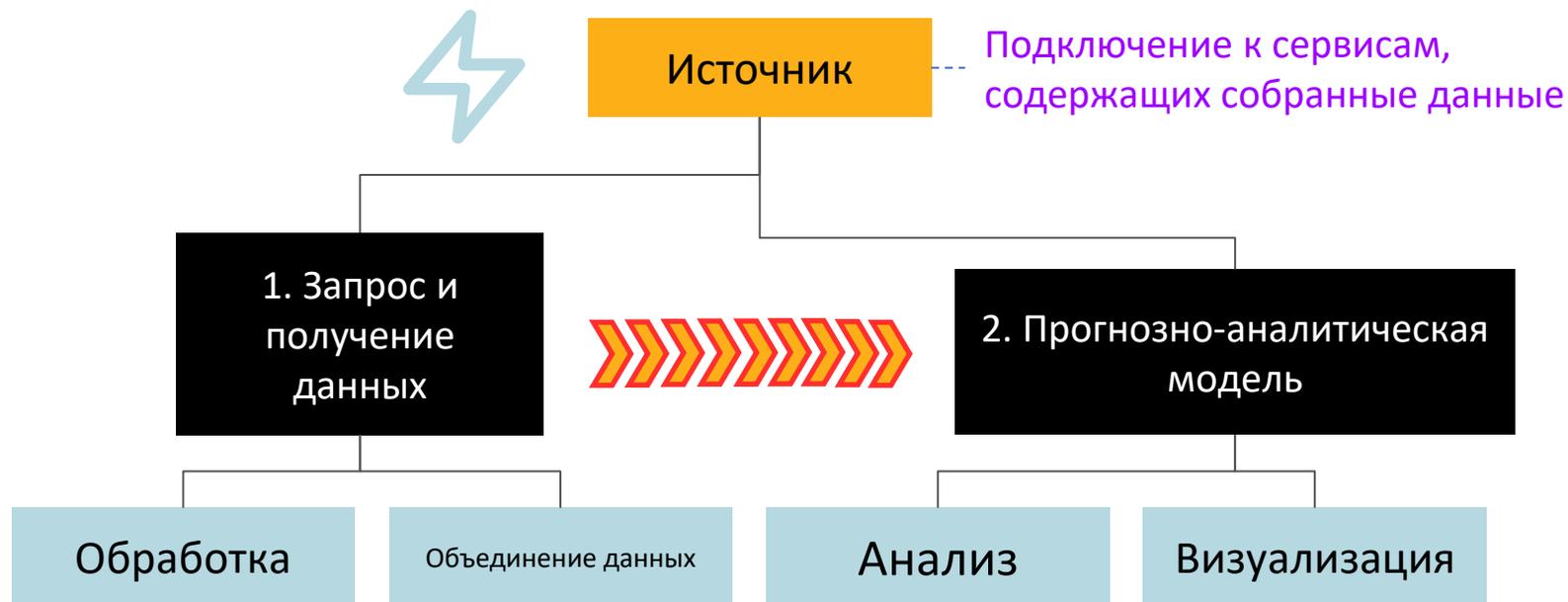
- ▶ Охарактеризовать этапы анализа данных
- ▶ Реализовать приложение для тестирования и сбора данных для обработки
- ▶ Создать базу данных и выполнить анализ
- ▶ Визуализировать полученные результаты



ОБЩИЙ ПРОЦЕСС АНАЛИЗА ДАННЫХ

Помимо функции опросника, приложение направлено на сбор не конфиденциальной информации, анализ которой позволяет определить неравенство и уровень цифровизации регионов.

При этом общая схема анализа включает механизм сбора, обработки и анализа данных.



АЛГОРИТМ ПРОГРАММЫ

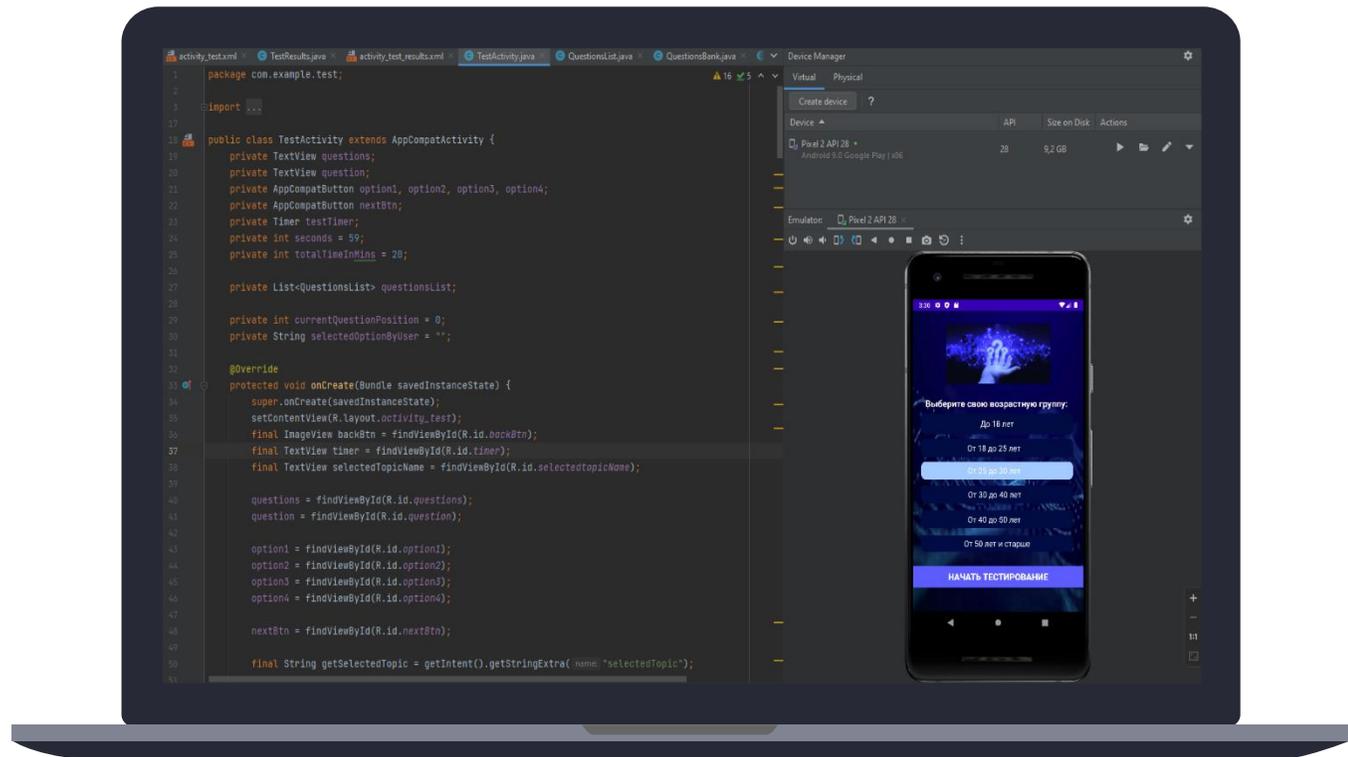


Программа создана на самой распространенной операционной системе – Android.

При этом для создания использовалось свободное ПО – Android studio.

Приложение доступно для 99,9% устройств, так как имеет минимально допустимую версию андроид – 4.4.

В алгоритм заложены несколько групп для опроса разных категорий населения.



ИНТЕРФЕЙС ПРИЛОЖЕНИЯ



- Программа имеет простой и понятный интерфейс
- Возможность локализации на различные языки
- Малый объем (3,80 МБ)
- Встроена система навигации и подсказок действий
- Группы вопросов, определяющие уровень цифровых компетенций различных групп населения

СОЗДАНИЕ БАЗЫ ДАННЫХ И АНАЛИЗ



Первый этап

После установки приложения-опросника, клиент проходит тестирование, выбирая нужную категорию под себя (возрастная группировка). По результатам прохождения тестирования результаты выдаются пользователю для ознакомления.

Цифровая грамотность для разных групп населения оценивается исходя из набора отдельно составленных вопросов, отвечая на которые и формируется исходная информация.



Второй этап

Результаты тестирования могут быть отправлены из приложения на прописанный хост-сервер или (доступный, не требующий покупки хоста вариант) отдельный созданный облачный диск (Google).

Собранные данные включают в себя IP-адрес, позволяющий определить регион страны и отсортировать результаты по регионам, ответы в виде набора цифр.



АНАЛИЗ ДАННЫХ, МЕТОДИКА РЕЙТИНГОВОЙ ОЦЕНКИ

		Indicators for 85 regions of the country (for 2021*)				Forecast by models		
Group	Factors	1998	1999	...	2020	2021F	...	2030F
(1) Economy [ec]	1 Gross regional product (in current prices, million rubles);	ec_{1p_0}	ec_{1p_1}	...	$ec_{1p_{20}}$	$ec_{1p_{21}}$...	$ec_{1p_{32}}$
	2 Investments in fixed assets (million rubles)	ec_{2p_0}	ec_{2p_1}	...	$ec_{2p_{20}}$	$ec_{2p_{21}}$...	$ec_{2p_{32}}$
	3 Advanced manufacturing technologies used (units)	ec_{3p_0}	ec_{3p_1}	...	$ec_{3p_{20}}$	$ec_{3p_{21}}$...	$ec_{3p_{32}}$
	4 Average annual number of people employed in the economy (thousand people)	ec_{4p_0}	ec_{4p_1}	...	$ec_{4p_{20}}$	$ec_{4p_{21}}$...	$ec_{4p_{32}}$
(2) Standard of living [sl]	1 Average per capita income of the population (rubles/month)	sl_{1p_0}	sl_{1p_1}	...	$sl_{1p_{20}}$	$sl_{1p_{21}}$...	$sl_{1p_{32}}$
	2 Residential area per inhabitant (m ²)	sl_{2p_0}	sl_{2p_1}	...	$sl_{2p_{20}}$	$sl_{2p_{21}}$...	$sl_{2p_{32}}$
	3 Consumer price index	sl_{3p_0}	sl_{3p_1}	...	$sl_{3p_{20}}$	$sl_{3p_{21}}$...	$sl_{3p_{32}}$
	4 Volume of communication services to the population (million rubles)	sl_{4p_0}	sl_{4p_1}	...	$sl_{4p_{20}}$	$sl_{4p_{21}}$...	$sl_{4p_{32}}$
(3) Medicine [med]	1 Capacity of medical outpatient clinics (thousand visits per shift)	med_{1p_0}	med_{1p_1}	...	$med_{1p_{20}}$	$med_{1p_{21}}$...	$med_{1p_{32}}$
	2 Number of hospital beds (thousand units)	med_{2p_0}	med_{2p_1}	...	$med_{2p_{20}}$	$med_{2p_{21}}$...	$med_{2p_{32}}$
	3 Number of doctors of all specialties (thousand people)	med_{3p_0}	med_{3p_1}	...	$med_{3p_{20}}$	$med_{3p_{21}}$...	$med_{3p_{32}}$
	4 Incidence per 1000 population	med_{4p_0}	med_{4p_1}	...	$med_{4p_{20}}$	$med_{4p_{21}}$...	$med_{4p_{32}}$
(4) Education and culture [eac]	1 Number of reported crimes per 100,000 population	eac_{1p_0}	eac_{1p_1}	...	$eac_{1p_{20}}$	$eac_{1p_{21}}$...	$eac_{1p_{32}}$
	2 Library fund of public libraries per 1000 population (copies)	eac_{2p_0}	eac_{2p_1}	...	$eac_{2p_{20}}$	$eac_{2p_{21}}$...	$eac_{2p_{32}}$
	3 Enrollment of children in preschool educational institutions (in % of the number of children of this age)	eac_{3p_0}	eac_{3p_1}	...	$eac_{3p_{20}}$	$eac_{3p_{21}}$...	$eac_{3p_{32}}$
	4 Number of students of state and municipal secondary specialized educational institutions per 10,000 population (persons)	eac_{4p_0}	eac_{4p_1}	...	$eac_{4p_{20}}$	$eac_{4p_{21}}$...	$eac_{4p_{32}}$
(5) Ecology [eco]	5 Emissions of pollutants into the atmospheric air from stationary sources (thousand tons)	eco_{5p_0}	eco_{5p_1}	...	$eco_{5p_{20}}$	$eco_{5p_{21}}$...	$eco_{5p_{32}}$
	6 Reforestation (thousand hectares)	eco_{6p_0}	eco_{6p_1}	...	$eco_{6p_{20}}$	$eco_{6p_{21}}$...	$eco_{6p_{32}}$
	7 Fresh water use (million m ³)	eco_{7p_0}	eco_{7p_1}	...	$eco_{7p_{20}}$	$eco_{7p_{21}}$...	$eco_{7p_{32}}$
	8 Crude mortality rate (number of deaths per 1000 people)	eco_{8p_0}	eco_{8p_1}	...	$eco_{8p_{20}}$	$eco_{8p_{21}}$...	$eco_{8p_{32}}$
Population, persons		PP_{p_0}	PP_{p_1}	...	$PP_{p_{20}}$	$PP_{p_{21}}$...	$PP_{p_{32}}$

Finding the growth rate [R = (p/p_{t-1})-1]

1998	1999	...	2020	2021F	...	2030F
$\{ - \frac{ec_{1p_1}}{ec_{1p_0}} - 1 \dots \frac{ec_{1p_{22}}}{ec_{1p_{21}}} - 1 \dots \frac{ec_{1p_{32}}}{ec_{1p_{31}}} - 1 \}$	$\{ - \frac{ec_{2p_1}}{ec_{2p_0}} - 1 \dots \frac{ec_{2p_{22}}}{ec_{2p_{21}}} - 1 \dots \frac{ec_{2p_{32}}}{ec_{2p_{31}}} - 1 \}$...	$\{ - \frac{ec_{4p_1}}{ec_{4p_0}} - 1 \dots \frac{ec_{4p_{22}}}{ec_{4p_{21}}} - 1 \dots \frac{ec_{4p_{32}}}{ec_{4p_{31}}} - 1 \}$	$\{ - \frac{PP_{p_1}}{PP_{p_0}} - 1 \dots \frac{PP_{p_{22}}}{PP_{p_{21}}} - 1 \dots \frac{PP_{p_{32}}}{PP_{p_{31}}} - 1 \}$...	$\{ - \frac{eco_{4p_1}}{eco_{4p_0}} - 1 \dots \frac{eco_{4p_{22}}}{eco_{4p_{21}}} - 1 \dots \frac{eco_{4p_{32}}}{eco_{4p_{31}}} - 1 \}$

[RN] Rationing

1999-2030F
$\{ - \max\{1; \max\{^e R\}\} \dots \max\{1; \max\{^e R\}\} \}$

Normalized increments [Rn = R/RN]

1999	...	2020	2021F	...	2030F
$\{ \frac{ec_{1Rn_1}}{ec_{1Rn_1}} \dots \frac{ec_{1Rn_{22}}}{ec_{1Rn_{22}}} \dots \frac{ec_{1Rn_{32}}}{ec_{1Rn_{32}}} \}$	$\{ \frac{ec_{2Rn_1}}{ec_{2Rn_1}} \dots \frac{ec_{2Rn_{22}}}{ec_{2Rn_{22}}} \dots \frac{ec_{2Rn_{32}}}{ec_{2Rn_{32}}} \}$...	$\{ \frac{ec_{4Rn_1}}{ec_{4Rn_1}} \dots \frac{ec_{4Rn_{22}}}{ec_{4Rn_{22}}} \dots \frac{ec_{4Rn_{32}}}{ec_{4Rn_{32}}} \}$	$\{ \frac{PP_{Rn_1}}{PP_{Rn_1}} \dots \frac{PP_{Rn_{22}}}{PP_{Rn_{22}}} \dots \frac{PP_{Rn_{32}}}{PP_{Rn_{32}}} \}$	$\{ \frac{eco_{4Rn_1}}{eco_{4Rn_1}} \dots \frac{eco_{4Rn_{22}}}{eco_{4Rn_{22}}} \dots \frac{eco_{4Rn_{32}}}{eco_{4Rn_{32}}} \}$

Finding the absolute difference with the mean [A = p_t - p^{RF}]

1999	...	2020	2021F	...	2030F
$\{ \frac{ec_{1p_1}}{ec_{1p_1}} - \frac{ec_{1p_1}^{RF}}{ec_{1p_1}^{RF}} \dots \frac{ec_{1p_{22}}}{ec_{1p_{22}}} - \frac{ec_{1p_{22}}^{RF}}{ec_{1p_{22}}^{RF}} \dots \frac{ec_{1p_{32}}}{ec_{1p_{32}}} - \frac{ec_{1p_{32}}^{RF}}{ec_{1p_{32}}^{RF}} \}$	$\{ \frac{ec_{2p_1}}{ec_{2p_1}} - \frac{ec_{2p_1}^{RF}}{ec_{2p_1}^{RF}} \dots \frac{ec_{2p_{22}}}{ec_{2p_{22}}} - \frac{ec_{2p_{22}}^{RF}}{ec_{2p_{22}}^{RF}} \dots \frac{ec_{2p_{32}}}{ec_{2p_{32}}} - \frac{ec_{2p_{32}}^{RF}}{ec_{2p_{32}}^{RF}} \}$...	$\{ \frac{ec_{4p_1}}{ec_{4p_1}} - \frac{ec_{4p_1}^{RF}}{ec_{4p_1}^{RF}} \dots \frac{ec_{4p_{22}}}{ec_{4p_{22}}} - \frac{ec_{4p_{22}}^{RF}}{ec_{4p_{22}}^{RF}} \dots \frac{ec_{4p_{32}}}{ec_{4p_{32}}} - \frac{ec_{4p_{32}}^{RF}}{ec_{4p_{32}}^{RF}} \}$	$\{ \frac{PP_{p_1}}{PP_{p_1}} - \frac{PP_{p_1}^{RF}}{PP_{p_1}^{RF}} \dots \frac{PP_{p_{22}}}{PP_{p_{22}}} - \frac{PP_{p_{22}}^{RF}}{PP_{p_{22}}^{RF}} \dots \frac{PP_{p_{32}}}{PP_{p_{32}}} - \frac{PP_{p_{32}}^{RF}}{PP_{p_{32}}^{RF}} \}$	$\{ \frac{eco_{4p_1}}{eco_{4p_1}} - \frac{eco_{4p_1}^{RF}}{eco_{4p_1}^{RF}} \dots \frac{eco_{4p_{22}}}{eco_{4p_{22}}} - \frac{eco_{4p_{22}}^{RF}}{eco_{4p_{22}}^{RF}} \dots \frac{eco_{4p_{32}}}{eco_{4p_{32}}} - \frac{eco_{4p_{32}}^{RF}}{eco_{4p_{32}}^{RF}} \}$

[AN] Normalized differences [An = A/AN]

1999	...	2020	2021F	...	2030F
$\{ \frac{ec_{1An_1}}{ec_{1An_1}} \dots \frac{ec_{1An_{22}}}{ec_{1An_{22}}} \dots \frac{ec_{1An_{32}}}{ec_{1An_{32}}} \}$	$\{ \frac{ec_{2An_1}}{ec_{2An_1}} \dots \frac{ec_{2An_{22}}}{ec_{2An_{22}}} \dots \frac{ec_{2An_{32}}}{ec_{2An_{32}}} \}$...	$\{ \frac{ec_{4An_1}}{ec_{4An_1}} \dots \frac{ec_{4An_{22}}}{ec_{4An_{22}}} \dots \frac{ec_{4An_{32}}}{ec_{4An_{32}}} \}$	$\{ \frac{PP_{An_1}}{PP_{An_1}} \dots \frac{PP_{An_{22}}}{PP_{An_{22}}} \dots \frac{PP_{An_{32}}}{PP_{An_{32}}} \}$	$\{ \frac{eco_{4An_1}}{eco_{4An_1}} \dots \frac{eco_{4An_{22}}}{eco_{4An_{22}}} \dots \frac{eco_{4An_{32}}}{eco_{4An_{32}}} \}$

Evaluation of each indicator [SC = (Rn_t + An_t) · 0.5]

1999	...	2020	2021F	...	2030F
$\{ \frac{ec_{1SC_1}}{ec_{1SC_1}} \dots \frac{ec_{1SC_{22}}}{ec_{1SC_{22}}} \dots \frac{ec_{1SC_{32}}}{ec_{1SC_{32}}} \}$	$\{ \frac{ec_{2SC_1}}{ec_{2SC_1}} \dots \frac{ec_{2SC_{22}}}{ec_{2SC_{22}}} \dots \frac{ec_{2SC_{32}}}{ec_{2SC_{32}}} \}$...	$\{ \frac{ec_{4SC_1}}{ec_{4SC_1}} \dots \frac{ec_{4SC_{22}}}{ec_{4SC_{22}}} \dots \frac{ec_{4SC_{32}}}{ec_{4SC_{32}}} \}$	$\{ \frac{eco_{4SC_1}}{eco_{4SC_1}} \dots \frac{eco_{4SC_{22}}}{eco_{4SC_{22}}} \dots \frac{eco_{4SC_{32}}}{eco_{4SC_{32}}} \}$	$\{ \frac{eco_{4SC_1}}{eco_{4SC_1}} \dots \frac{eco_{4SC_{22}}}{eco_{4SC_{22}}} \dots \frac{eco_{4SC_{32}}}{eco_{4SC_{32}}} \}$

Group average [SCG = (ΣSC_t / n)]

1999	...	2020	2021F	...	2030F
$\{ \frac{\sum_{k=1}^4 ec_{kSC_1}}{4} \dots \frac{\sum_{k=1}^4 ec_{kSC_{22}}}{4} \dots \frac{\sum_{k=1}^4 ec_{kSC_{32}}}{4} \}$	$\{ \frac{\sum_{k=1}^4 eco_{kSC_1}}{4} \dots \frac{\sum_{k=1}^4 eco_{kSC_{22}}}{4} \dots \frac{\sum_{k=1}^4 eco_{kSC_{32}}}{4} \}$...	$\{ \frac{\sum_{k=1}^4 ec_{kSC_{23}}}{4} \dots \frac{\sum_{k=1}^4 eco_{kSC_{23}}}{4} \}$	$\{ \frac{\sum_{k=1}^4 ec_{kSC_{32}}}{4} \dots \frac{\sum_{k=1}^4 eco_{kSC_{32}}}{4} \}$	$\{ \frac{\sum_{k=1}^4 ec_{kSC_{32}}}{4} \dots \frac{\sum_{k=1}^4 eco_{kSC_{32}}}{4} \}$

Region	Region score [I = 100 * (ΣSCG _t + ΣSC _t ¹ + ΣSC _t ² + ΣSC _t ³ + ΣSC _t ⁴) / (5*2)]					
	1999	...	2020	2021F	...	2030F
1 Altai region	I_1	...	I_{12}	I_{123}	...	I_{132}
2 Amur region	I_2	...	I_{22}	I_{23}	...	I_{232}
...
84 Yamalo-Nenets Autonomous Okrug	I_{841}	...	I_{842}	I_{8423}	...	I_{8432}
85 Yaroslavl region	I_{851}	...	I_{852}	I_{8523}	...	I_{8532}

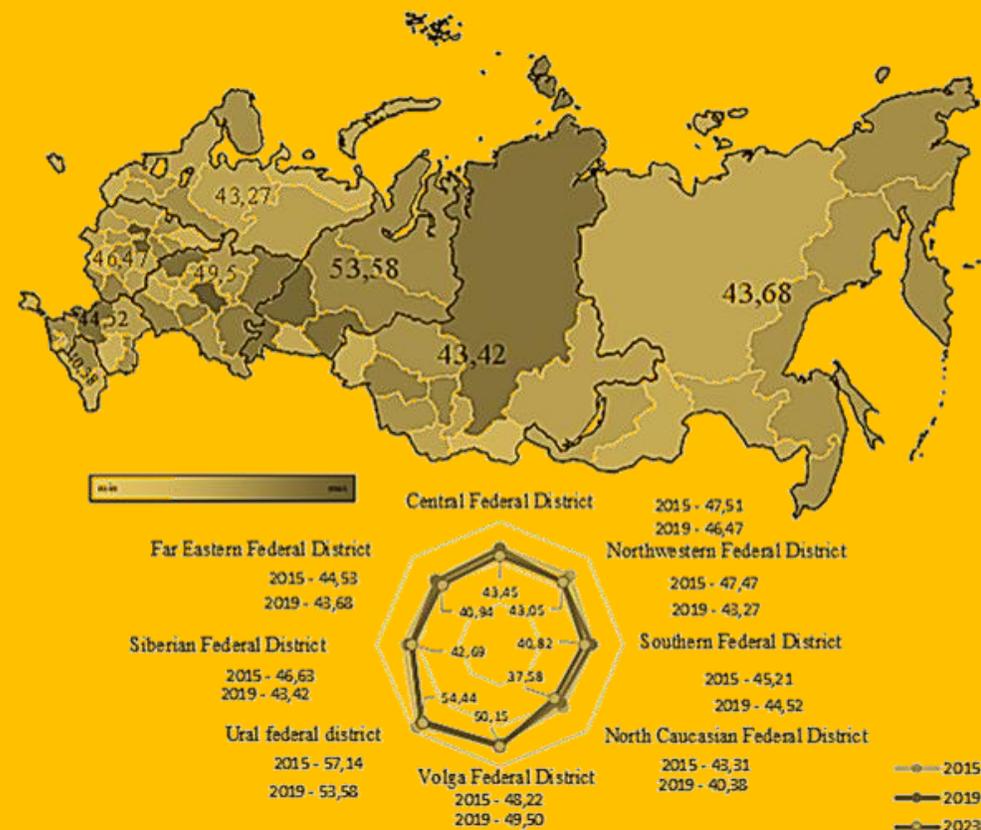
Разработанный алгоритм анализа данных может быть выполнен облачными вычислениями. Он позволяет агрегировать набор данных по регионам (на основе опросов и статистических данных) и создать рейтинговую шкалу.

*** При этом используя бесплатный доступный сервис Google возможно настроить автоматические вычисления (выполнение макроса – 1 раз/минуту)

АНАЛИЗ ДАННЫХ

Анализ цифровизации, отличается от обычного поскольку имеет дело с анализом структуры населения и компаний, а объем используемых данных может быть очень большим.

При этом анализ на основе опросника позволит точно выделить рейтинговую шкалу по каждому региону, благодаря которой так же доступно углубленное изучение драйверов данной динамики (анализ факторов, влияющих на совокупный результат).



ДАШБОРД ЦИФРОВИЗАЦИИ И ЦИФРОВЫХ КОМПЕТЕНЦИЙ НАСЕЛЕНИЯ



• Все звенья цепи по сбору, обработке и анализу, а так же визуализации данных выполняются в рамках разработанного алгоритма. На основе полученных результатов становится возможным построение системы визуализации данных.

• Анализ статистических данных дополняет проведенное исследование. В сравнении с общими показателями оценки цифровизации регионов, углубленное изучение цифровой грамотности населения в разрезе возрастных групп дает более точную оценку цифровизации.

• Модернизация приложения-опросника может быть направлена на создание универсального приложения, позволяющего проходить опросы по различным тематикам, аккумулируя необходимые данные для анализа и статистики регионов с соблюдением требований конфиденциальности.

ВЫВОДЫ И РЕЗУЛЬТАТЫ



1

Главным выводом из полученных результатов анализа является неоднородность развития регионов. Регионы, попавшие в лидеры данной рейтинговой шкалы, имеют большое количество организаций, которые производят и выделяют средства на исследования и разработки, обладают преобладающей долей населения, активно пользующимися интернетом, кроме того, использование информационных и коммуникационных технологий в организациях данных регионах используют передовые вычислительные средства.

2

Построенный рейтинг позволяет оценить неравенство в цифровизации между регионами России и выявить источники диспропорций. Данное исследование показывает, что не все регионы России одинаково готовы к цифровой экономике. Если организации в целом демонстрируют относительно высокий уровень цифровизации и компьютеризации, то доля активных пользователей интернета не везде одинаково высока. Полученные данные с помощью опросника позволяют выделить группы населения, активно пользующиеся интернетом и имеющие высокую цифровую грамотность.

3

Можно заключить, что, применяя для анализа широкий объем данных, возможным становится анализ совершенно любых показателей, отражающих динамику того или иного явления. Рассмотренная на примере анализа данных цифровизации и цифровых компетенций населения показала эффективность разработанного алгоритма и его применимость для использования большого массива информации.

ПРИЛОЖЕНИЕ ДОСТУПНО ПО QR-КОДУ



Пройди тест и оцени уровень своих цифровых компетенций!



***Приложение имеет уникальную цифровую подпись, поэтому проигнорируйте оповещение Google-Play защитника.



СПАСИБО ЗА
ВНИМАНИЕ!

