

What the Yandex Leak Tells Us About How Big Tech Uses Your Data

Kaileigh McCrea, Privacy Engineer @ Confiant

August 2023



About Me

Kaileigh McCrea

- Privacy Engineer at Conflant (3 yrs)
- Software Engineer (6 years)
- Cybersecurity Nerd
- Recovering Political Science major
- Twitter: @kaileighrose

What we're talking about

Yandex

The screenshot shows a post on the BreachForums forum. The title is "yandex git sources" and it was posted by "breachforums" on Wednesday January 25, 2023 at 03:48 PM. The post content is a link to a GitHub repository containing source code for Yandex services. The user is identified as a "BreachForums User" and has the status "MEMBER". Below the post, there is a summary of activity: Posts: 1, Threads: 1, Joined: Jan 2023, Reputation: 0. At the bottom of the post area, there are buttons for "PM", "Find", "Reply", "Quote", and "Report".

YANDEX SERVICES SOURCE CODE LEAK

SHORT OVERVIEW OF BREACH CONTENTS

PUBLISHED THU, JAN 26, 2023 BY ARSENIY SHESTAKOV

The screenshot shows a news article from BleepingComputer. The main headline is "Yandex denies hack, blames source code leak on former employee". The article is by Bill Toulas and was published on January 26, 2023, at 09:44 AM. The article URL is [https://www.bleepingcomputer.com/news/security/yandex-denies-hack-blames-source-code-leak-on-former-employee/](#). The BleepingComputer navigation bar includes links for NEWS, DOWNLOADS, VPNs, VIRUS REMOVAL GUIDES, TUTORIALS, and DEALS. There is also a search bar and social media sharing icons.

Roadmap

- Background on Yandex Leak
- Dive into code:
 - What data Yandex is collecting
 - What Yandex is doing with that data
 - Who Yandex is sharing that data with
- Conclusions and wrap up
- Q&A

Roadmap

- Background on Yandex Leak
- Dive into code:
 - What data Yandex is collecting
 - What Yandex is doing with that data
 - Who Yandex is sharing that data with
- Conclusions and wrap up
- Q&A

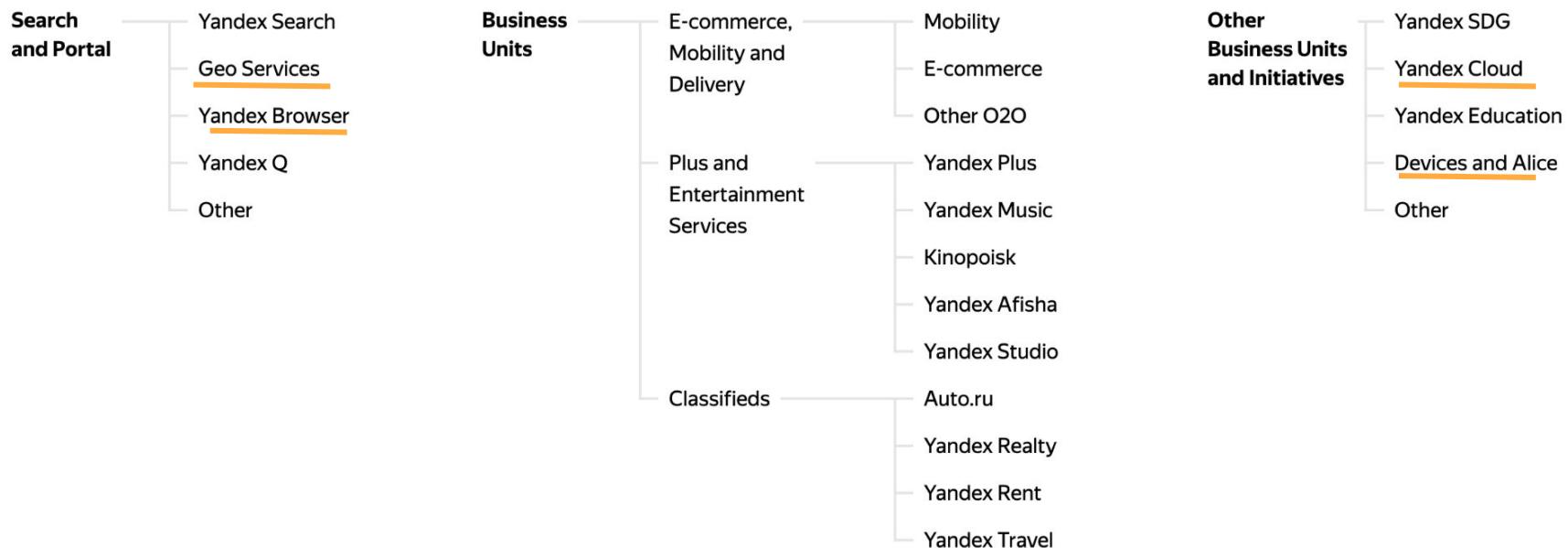
Yandex 101

Yandex 101

The screenshot shows the Yandex homepage with a dark theme. At the top left is the Yandex logo. On the right are buttons for "Log in" and a menu icon. Below the header is a row of service icons: Games (purple controller), Images (camera), Video (play button), Meteum (weather symbol), Maps (location pin), Mail (envelope), and Translate (two languages). A large search bar in the center contains the placeholder text "finds everything". Below the search bar, the current weather is displayed as "26°" with a sun icon. At the bottom right, there are three small gray squares.

Yandex 101

Key Businesses



Yandex 101



AppMetrica

In-depth analytics
for product and
growth teams



Crypto

helps to identify
important user
characteristics
for advertisers



Audiences

allows you to pull
data from several
sources to
generate your own
targeted segments

Yandex 101

Yandex LLC

Head office in Russia: Moscow

Head office

16, Leo Tolstoy St., Moscow, Russia
119021
tel.: +7 495 739-70-00
fax: +7 495 739-70-70

Advertising clients

tel.: +7 495 739-37-77
fax: +7 495 739-23-32
adv@yandex-team.ru

Investor Relations

tel.: +7 495 974-35-38
askIR@yandex-team.ru

Public relations

pr@yandex-team.ru

Corporate Secretary

secretary@yandex-team.ru

Sustainability

sustainability@yandex-team.com

Yandex N.V.

Registered office in Amsterdam

Schiphol Boulevard 165, 1118 BG Schiphol, The Netherlands
tel.: +31 0 20 206 6970

Official Telegram channel for individual investors https://t.me/yndx_forinvestors (in Russian only)

Yandex: A Drama

Support us →

Russia

• This article is more than 3 years old

Advertisement

Russian internet giant grants veto powers to Kremlin-linked body

Yandex agrees to corporate restructuring in move likely to increase government oversight

Andrew Roth in Moscow

Mon 18 Nov 2019 06.30 EST



Arkady Volozh, the chief executive of Yandex, said the company would maintain control over its "Golden Share" operations. Photograph: Milen Tass

2019

October

November

December

January

February

March

April

2020

• This article is more than 1 year old

Advertisement

Warnings raised over Russian tech giant Yandex's UK operation

MPs want restrictions placed on the company, known as Russia's Google, which also runs the Yango Deli grocery service

• [Russia-Ukraine war: live news](#)



2022

Russia invades
Ukraine

Alarm About Sending
Data to Russian Servers

September

November

December

January

February

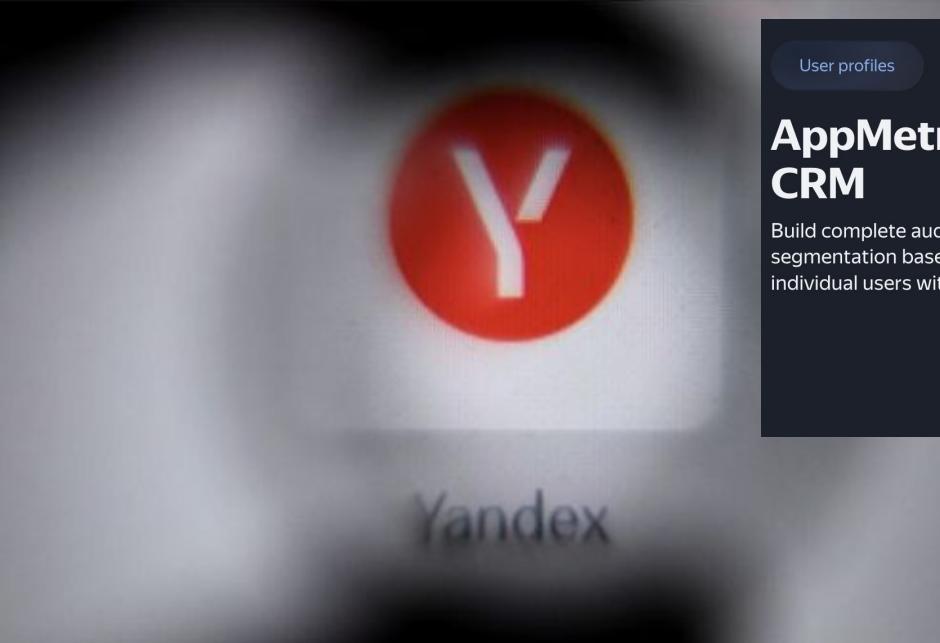
March

April

Data-harvesting code in mobile apps sends user data to “Russia’s Google”

Data from apps on Apple- and Google-powered mobile devices is sent to Russian servers.

PATRICK MCGEE, FINANCIAL TIMES - 3/29/2022, 7:18 AM



User profiles

AppMetrica: Your app's CRM

Build complete audience knowledge with segmentation based on profile data or dive into individual users with profile cards.

♂ 18-24

📍 ⓘ

Today

- Launch app
- Start onboarding
- Go to catalog
- View item
- Add to cart

2022

September

November

December

January

Russia invades Ukraine

Alarm About Sending Data to Russian Servers

February

March

April

““Yandex has acknowledged its software collects “device, network and IP address” information ... but it called this data “non-personalised and very limited”. It added: “Although theoretically possible, in practice it is extremely hard to identify users based solely on such information collected. Yandex definitely cannot do this.””

– *Financial Times*

Featured Article

Russia's war hits Yandex, the 'Google of Russia'

Sources say the company is seeking a media exit as top exec hit with sanctions over propaganda charge

Natasha Lomas, Ingrid Lunden / 12:20 PM PDT • March 16, 2022

 Comment



Premium HOME > TECH

'I bought a plane ticket and left 12 hours later': Engineers at Yandex, Russia's Google rival, are fleeing abroad and leaving spouses and salaries behind

Rosie Bradbury Apr 12, 2022, 3:35 AM PDT



2022

Russia invades Ukraine

Exec hit with EU propaganda sanctions

Alarm About Sending Data to Russian Servers

Thousands of Engineers have Fled Russia

October

November

December

January

February

March

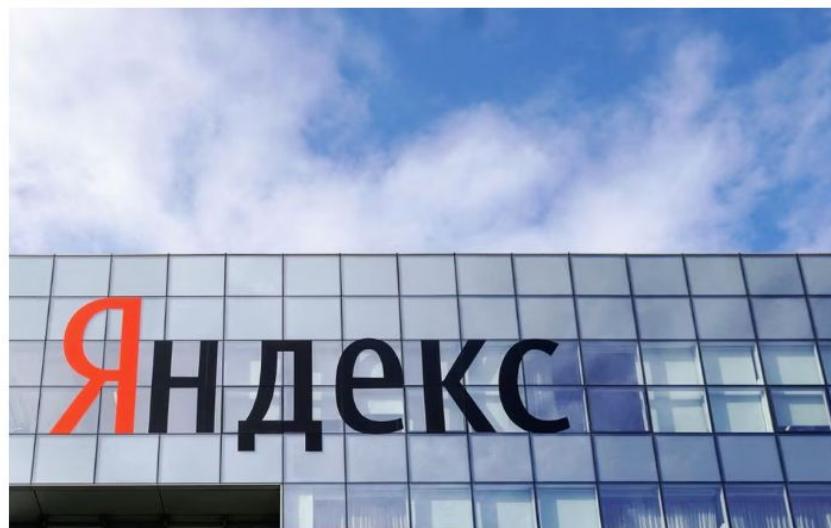
April

Technology

Yandex CEO resigns after being targeted by EU sanctions

Reuters

June 3, 2022 7:35 AM PDT · Updated a year ago



The logo of Russian internet group Yandex is pictured at the company's headquarter in Moscow, Russia October 4, 2018. REUTERS/Shamil Zhumatov

June 3, 2022

Russia invades Ukraine

Exec hit with EU propaganda sanctions

down as CEO and left the board of directors

sanc

Alarm About Sending Data to Russian Servers

Yandex (YNDX.O) said on Friday that Arkady Volozh had stepped down as CEO and left the board of directors

him on its latest list of

Thousands of Engineers have Fled Russia

CEO steps down

2022

January

February

March

April

May

June

July

[Join TechCrunch+](#)[Login](#)[Search Q](#)[TechCrunch+](#)[Startups](#)[Venture](#)[Security](#)[AI](#)[Crypto](#)[Apps](#)[Events](#)[More](#)

Yandex's sale of media assets to VK includes yandex.ru homepage

Natasha Lomas @riptari / 12:05 AM PDT • August 23, 2022

 Comment



Exec hit with EU
propaganda sanctions

Thousands of Engineers
have Fled Russia

CEO steps down

Yandex sells media assets to
VK

2022

March

April

May

June

July

August

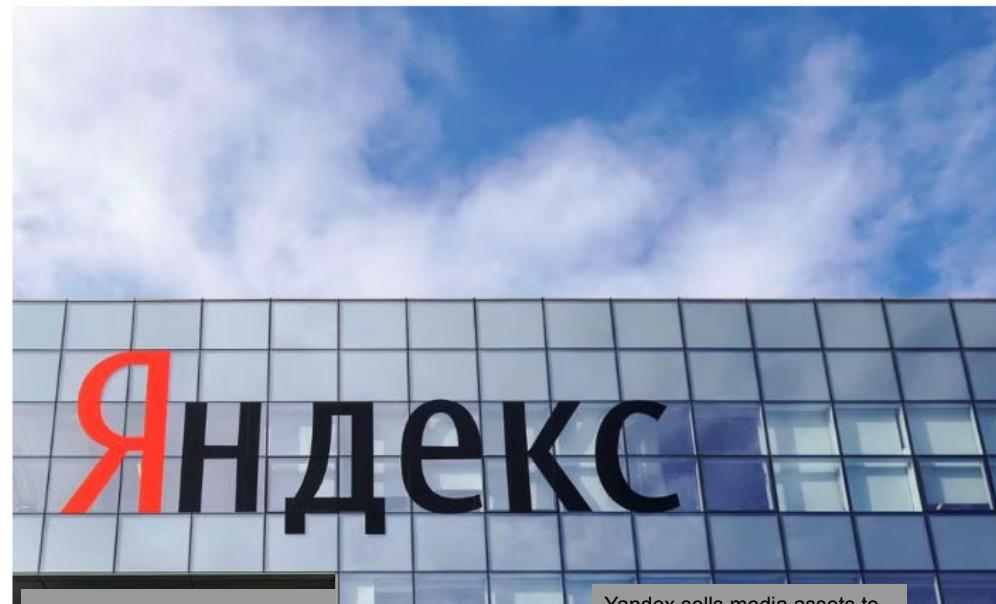
September

Deals

Yandex parent to review ownership of Russian tech giant, seek divestment

By Alexander Marrow, Darya Korsunskaya and Polina Devitt

November 25, 2022 7:31 AM PST · Updated 8 months ago



CEO steps down

Yandex sells media assets to VK

utch parent company restructures
Putin Ally steps in

2022

May

June

July

August

September

October

November

20

Putin, Kudrin touch on future of Yandex in late-night meeting -sources

Reuters

November 25, 2022 4:19 AM PST · Updated 8 months ago



[1/2] The logo of Russian internet group Yandex is pictured at the company's headquarter in Moscow, Russia October 4, 2018. REUTERS/Shamil Zhumatov



CEO steps down

Yandex sells media assets to VK

Dutch parent company restructures,
Putin Ally steps in

2022

May

June

July

August

September

October

November

Kremlin Ally Kudrin Confirms Move to Tech Giant Yandex

Dec. 5, 2022



Dutch parent company
restructures, Putin Ally steps in

2023

Yandex Source Code Leak

November

December

January

February

March

April

May



Russian Billionaires Line Up to Buy Yandex – Reports

May 4, 2023



Yandex Source Code Leak

Oligarchs look to buy Yandex
Russia Loses Record \$253Bln in War Flight

2023

January

February

March

April

May

June

July

MOST READ JUST IN

- 1 [OVERNIGHT STRIKE](#)
Russia Says Ukrainian Drones Hit Crimea
- 2 [NO PASSAGE](#)
Russia Blocks Cargo Ship Over 'Explosives' Traces'
- 3 [POLITICAL PRISONER](#)
Navalny Ally Jailed 9 Years for 'Extortion'
- 4 [MORE MANPOWER](#)
Russia Raises Upper-Age Limit for Reservists
- 5 [MONEY DRAIN](#)
Russia Loses Record \$253Bln in War Flight

Technology

Russia's Yandex fined for refusing to share user information with security services

Reuters

June 18, 2023 3:20 PM PDT · Updated a month ago



Yandex fined for not sharing data with security services

Oligarchs look to buy Yandex

Yandex Source Code Leak

2023

January

February

March

April

May

June

24

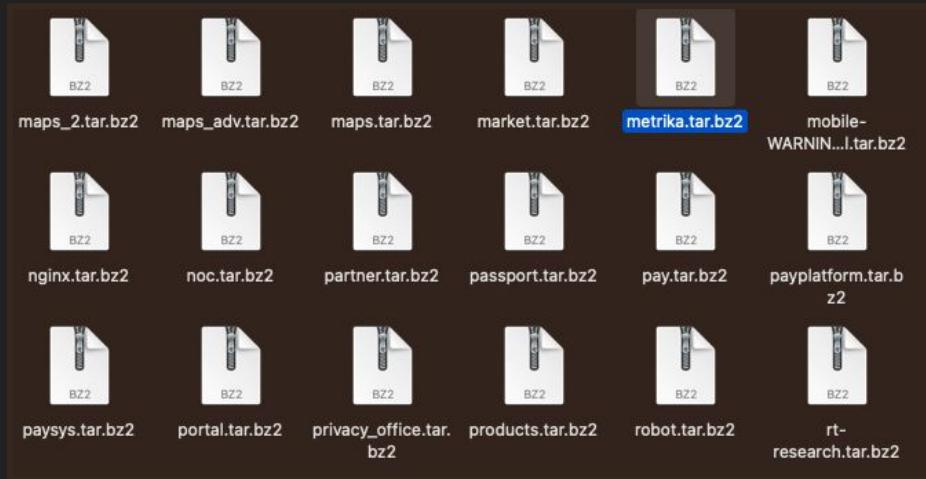
July

Roadmap

- Background on Yandex Leak
- Dive into code:
 - What data Yandex is collecting
 - What Yandex is doing with that data
 - Who Yandex is sharing that data with
- Conclusions and wrap up
- Q&A

Yandex Codebase

Codebase



Metrika

[Solutions](#)[Features](#)[Verticals](#)[Resources](#)

Supercharge app metrics with data insights

with a one-stop solution for analytics and marketing

[Get started](#)[Take a tour](#)

Yandex Metrica

[Features](#)[Resources](#)[Pricing](#)

All-Round Web Analytics

From traffic trends to mouse movements – get a comprehensive understanding of your online audience and drive business growth.

[Get started](#)[Try live demo](#)

Example Raw Data Fields that AppMetrica Logs

analytics > appmetrica-location-log-anonymizer > convert_log.yql

```
89  insert into `//home/metrika-logs/anonym-appmetrica-location-log/1d/{table_date}`  
90  with truncate  
91  select  
92      String::HexEncode(Digest::Blake2B(`DeviceID`, seed)) as `DeviceID`,  
93      String::HexEncode(Digest::Blake2B(`ADVID`, seed)) as `ADVID`,  
94      String::HexEncode(Digest::Blake2B(`IFA`, seed)) as `IFA`,  
95      String::HexEncode(Digest::Blake2B(`UUID`, seed)) as `UUID`,  
96      String::HexEncode(Digest::Blake2B(`AndroidID`, seed)) as `AndroidID`,  
97      `APIKey`,  
98      `AppBuildNumber`,  
99      `AppFramework`,  
100     `AppID`,  
101     `AppPlatform`,  
102     `AppVersionName`,  
103     `Cells_AreConnected`,  
104     `Cells_CellsIDs`,  
105     `Cells_CountriesCodes`,  
106     `Cells_Lacs`,  
107     `Cells_LastVisibleTimeOffset`,  
108     `Cells_OperatorsIDs`,  
109     `Cells_OperatorsNames`,  
110     `Cells_PhysicalCellsIDs`,  
111     `Cells_SignalsStrengths`,  
112     `Cells_Types`,  
113     `ChargeType`,  
114     `ClientIP`,  
115     `ClientIPHash`,  
116     `CollectTimestamp`,  
117     `CollectTimestampBootOffset` , |  
118     `CollectionMode`,  
119     `DeviceType`,  
120     `EventID`,  
121     `IncrementalID`,  
122     `IsExtraLocationEvent`,  
123     `IsRooted`,  
124     `KitBuildNumber`,  
125     `KitBuildType`,  
126     `KitVersion` ,
```

analytics > appmetrica-location-log-anonymizer > convert_log.yql

```
122     `IsExtraLocationEvent`,  
123     `IsRooted`,  
124     `KitBuildNumber`,  
125     `KitBuildType`,  
126     `KitVersion`,  
127     `Latitude`,  
128     `LatitudeLBS`,  
129     `LocationAltitude`,  
130     `LocationDirection`,  
131     `LocationEnabled`,  
132     `LocationPrecision`,  
133     `LocationPrecisionLBS`,  
134     `LocationSource`,  
135     `LocationSpeed`,  
136     `LocationTimestamp`,  
137     `LocationTimestampBootOffset`,  
138     `Longitude`,  
139     `LongitudeLBS`,  
140     `OSApiLevel`,  
141     `OSVersion`,  
142     `OperatingSystem`,  
143     `OriginalCollectTimestamp`,  
144     `OriginalLocationTimestamp`,  
145     `ReceiveTimestamp`,  
146     `RequestID`,  
147     `SendTimestamp`,  
148     `Wifi_AreConnected`,  
149     `Wifi_LastVisibleTimeOffset`,  
150     `Wifi_Macs`,  
151     `Wifi_SignalsStrengths`,  
152     `Wifi_Ssids`,  
153     `_logfeller_index_bucket`,  
154     `_logfeller_timestamp`,  
155     `_rest`,  
156     `_stbx`,  
157     `iso_eventtime`,  
158     `source_uri`,  
159     `subkey` ,
```

Anonymized Identifiers

```
91    select
92        String::HexEncode(Digest::Blake2B(`DeviceID`, seed)) as `DeviceID`,
93        String::HexEncode(Digest::Blake2B(`ADVID`, seed)) as `ADVID`,
94        String::HexEncode(Digest::Blake2B(`IFA`, seed)) as `IFA`,
95        String::HexEncode(Digest::Blake2B(`UUID`, seed)) as `UUID`,
96        String::HexEncode(Digest::Blake2B(`AndroidID`, seed)) as `AndroidID`,
```

Location Fields

```
126 `KitVersion`,  
127 `Latitude`,  
128 `LatitudeLBS`,  
129 `LocationAltitude`,  
130 `LocationDirection`,  
131 `LocationEnabled`,  
132 `LocationPrecision`,  
133 `LocationPrecisionLBS`,  
134 `LocationSource`,  
135 `LocationSpeed`,  
136 `LocationTimestamp`,  
137 `LocationTimestampBootOffset`,  
138 `Longitude`,  
139 `LongitudeLBS`,  
140 `OSApiLevel`,  
141 `OSVersion`,  
142 `OperatingSystem`,  
143 `OriginalCollectTimestamp`,  
144 `OriginalLocationTimestamp`,  
145 `ReceiveTimestamp`,  
146 `RequestID`,
```

Wifi Fields Collected by AppMetrica

```
147     `SendTimestamp` ,  
148     `Wifi_AreConnected` ,  
149     `Wifi_LastVisibleTimeOffset` ,  
150     `Wifi_Macs` ,  
151     `Wifi_SignalsStrengths` ,  
152     `Wifi_Ssids` ,  
153     `loafeller index bucket` .
```

Those Same Fields in Crypta

graph > fuzzy > lib > yql > ≡ export_ssid_devid_day_table.yql

```
26
27 $list_metrika_log = (
28     select coalesce(DeviceID, "") as DeviceID,
29             coalesce(OriginalDeviceID, "") as OriginalDeviceID,
30             $MakeStringList(Wifi_Macs) as Wifi_Macs,
31             $MakeStringList(Wifi_Ssids) as Wifi_Ssids,
32             $MakeIntList(Wifi_SignalsStrengths) as Wifi_SignalsStrengths,
33             $MakeIntList(Wifi_AreConnected) as Wifi_AreConnected
34     from `'{source_mmetric_table}`
35     where DeviceID is not null
36 );
```

Dev Id and SSID Associated with Yandex UID

```
graph > fuzzy > lib > yql >  ≡ export_ssids_yuids.yql
6   $mobile_all_table = (
7     select distinct mmetric_devid, ssid
8     from concat({sources})
9   );
10
11 $mmetric_to_devid = (
12   select mmetric_devid, devid,
13     |   coalesce(cast(yuid as uint64), 0) as yuid
14   from `'{source_nolimit_table}`
15 );
16
```

Click Event Data Being Matched to Existing Users

core > programs > clicklogd-mobile > src > `C event_indexed_pool.h` > `TEventIndexedPool` > `GetIndex<TMatchCriteria>()`

```
58     private:
59         template <class TMatchCriteria>
60         TIndex<TMatchCriteria>& GetIndex() {
61             if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TAndroidId>) {
62                 return AndroidId_;
63             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TAndroidIdMd5>) {
64                 return AndroidIdMd5_;
65             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TAndroidIdSha1>) {
66                 return AndroidIdSha1_;
67             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TDeviceIdHash>) {
68                 return DeviceIdHash_;
69             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TFingerprint>) {
70                 return Fingerprint_;
71             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TGoogleAid>) {
72                 return GoogleAid_;
73             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TGoogleAidMd5>) {
74                 return GoogleAidMd5_;
75             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TGoogleAidSha1>) {
76                 return GoogleAidSha1_;
77             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TIfa>) {
78                 return Ifa_;
79             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TIfaMd5>) {
80                 return IfaMd5_;
81             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TIfaSha1>) {
82                 return IfaSha1_;
83             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TWindowsAid>) {
84                 return WindowsAid_;
85             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TWindowsAidMd5>) {
86                 return WindowsAidMd5_;
87             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TWindowsAidSha1>) {
88                 return WindowsAidSha1_;
89             } else if constexpr (std::is_same_v<TMatchCriteria, NMatchCriteria::TYmTrackingId>) {
90                 return YmTrackingId_;
91             }
92         }
```

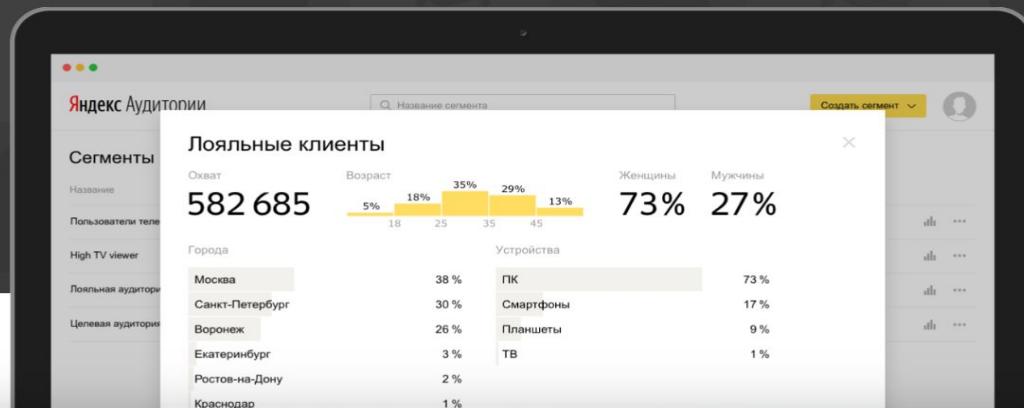
Socio-Demographic Attributes for DevID Being Updated

```
core > programs > socdem-updated-mobile > src > UserIdAndInfoParser.cpp > ...
45     t"0_1", AgeIntervalsCrypt::LessThan18},
46     {"18_24", AgeIntervalsCrypt::Between18and24},
47     {"25_34", AgeIntervalsCrypt::Between25and34},
48     {"35_44", AgeIntervalsCrypt::Between35and44},
49     {"45_54", AgeIntervalsCrypt::Between45and54},
50     {"55_99", AgeIntervalsCrypt::MoreThan55}
51 };
52 ::setValue(value, exact_socdem_node, key, json_keys_to_ages_intervals);
53 }
54
55 void UserIdAndInfoParser::setValue(
56     SexTypesCrypt & value,
57     const NYT::TNode & exact_socdem_node,
58     const TString key)
59 {
60     static const std::map<TString, SexTypesCrypt> json_keys_to_sex_types =
61     {
62         {"f", SexTypesCrypt::Female},
63         {"m", SexTypesCrypt::Male}
64     };
65 ::setValue(value, exact_socdem_node, key, json_keys_to_sex_types);
66 }
67
68 std::string UserIdAndInfoParser::parse(const NYT::TNode & user_record)
69 {
70     const TString & device_id = user_record["appmetrica_devid"].AsString();
71     UserInfo user_info;
72
73     const auto & exact_socdem = user_record["exact_socdem"];
74     setValue(user_info.age, exact_socdem, "age_segment");
75     setValue(user_info.sex, exact_socdem, "gender");
76
77     static const auto tail = getConstTail();
78
79     std::ostringstream buffer;
80     buffer <<
81         sipHash64(device_id.data(), device_id.size()) << '\t' <<
82         static_cast<int>(user_info.age) << '\t' <<
```



Create segments based on
offline and online data

Create Segment



Crypta

The screenshot shows the Yandex website with a red header bar containing the Yandex logo and navigation links: COMPANY, JOBS, FOR DEVELOPERS, FOR ADVERTISERS, FOR INVESTORS, and a menu icon. Below the header is a grey navigation bar with links: About, History, Privacy, Press Releases, Blog, Contact, and a search icon. The main content area has a white background. At the top left of this area, there is a breadcrumb trail: Technologies / Crypta. The word "Crypta" is prominently displayed in large black text. Below the title, there is a paragraph of text describing the technology.

Technologies /

Crypta

Every day, millions of web users are exposed to banner ads on the pages of Yandex's sites. Advertisers on Yandex can opt to show their ads only to that part of the viewer audience that is potentially interested in seeing them, such as people of a certain age or gender. To enable advertisers to target their ads to a specific audience, Yandex uses its own proprietary behavior analytics technology called Crypta. This technology allows classification of web users based on their online behavior. Their behaviour just has to differ somehow.

Example Segments

CRYPTA
> test
age_segment_18_20.py
alice_users.py
apartment_room_number.py
apps_users.py
artists.py
auto_interactions.py
avia_travellers.py
bank_cards.py
bought_two_tickets.py
business_travellers.py
children_age_segment_clarification.py
compulsory_auto_insurance.py
connection_type.py
console_gamers.py
contest.py
devices_without_google_services.py
digital_viewers.py
direct_clients_by_industry.py
direct_product_users.py
disk_users.py
ecommerce_owners.py
edadeal_offline_purchases_lal.py
expensive_car_customers.py
film_lovers_by_genres.py
gas_stations.py
industryRepresentatives.py
kbt_customers.py
kfc_visitors.py
kinopoisk_logins.py
kinopoisk_movie_watchers.py
> CRYPTA
kinopoisk_movie_watchers.py
kz_users.py
laptop_users.py
logged_in_for_plus.py
longterm_interest_mobile_gamers.py
loyal_to_launcher_install.py
macos_users.py
mail_data.py
manufacturer_phone_owners.py
marketplaces_ltv_users.py
mobile_gamers.py
mobile_operators_users_by_prefix.py
mobile_operators_users.py
multidevice_puid.py
multidevice.py
music_genres_listeners.py
nestle_regions.py
phone_buyers.py
phone_owners.py
phone_with_esim_owners.py
potential_aon_android_users.py
potential_aon_ios_users.py
preinstalled_apps.py
prism.py
proleads.py
realty_interactions.py
recent_passport_accounts.py
score_users_for_telephony.py
searched_for_phone_numbers.py
searched_radisson_on_maps.py
seo_specialists.py
seo_users.py
smart_gadgets_customers.py
smokers.py
summer_residents.py
travellers.py
video_bloggers.py
want_to_change_the_provider.py
webmaster.py
widgets.py
with_children_by_ages.py
via_make
INTERESTS
sections.js
webportal/src/public-info/sections/interests...

Example Segments

- ⌚ smart_gadgets_customers.py
- ⌚ smokers.py
- ⌚ summer_residents.py
- ⌚ travellers.py
- ⌚ video_bloggers.py
- ⌚ want_to_change_the_provider.py
- ⌚ webmaster.py
- ⌚ widgets.py
- ⌚ with_children_by_ages.py
- ≡ va make

Travellers

profile > runners > segments > lib > coded_segments > 🗃 travellers.py > ...

```
82
83     INSERT INTO `'{output_table}`` WITH TRUNCATE
84     SELECT
85         id,
86         id_type,
87         segment_name
88     FROM(
89         SELECT
90             crypta_id AS id,
91             'crypta_id' AS id_type,
92             CASE
93                 WHEN Geo::RoundRegionById(region, "country").id != Geo::RoundRegionById(CAST(main_region AS Int32), "country").id THEN 'international'
94                 ELSE 'domestic'
95             END AS segment_name,
96             MAX(`date`) AS last_seen,
97             MIN(`date`) AS first_seen,
98             region,
99             week_end_date,
100            FROM $travell_visits
101            GROUP BY region, main_region, crypta_id, week_end_date
102        )
103        WHERE
104            last_seen <= week_end_date AND
105            DateTime::ToDays(DateTime::MakeTimestamp($parse(last_seen)) - DateTime::MakeTimestamp($parse(first_seen))) > 0 AND
106            DateTime::ToDays(DateTime::MakeTimestamp($parse(week_end_date)) - DateTime::MakeTimestamp($parse(first_seen))) <= 7
107        GROUP BY id, id_type, segment_name
108        """
109
110
```

Mail Data

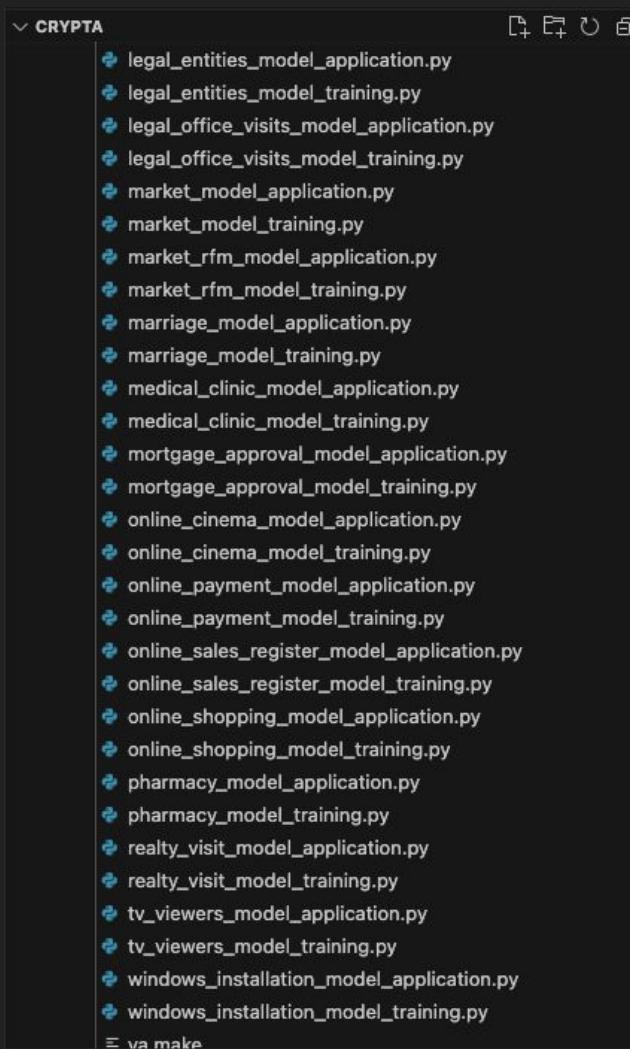
```
profile > runners > segments > lib > coded_segments > mail_data.py > ...
12
13
14     segment_query = """
15     INSERT INTO `{{output_table}}` WITH TRUNCATE
16     SELECT id, id_type, segment_name
17     FROM `{{mail_data_table}}`;
18
19     INSERT INTO `{{sample_table}}` WITH TRUNCATE
20     SELECT
21         yandexuid,
22         segment_name
23     FROM (
24         SELECT matching.yandexuid AS yandexuid, mail_data.segment_name AS segment_name
25         FROM `{{mail_data_table}}` AS mail_data
26         INNER JOIN `{{indevice_yandexuid_matching}}` AS matching
27         USING (id, id_type)
28     )
29     GROUP BY yandexuid, segment_name
30     """
31
32
33 class PrepareMailSampleForLalSegments(RegularSegmentBuilder):
34     keyword = 549
35     name_segment_dict = {
36         'aviaticket': 1404,
37         'boardingpass': 1405,
38         'hotel': 1406,
39     }
```

Gas Stations

profile > runners > segments > lib > coded_segments >  gas_stations.py > ...

```
92     class ProcessedDeepVisitLogForGasStations(DayProcessor):
93         def requires(self):
94             return deep_visits.org_visits_deep_external_input(self.date)
95
96         def process_day(self, inputs, output_path):
97
98             self.yql.query(
99                 gas_stations_query_template.format(
100                     organization_categories=config.ORGANIZATION_CATEGORIES,
101                     deep_visits=inputs.table,
102                     matching_idfa=get_matching_table('idfa', 'crypto_id'),
103                     matching_gaid=get_matching_table('gaid', 'crypto_id'),
104                     name_to_variable=',\n'.join(
105                         [u'("{}", "{}")'.format(key, value)
106                         | for key, value in name_to_variable.iteritems()]
107                         ),
108                     output_table=output_path,
109                     ),
110                     transaction=self.transaction,
111             )
```

Example ML Model Types



The screenshot shows a file explorer window with a dark theme. At the top, there are icons for file operations: copy, paste, move, and delete. Below the title bar, the directory path is shown as 'CRYPTA'. The main area displays a list of files, each preceded by a small blue circular icon with a white symbol. The files are listed in alphabetical order:

- legal_entities_model_application.py
- legal_entities_model_training.py
- legal_office_visits_model_application.py
- legal_office_visits_model_training.py
- market_model_application.py
- market_model_training.py
- market_rfm_model_application.py
- market_rfm_model_training.py
- marriage_model_application.py
- marriage_model_training.py
- medical_clinic_model_application.py
- medical_clinic_model_training.py
- mortgage_approval_model_application.py
- mortgage_approval_model_training.py
- online_cinema_model_application.py
- online_cinema_model_training.py
- online_payment_model_application.py
- online_payment_model_training.py
- online_sales_register_model_application.py
- online_sales_register_model_training.py
- online_shopping_model_application.py
- online_shopping_model_training.py
- pharmacy_model_application.py
- pharmacy_model_training.py
- realty_visit_model_application.py
- realty_visit_model_training.py
- tv_viewers_model_application.py
- tv_viewers_model_training.py
- windows_installation_model_application.py
- windows_installation_model_training.py

At the bottom right of the file list, there is a small text entry field with the placeholder 'va make'.

Basic Example of Household Details

```
graph > metrics > household > query.sql
42  DEFINE SUBQUERY $hh_size_by_yuid($title) AS
43    SELECT ($title || key) AS key, COUNT(1) AS hh_size_by_yuids
44    FROM $composition
45    GROUP BY $size_to_range(size) AS key;
46  END DEFINE;
47
48  DEFINE SUBQUERY $hh_size_by_crypta_id($title, $predicat) AS
49    SELECT
50      ($title || key) AS key,
51      COUNT(1) AS hh_size_by_crypta_id
52    FROM $composition
53    WHERE $predicat(size, socdems)
54    GROUP BY CAST(Yson::GetLength(Yson::Lookup(data, 'crypta_ids')) AS String) AS key;
55  END DEFINE;
56
57  DEFINE SUBQUERY $hh_by_socdems($title, $predicat) AS
58    $hh_socdem = (
59      SELECT
60        hhid,
61        size,
62        IF((Yson::LookupInt64(info, 'female') != 0), 'female', Null) AS has_female,
63        IF((Yson::LookupInt64(info, 'male') != 0), 'male', Null) AS has_male,
64        IF((Yson::LookupInt64(info, 'grand') != 0), 'grand', Null) AS has_old,
65        IF((Yson::LookupInt64(info, 'child') != 0), 'child', Null) AS has_child
66      FROM $composition
67      WHERE $predicat(size, socdems)
68    );
69
70    SELECT ($title || groups) AS key, hh_c AS hh_socdem_count
71    FROM (
72      SELECT groups, SUM(size) AS hh_c
73      FROM $hh_socdem
74      GROUP BY String::JoinFromList(
75        ListSort(AsList(has_female, has_male, has_old, has_child)),
76        '_') AS groups
77    ) WHERE groups != "";
    }
```

Appmetrica Being Used to Pull Wifi Connection Types:

```
profile > runners > segments > lib > coded_segments > connection_type.py > ...
59     connection_type_query = """
60     INSERT INTO `'{output_table}'` WITH TRUNCATE
61     SELECT
62         id,
63         'mm_device_id' AS id_type,
64         CASE
65             WHEN types == AsSet('3g') THEN '3g'
66             WHEN types == AsSet('4g') THEN '4g'
67             ELSE '3g_4g'
68         END AS segment_name
69     FROM (
70         SELECT
71             id,
72             ToSet(AGGREGATE_LIST_DISTINCT(segment_name)) AS types
73         FROM `'{input_table}'`
74         GROUP BY id
75     );
76     """
77
78
79     class ConnectionType(RegularSegmentBuilder):
80         name_segment_dict = {
81             '3g': (557, 17823841),
82             '4g': (557, 17823853),
83             '3g_4g': (557, 17823847),
84         }
85
86         number_of_days = 35
87
88         def requires(self):
89             return [
90                 'AppMetrica': LogProcessor(
91                     ProcessAppMetricaForConnectionType,
92                     self.date,
93                     self.number_of_days,
94                 ),
95             ]
96     
```

Appmetrica Data Being Used to Identify Users with Common SSIDS (Wifi Networks)

```
class ImportSsidMobileMetrikaTask(BaseTask):
    date = DateParameter()
    SSID_THRESHOLD = 20
    YUID_THRESHOLD = 20
    DAYS_IN_MONTH = 7

    def requires(self):
        .....
        this tasks must be done to complete this task
        .....
        task_list = [
            ImportSsidMobileMetrikaDayTask(date=self.date, target_date=target_date, ssid_threshold=self.SSID_THRESHOLD)
            for target_date in days_range_back(self.date, self.DAYS_IN_MONTH)
        ]
        return task_list
```

Appmetrica Data Being Used to Identify Users with Common SSIDS (Wifi Networks)

```
def _run(self):
    self.yt.create_table_with_schema(
        self.destination, self.destination_schema, strict=True, recreate_if_exists=True
    )
    with self.yt.TempTable() as unexploded, self.yt.TempTable() as not_unique:
        self.yql.execute(self.query(unexploded), syntax_version=1)
        run_native_reduce(
            reducer_name="NCommonWifiAP::TExploder",
            source=unexploded,
            destination=not_unique,
            proxy=self.yt.proxy,
            transaction=self.yt.transaction_id,
            pool=conf.Yt.POOL,
            title="Explode yandexuids with common wifi access point",
            reduce_by=[{"ssid"}],
        )
        yuid_pair = [conf.Constants.YUID_LEFT, conf.Constants.YUID_RIGHT]
        self.yt.run_sort(not_unique, not_unique, sort_by=yuid_pair)
        run_native_reduce(
            reducer_name="NCommonWifiAP::TUnique",
            source=not_unique,
            destination=self.destination,
            proxy=self.yt.proxy,
            transaction=self.yt.transaction_id,
            pool=conf.Yt.POOL,
            title="Make yandexuids with common wifi access point unique",
            reduce_by=yuid_pair,
        )
        self.yt.run_sort(self.destination, sort_by=yuid_pair)
    self.yt.set(self.destination + "/@generate_date", self.date.isoformat())
```

Sources

```
graph > fuzzy > lib > config.py > GeoPaths
59
60     class SourceTypes(object):
61         EMAIL_LOGIN = "EMAIL_LOGINS"
62         EMAIL_SIMILAR = "EMAIL_SIMILAR"
63         GEO_HOMEWORK = "GEO_HOMEWORK"
64         HOUSEHOLD = "HOUSEHOLD"
65         REQANS_LOG = "REQANS_LOG" ← Search Data
66         SSID = "SSID" ← Wifi
67
68
```

Yandex IDs Associated with Email

```
class EmailPaths(object):
    ROOT = ROOT
    # Emails
    BASE = "{root}/email".format(root=ROOT)

    ALL_EMAILS_TABLE = "{base}/all_emails".format(base=BASE)
    ALL_EMAIL_LOGINS_TABLE = "{base}/all_email_logins".format(base=BASE)
    ALL_EMAILS_SORTED_BY_LOGIN = "{base}/all_email_logins.sorted_by_login".format(base=BASE)
    ALL_EMAIL_LOGINS_PAIRS_TABLE = "{base}/all_email_logins.pairs".format(base=BASE)
    ALL_EMAILS_GROUPED_BY_LOGIN = "{base}/all_email_logins.groups".format(base=BASE)
    ALL_YUID_PAIRS_FROM_EMAIL_LOGIN = "{base}/all_yuid_pairs_from_email_logins_matching".format(base=BASE)
    ALL_YUID_PAIRS_FROM_SIMILAR_EMAILS = "{base}/all_yuid_pairs_from_similar_emails".format(base=BASE)

    ALL_EMAILS_TABLE_SCHEMA = {"email": "string", "yuids": "any"}
    ALL_EMAIL_LOGINS_TABLE_SCHEMA = {"login": "string", "email": "string", "yuids": "any"}
    ALL_EMAILS_SORTED_BY_LOGIN_SCHEMA = {"login": "string", "email": "string", "yuids": "any"}
    ALL_EMAILS_GROUPED_BY_LOGIN_SCHEMA = {"login": "string", "all_emails": "any", "howmany": "uint64"}
    ALL_EMAIL_LOGINS_PAIRS_TABLE_SCHEMA = {
        "email_1": "string",
        "email_2": "string",
        "login": "string",
        "yuids_1": "any",
        "yuids_2": "any",
    }
    ALL_YUID_PAIRS_FROM_EMAIL_LOGIN_SCHEMA = {
        Constants.YUID_LEFT: ("uint64", True),
        Constants.YUID_RIGHT: ("uint64", True),
        "match": "any",
    }
    ALL_YUID_PAIRS_FROM_SIMILAR_EMAILS_SCHEMA = {
        Constants.YUID_LEFT: "uint64",
        Constants.YUID_RIGHT: "uint64",
        "email_left": "string",
        "email_right": "string",
        "fragment": "string",
    }
```

Login Data

```
graph > fuzzy > lib > tasks > sources > visitlog_logins > extract.py > ...
59     def filter_rare_logins_options(self):
60         return TFILTERRareLoginsOptions(Threshold=self.threshold).SerializeToString()
61
62     @property
63     def filter_keys_options(self):
64         return TFILTERKeysOptions(
65             Keywords=[
66                 "login",
67                 "user",
68                 "userid",
69                 "clientid",
70                 "uid",
71                 "email",
72                 "emailhash",
73                 "\u043b\u043e\u0433\u0438\u043d",
74                 "computerid",
75                 "cid",
76                 "suserid",
77             ]
78         ).SerializeToString()
79
```

} Extracting multiple types of identifiers

Passport

The screenshot shows a web browser window with the URL <https://passport.yandex.com/registration?retpath=https://audience.yandex.com/>. The page features a large red Yandex logo at the top left. Below it are five smaller icons: a red location pin, a white Yandex logo, a red and black document with the letter 'A' and Chinese characters, a yellow envelope, and a blue globe. A text block below these icons states: "With a single account, you can search, write emails, save and share your files, find stuff you want, get directions and use other services on all your devices and platforms". On the right side, there is a registration form titled "Registration". The form includes fields for "First name", "Last name", "Enter a login", "Enter a password" (with a clear button icon), "Confirm password", and "Mobile phone number". At the bottom is a large yellow "Register" button. A link to "Help" and the copyright notice "© 2023, Yandex" are at the very bottom.

Already have an account? [Log in](#)

Registration

First name

Last name

Enter a login

Enter a password

Confirm password

Mobile phone number

Register

Help © 2023, Yandex

Passport User ID Associated with Phone

```
graph > data_import > passport > lib > query > ┌(passport.sql
33
34 $out_login_tbl = $soup_output_dir || $edge(IdType::PUID(), IdType::PHONE(), SourceType::PASSPORT_PROFILE(), LogSource::PASSPORT_PHONE_DUMP())
35 INSERT INTO $out_login_tbl WITH TRUNCATE
36 SELECT
37     id1,
38     IdType::PUID() AS id1Type,
39     id2,
40     IdType::PHONE() AS id2Type,
41     SourceType::PASSPORT_PROFILE() AS sourceType,
42     LogSource::PASSPORT_PHONE_DUMP() AS logSource,
43     ListCreate(String) AS dates
44 FROM (
45     SELECT DISTINCT puid, phone
46     FROM $log FLATTEN LIST BY phones AS phone
47 ) WHERE Identifiers::IsSignificantPhone(phone)
48 GROUP BY
49     puid AS id1,
50     phone AS id2
51 ;
52
```

```
graph > data_import > passport > tests > fixtures > {} passport.json > ...
```

```
1  {"uid": "11111", "login": "aashinova"}  
2  {"uid": "11112", "login": "andrei-ponomareff-1997"}  
3  {"uid": "11113", "login": "anoshko-av"}  
4  {"uid": "11114", "login": "bars12@161.ru"}  
5  {"uid": "11115", "login": "ev0ngertlt"}  
6  {"uid": "11116", "login": "evarcher"}  
7  {"uid": "11117", "login": "lagutin2008"}  
8  {"uid": "11118", "login": "login-for-avito"}  
9  {"uid": "11119", "login": "modsever"}  
10 {"uid": "11120", "login": "mouradian"}  
11 {"uid": "11121", "login": "perschina-olga2013"}  
12 {"uid": "11122", "login": "r.amiraslanov@dveri.ru"}  
13 {"uid": "11123", "login": "saprovec2015"}  
14 {"uid": "11124", "login": "stoltat"}  
15 {"uid": "11125", "login": "sveta-aleshina2015"}  
16 {"uid": "11126", "login": "watchradius"}  
17 {"uid": "11127", "login": "watchradius"}  
18 {"uid": "11128", "login": "sveta-aleshina2015"}  
19 {"uid": "11129", "login": "watchradiusmob"}  
20 {"uid": "123456", "login": "abc123", "phone_numbers": []}  
21 {"uid": "123457", "login": "abc127", "phone_numbers": ["+1234567890", "+71111234567"]}  
22 {"uid": "123458", "login": "", "phone_numbers": ["+9393939393", "+71202020201"]}  
23 {"uid": "134614616", "login": "roscosh8"}  
24 {"uid": "134648582", "login": "e222mn"}  
25 {"uid": "15033290", "login": "mouradian"}  
26 {"uid": "194502233", "login": "ingvr80"}  
27 {"uid": "2687", "login": "govshit"}  
28 {"uid": "2687", "login": "GOVSHIT"}  
29 {"uid": "76667777", "login": "g8jkqqaaaaaaaaah"}  
30 {"uid": "766679666", "login": "d6fqqaaaaaaaaah"}  
31 {"uid": "766679777", "login": "mdmozn45"}
```

32

Geo Graphs

Using lat/long
data
associated
with “predicted
home”, linked
to Yandex UID

```
graph > fuzzy > lib > tasks > sources > geo > C geo_operations.h
61     for (; input->IsValid(); input->Next()) {
62         const auto& row = input->GetRow();
63         if (not IsRowValid(row)) {
64             continue;
65         }
66
67         const ui64 yandexuid = FromString<ui64>(row["yandexuid"].AsString());
68         const auto& homeCoordinates = row["predicted_home"];
69         const auto latitude = homeCoordinates["latitude"].AsDouble();
70         const auto longitude = homeCoordinates["longitude"].AsDouble();
71         const auto& square = computeSquare({.Lat = latitude, .Lon = longitude}, State->radius());
72
73         /**
74          * + + -
75          * + + -
76          * + - -
77          */
78         for (int beltOffset : {-1, 0, 1}) {
79             for (int sqOffset : {-1, 0}) {
80                 if (beltOffset == -1 && sqOffset == 0) {
81                     continue;
82                 }
83                 const ui64 square_idx = ConvertSquareToIdx({.Belt = square.Belt + beltOffset, .Sq = square.Sq + sqOffset});
84                 TGeoSquare out;
85                 out.set_yandexuid(yandexuid);
86                 out.set_lat(latitude);
87                 out.set_lon(longitude);
88                 out.set_squareidx(square_idx);
89                 output->AddRow(out);
90             }
91         }
    }
```

Geo Graphs

Then using that data to find literal neighbors within a certain radius of that home

```
graph > fuzzy > lib > tasks > sources > geo > C geo_operations.h
106
107 class TFindNeighbors : public IReducer<TTableReader<TGeoSquare>, TTableWriter<TNeighborsDistance>> {
108 public:
109     TFindNeighbors()
110         : State()
111     {
112     }
113
114     TFindNeighbors(const TBuffer& buffer)
115         : State(buffer)
116     {
117     }
118
119     void Do(TTableReader<TGeoSquare>* input, TTableWriter<TNeighborsDistance>* output) override {
120         const double radius = State->radius();
121         TVector<TGeoSquare> candidates;
122         for (; input->IsValid(); input->Next()) {
123             const auto& row = input->GetRow();
124             candidates.push_back(row);
125         }
126         for (auto i : xrange(candidates.size())) {
127             for (auto j : xrange(i + 1, candidates.size())) {
128                 const auto& left = candidates.at(i);
129                 const auto& right = candidates.at(j);
130                 if (left.yandexuid() == right.yandexuid()) {
131                     continue;
132                 }
133                 double distance = computeDistance({.Lat = left.lat(), .Lon = left.lon()}, {.Lat = right.lat(), .Lon = right.lon()});
134                 if (distance > radius) {
135                     continue;
136                 }
137                 TNeighborsDistance out;
138                 out.set_distance(distance);
139                 out.set_yandexuidleft(Min(left.yandexuid(), right.yandexuid()));
140                 out.set_yandexuidright(Max(left.yandexuid(), right.yandexuid()));
141                 output->AddRow(out);
142             }
143         }
144     }
145 }
```

Appmetrica and Taxi Data Being Used to Generate Segments about Households with Children

```
profile > runners > segments > lib > coded_segments > with_children_by_ages.py > ...
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
def prepare_with_children_by_app(self, with_children_by_app_table):
    app_segment_name_tuples = []

    for segment_name, apps in SEGMENT_APP_DICT.iteritems():
        for app in apps:
            app_segment_name_tuples.append('AsTuple("{}", "{}").format(app, segment_name)')

    self.yql.query(
        app_metric_query.format(
            devid_by_app_table=self.input()['DevidByApp'].table,
            output_table=with_children_by_app_table,
            app_to_segment_name='\\n'.join(app_segment_name_tuples),
        ),
        transaction=self.transaction,
    )

def build_segment(self, inputs, output_path):
    with self.yt.TempTable() as taxi_puid_table, \
        self.yt.TempTable() as app_metric_table:
        self.yt.run_map(
            extract_children_from_taxi,
            inputs['TaxiData'].table,
            taxi_puid_table,
        )

        self.prepare_with_children_by_app(app_metric_table)

        self.yql.query(
            with_children_query_template.format(
                metrics_table=inputs['ProcessedMetrics'].table,
                reqans_table=inputs['ProcessedReqans'].table,
                app_metric_table=app_metric_table,
                taxi_data_table=taxi_puid_table,
                id_to_crypt_id_table=config.VERTICES_NO_MULTI_PROFILE,
                crypt_id_to_hhid_table=config.HOUSEHOLD_CRYPT_ID_TO_HHID,
                yandexuid_to_hhid_table=config.HOUSEHOLD_REVERSED_TABLE,
                hhid_to_yandexuid_table=config.HOUSEHOLD_ENRICH_TABLE,
                output_table=output_path,
            ),
            transaction=self.transaction,
        )
```

Appmetrica and Taxi Data Being Used to Generate Segments about Households with Children

```
456
457     def build_segment(self, inputs, output_path):
458         with self.yt.TempTable() as taxi_puid_table, \
459             self.yt.TempTable() as app_metrica_table:
460             self.yt.run_map(
461                 extract_children_from_taxi,
462                 inputs['TaxiData'].table,
463                 taxi_puid_table,
464             )
465
466             self.prepare_with_children_by_app(app_metrica_table)
467
```

ID Mapping Associations

```
self.yql.query(  
    with_children_query_template.format(  
        metrics_table=inputs['ProcessedMetrics'].table,  
        reqans_table=inputs['ProcessedReqans'].table,  
        app_metrica_table=app_metrica_table,  
        taxi_data_table=taxi_puid_table,  
        id_to_crypta_id_table=config.VERTICES_NO_MULTI_PROFILE,  
        crypta_id_to_hhid_table=config.HOUSEHOLD_CRYPTA_ID_TO_HHID,  
        yandexuid_to_hhid_table=config.HOUSEHOLD_REVERSED_TABLE,  
        hhid_to_yandexuid_table=config.HOUSEHOLD_ENRICH_TABLE,  
        output_table=output_path,  
    ),  
    transaction=self.transaction,  
)
```

Profiles Integrate Biometric Data, Most Likely from Smart Speakers That Use Yandex's Alice Smart Assistant

Yandex

COMPANY

JOBs

FOR DEVELOPERS

FOR ADVERTISERS

FO

About

History

Privacy

Press Releases

Blog

Contact

Press Releases / 2022 /

Yandex Launches Smart Devices With Alice in Uzbekistan

Internet, November 22, 2022. Uzbekistan's local retailers Stations with Alice. Upon purchasing a smart speaker, us



Stations are smart speakers with Alice. A single Yandex Plus subscription allows you to play music, podcasts, or playlists with personal recommendations on Yandex Station. Alice, the voice assistant on board, can entertain children with an educational game or compose a fairy tale together. She will tell you about the weather or remind you to buy groceries. Alice loves talking and will hold a conversation with ease: currently, in Russian only.

Possible Children by Voice

```
profile > runners > segments > lib > coded_segments > children_age_segment_clarification.py > ...
12
13     clarify_children_yql_template = """
14     $possible_children_by_voice = (
15         SELECT `uuid`, TableName() AS `date`, '0_12' AS segment_name
16         FROM RANGE(`{biometry_folder}`, '{biometry_first_date}', '{biometry_last_date}')
17         WHERE bio_child > 0.8
18     );
19
20     $possible_children_by_voice = (
21         SELECT DISTINCT `uuid` , `date` , segment_name
22         FROM $possible_children_by_voice
23     );
24
25     $possible_children_by_voice = (
26         SELECT `uuid` , segment_name
27         FROM $possible_children_by_voice
28         GROUP BY `uuid` , segment_name
29         HAVING COUNT(*) >= 2
30     );
31
32     $sources_new_age = (
33         SELECT matching.cryptoId AS cryptoId,
34             CASE
35                 WHEN socdem_storage.birth_date > '{thirteenth_birthday}' THEN '0_12'
36                 WHEN '{thirteenth_birthday}' >= socdem_storage.birth_date AND
37                     socdem_storage.birth_date > '{eighteenth_birthday}' THEN '13_17'
38                 ELSE '18_99'
39             END AS segment_name
40         FROM `{socdem_storage_table}` AS socdem_storage
41         INNER JOIN `{id_to_crypto_id_table}` AS matching
42         ON socdem_storage.id == matching.id AND socdem_storage.id_type == matching.id_type
43         WHERE socdem_storage.birth_date is not Null
44         UNION ALL
45         SELECT matching.cryptoId AS cryptoId, biometry.segment_name AS segment_name
46         FROM $possible_children_by_voice AS biometry
47         INNER JOIN `{id_to_crypto_id_table}` AS matching
48         ON biometry.`uuid` == matching.id
49         WHERE matching.id_type == 'uuid'
50     );

```

UI for Infographics Card

```
27
28     const marriedText = convertMarriedToSingleText(exactDemographics.gender, married);
29     const incomeText = convertIncomeSegmentToText(exactDemographics.income);
30     const hasChildrenText = convertHasChildrenToText(hasChildren);
31
32     return (
33         <div className="BasicInfoGraphics">
34             <img alt="" className="BasicInfoGraphics-Image" src={images[exactDemographics.gender]}/>
35             <div className="BasicInfoGraphics-Bubble BasicInfoGraphics-Bubble_family">{marriedText}</div>
36             <div className="BasicInfoGraphics-Bubble BasicInfoGraphics-Bubble_income">{incomeText}</div>
37             <div className="BasicInfoGraphics-Bubble BasicInfoGraphics-Bubble_children">{hasChildrenText}</div>
38             <div className="BasicInfoGraphics-Interest BasicInfoGraphics-Interest_first">
39                 <div className="BasicInfoGraphics-InterestIcon"
40                     style={{ backgroundImage: `url(${interestIcons[0]})` }}/>
41             </div>
42             <div className="BasicInfoGraphics-Interest BasicInfoGraphics-Interest_second">
43                 <div className="BasicInfoGraphics-InterestIcon"
44                     style={{ backgroundImage: `url(${interestIcons[1]})` }}/>
45             </div>
46             <div className="BasicInfoGraphics-Interest BasicInfoGraphics-Interest_third">
47                 <div className="BasicInfoGraphics-InterestIcon"
48                     style={{ backgroundImage: `url(${interestIcons[2]})` }}/>
49             </div>
50         </div>
51     );
52 }
```

- > icons
- > apps
- > interests
 - ─ agro.svg
 - ─ animals.svg
 - ─ appliances.svg
 - ─ beauty.svg
 - ─ business.svg
 - ─ clothes.svg
 - ─ construction.svg
 - ─ education.svg
 - ─ electronics.svg
 - ─ entertainments.svg
 - ─ family.svg
 - ─ finance.svg
 - ─ food.svg
 - ─ gifts.svg
- JS index.js
 - ─ job.svg
 - ─ realty.svg
 - ─ rest.svg
 - ─ sport.svg
 - ─ stationery.svg
 - ─ telecom.svg
 - ─ transport.svg

Search Profile by ID

```
web > portal > src > graph > search > JS SearchPanel.js > ...
112     return (
113         <div key={"inputs-" + suffix} className="input-group">
114             <div style={{display: showIdInput ? "block" : "none"}}>
115                 <ValueInput
116                     key="id_value"
117                     placeholder="[id value]"
118                     value={parameters["idValue" + suffix]}
119                     onChange={changeParameter("uid" + suffix)}>
120             </ValueInput>
121         </div>
122     )
123
124     return (
125         experiments.status !== 403 && (
126             <div className="experiments-bar">
127                 <div className="experiments-select-uid-type">
128                     <RadioButton
129                         value={activeUid}
130                         size="s"
131                         view="default"
132                         className="select-sorting"
133                         onChange={(event) => selectUidType(event.target.value)}
134                         options={[
135                             { value: "uid", children: t("by") + " yandexuid" },
136                             { value: "cryptoId", children: t("by") + " CryptoID" },
137                         ]}
138                 </RadioButton>
139             </div>
140         </div>
141     )
142
143     return (
144         <div>
145             <div>
146                 <div>
147                     <div>
148                         <div>
149                             <div>
150                                 <div>
151                                     <div>
152                                         <div>
153                                             <div>
154                                                 <div>
155                                                     <div>
156                                                         <div>
157                                                             <div>
158                                                                 <div>
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
389
390
391
392
393
394
395
396
397
398
399
399
400
401
402
403
404
405
406
407
408
409
409
410
411
412
413
414
415
416
417
418
419
419
420
421
422
423
424
425
426
427
428
429
429
430
431
432
433
434
435
436
437
438
439
439
440
441
442
443
444
445
446
447
448
449
449
450
451
452
453
454
455
456
457
458
459
459
460
461
462
463
464
465
466
467
468
469
469
470
471
472
473
474
475
476
477
478
479
479
480
481
482
483
484
485
486
487
488
489
489
490
491
492
493
494
495
496
497
498
499
499
500
501
502
503
504
505
506
507
508
509
509
510
511
512
513
514
515
516
517
518
519
519
520
521
522
523
524
525
526
527
528
529
529
530
531
532
533
534
535
536
537
538
539
539
540
541
542
543
544
545
546
547
548
549
549
550
551
552
553
554
555
556
557
558
559
559
560
561
562
563
564
565
566
567
568
569
569
570
571
572
573
574
575
576
577
578
579
579
580
581
582
583
584
585
586
587
588
589
589
590
591
592
593
594
595
596
597
598
599
599
600
601
602
603
604
605
606
607
608
609
609
610
611
612
613
614
615
616
617
618
619
619
620
621
622
623
624
625
626
627
628
629
629
630
631
632
633
634
635
636
637
638
639
639
640
641
642
643
644
645
646
647
648
649
649
650
651
652
653
654
655
656
657
658
659
659
660
661
662
663
664
665
666
667
668
669
669
670
671
672
673
674
675
676
677
678
679
679
680
681
682
683
684
685
686
687
688
689
689
690
691
692
693
694
695
696
697
698
699
699
700
701
702
703
704
705
706
707
708
709
709
710
711
712
713
714
715
716
717
718
719
719
720
721
722
723
724
725
726
727
728
729
729
730
731
732
733
734
735
736
737
738
739
739
740
741
742
743
744
745
746
747
748
749
749
750
751
752
753
754
755
756
757
758
759
759
760
761
762
763
764
765
766
767
768
769
769
770
771
772
773
774
775
776
777
778
779
779
780
781
782
783
784
785
786
787
788
789
789
790
791
792
793
794
795
796
797
798
799
799
800
801
802
803
804
805
806
807
808
809
809
810
811
812
813
814
815
816
817
818
819
819
820
821
822
823
824
825
826
827
828
829
829
830
831
832
833
834
835
836
837
838
839
839
840
841
842
843
844
845
846
847
848
849
849
850
851
852
853
854
855
856
857
858
859
859
860
861
862
863
864
865
866
867
868
869
869
870
871
872
873
874
875
876
877
878
879
879
880
881
882
883
884
885
886
887
888
889
889
890
891
892
893
894
895
896
897
898
899
899
900
901
902
903
904
905
906
907
908
909
909
910
911
912
913
914
915
916
917
918
919
919
920
921
922
923
924
925
926
927
928
929
929
930
931
932
933
934
935
936
937
938
939
939
940
941
942
943
944
945
946
947
948
949
949
950
951
952
953
954
955
956
957
958
959
959
960
961
962
963
964
965
966
967
968
969
969
970
971
972
973
974
975
976
977
978
979
979
980
981
982
983
984
985
986
987
988
989
989
990
991
992
993
994
995
996
997
998
999
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1089
1090
1091
1092
1093
1094
1095
1096
1097
1097
1098
1099
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1187
1188
1189
1189
1190
1191
1192
1193
1194
1195
1195
1196
1197
1198
1199
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1297
1298
1299
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1397
1398
1399
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1497
1498
1499
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1597
1598
1599
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1697
1698
1699
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1797
1798
1799
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1889
1890
1891
1892
1893
1894
1895
1896
1897
1897
1898
1899
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1989
1990
1991
1992
1993
1994
1995
1996
1997
1997
1998
1999
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2089
2090
2091
2092
2093
2094
2095
2096
2097
2097
2098
2099
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2189
2190
2191
2192
2193
2194
2195
2196
2197
2197
2198
2199
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2289
2290
2291
2292
2293
2294
2295
2296
2297
2297
2298
2299
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2389
2390
2391
2392
2393
2394
2395
2396
2397
2397
2398
2399
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2469
2470
2471
2472
2473
2474
2475
2476

```

UI - Available App Icons

```
✓ CRYPTA
  JS TimedInterests.js
  JS utils.js
  > pages
  ✓ public-info
  > components
  ✓ icons
    ✓ apps
      ✓ com.yandex.browser
        ↗ active.svg
        ↗ disabled.svg
      > com.yandex.browser.lite
      > com.yandex.lavka
      > com.yandex.mobile.drive
      > com.yandex.music.auto
      > com.yandex.music.xiaomi
      > com.yandex.toloka.androidapp
    ✓ com.yandex.zen
      ↗ active.svg
      ↗ disabled.svg
      > default
      > ru.yandex.androidkeyboard
      > ru.yandex.androidkeyboard.auto
      > ru.yandex.androidkeyboard.tv
      > ru.yandex.blue.market
      > ru.yandex.disk
      > ru.yandex.disk.notificationserviceext
      > ru.yandex.disk.shareext
      > ru.yandex.lavka
      > ru.yandex.mail
      > ru.yandex.mail.notificationserviceextension
      > ru.yandex.market
```

```
✓ CRYPTA
  > ru.yandex.mail.notificationserviceextension
  > ru.yandex.market
  > ru.yandex.metro
  > ru.yandex.mobile
  > ru.yandex.mobile.drive
  > ru.yandex.mobile.drive.notification
  > ru.yandex.mobile.keyboard
  > ru.yandex.mobile.keyboard.extension
  > ru.yandex.mobile.KeyboardExtension
  > ru.yandex.mobile.metro
  > ru.yandex.mobile.music
  > ru.yandex.mobile.music.push-extension
  > ru.yandex.mobile.music.widget-extension
  > ru.yandex.mobile.navigator
  > ru.yandex.mobile.NotificationService
  > ru.yandex.mobile.search
  > ru.yandex.mobile.toloka
  > ru.yandex.mobile.translate
  > ru.yandex.mobile.weather-v2
  > ru.yandex.music
  > ru.yandex.music.samsung
  > ru.yandex.searchplugin
  > ru.yandex.taxi
  > ru.yandex.telemost
  > ru.yandex.traffic
  > ru.yandex.translate
  > ru.yandex.uber
  > ru.yandex.uber-kz
  > ru.yandex.weatherplugin
  > ru.yandex.yandexmaps
  > ru.yandex.yandexnavi
```

✓ ru.yandex.yandexmaps
✓ ru.yandex.yandexnavi
✓ ru.yandex.ymarket
✓ ru.yandex.ytaxi
JS index.js
✓ interests
 ↗ agro.svg
 ↗ animals.svg
 ↗ appliances.svg
 ↗ beauty.svg

IDs Associated with Social Media Accounts

```
1 import React, { useState } from "react";
2 import { useSelector } from "react-redux";
3 import { getPublicGraph, getPublicGraphLoading } from "../../store/selectors";
4 import { Graph, GraphSkeleton } from "../../components/Graph/Graph";
5 import { Section } from "../../components/Section/Section";
6 import { getServiceIcon } from "../../icons/services";
7 import { getAppIcon } from "../../icons/apps";
8
9 import "./GraphSection.scss";
10
11 import noData from "./no-data.svg";
12
13 const IMAGE_SIZE_XS = 12;
14 const IMAGE_SIZE_S = 36;
15 const IMAGE_SIZE_M = 56;
16 const IMAGE_SIZE_L = 80;
17
18 const NODE_MAPPING = {
19   email: {
20     imageSize: IMAGE_SIZE_M,
21     imageHref: "mail",
22   },
23   yandexuid: {
24     imageSize: IMAGE_SIZE_XS,
25     imageHref: "yandexuid",
26   },
27   idfa: {
28     imageSize: IMAGE_SIZE_L,
29     imageHref: "ios",
30   },
31   gaid: {
32     imageSize: IMAGE_SIZE_L,
33     imageHref: "android",
34   },
35   oaid: {
36     imageSize: IMAGE_SIZE_L,
37     imageHref: "android",
38   },
39   login: {
40     imageSize: IMAGE_SIZE_M,
41     imageHref: "key",
42   },
43   puid: {
44     imageSize: IMAGE_SIZE_M,
45     imageHref: "key",
46   },
47   instagram_login: {
48     imageSize: IMAGE_SIZE_M,
49     imageHref: "instagram"
50   },
51   instagram_id: {
52     imageSize: IMAGE_SIZE_M,
```

```
web > portal > src > public-info > sections > GraphSection > JS GraphSection.js > useEffect
  ...
  51     instagram_id: {
  52       imageSize: IMAGE_SIZE_M,
  53       imageHref: "instagram"
  54     },
  55     fb_id: {
  56       imageSize: IMAGE_SIZE_M,
  57       imageHref: "facebook"
  58     },
  59     ok_id: {
  60       imageSize: IMAGE_SIZE_M,
  61       imageHref: "ok"
  62     },
  63     vk_id: {
  64       imageSize: IMAGE_SIZE_M,
  65       imageHref: "vk"
  66     },
  67     vk_name: {
  68       imageSize: IMAGE_SIZE_M,
  69       imageHref: "vk"
  70     },
  71     kp_id: {
  72       imageSize: IMAGE_SIZE_M,
  73       imageHref: "kinopoisk"
  74     },
  75   },
  76 }
  77 function getNodeMapping(item) {
  78   if (item.idType === 'uuid') {
  79     // App
  80     return { imageHref: item.icon, imageSize: IMAGE_SIZE_S };
  81   }
  82
  83   return NODE_MAPPING[item.icon] ?? { imageHref: "default", imageSize: IMAGE_SIZE_XS };
  84 }
  85
  86 function getImage(item) {
  87   const disabled = !item.isActive;
  88
  89   if (item.idType === 'uuid') {
  90     return getAppIcon(item.imageHref, disabled)
  91       .catch(() => getAppIcon("default", disabled));
  92   }
  93
  94   return getServiceIcon(item.imageHref, disabled)
  95     .catch(() => getServiceIcon("default", disabled));
  96 }
```

Matcher

Matcher

```
✓ matcher
  > bin
  > bundle
  ✓ lib
    > config
    ✓ matchers
      > base_matcher
      > beeline_matcher
      > er_telecom_matcher
      > intentai_matcher
      > mts_matcher
      > rostelecom_matcher
      ≡ ya.make
      ⚡ parser.cpp
      C parser.h
```

Rostelecom Matcher

```
ext_fp > matcher > lib > matchers > rostelecom_matcher > rostelecom_matcher.cpp
 31 void TRostelecomMatcher::AddConnection(const TFpEvent& event) {
 32     auto connection = MakeConnection(event);
 33
 34     Stats.Count->Add("events.incoming.rostelecom.count");
 35     Request += TStringBuilder() << connection.Ip << '\t'
 36             << connection.Port << '\t'
 37             << connection.Timestamp << '\t'
 38             << connection.Domain << '\n';
 39 }
 40
 41 TMatches TRostelecomMatcher::GetMatches() {
 42     if (Request.length() == 0) {
 43         return TMatches();
 44     }
 45     const auto& requestId = CreateGuidAsString();
 46     Log->info("Rostelecom request {} body:\n{}", requestId, Request);
 47
 48     NNeh::TMessage message(GetApiUrl(), "");
 49     Y_ENSURE(NNeh::NHttp::MakeFullRequest(message, "", Request, "text/plain"), "Failed to build request to Rostelecom API");
 50
 51     Stats.Count->Add("api.calls.rostelecom.count");
 52     const auto& resp = MakeRequest(Client, message, TDuration::MilliSeconds(Config.GetApiCallTimeoutMs()), "Rostelecom", requestId, Log);
 53
 54     return ParseResponse(resp->Data);
 55 }
 56
 57 TString TRostelecomMatcher::GetApiUrl() const {
 58     return "post://" + Config.GetApiUrl();
 59 }
 60
 61 TMatches TRostelecomMatcher::ParseResponse(const TString& response) {
 62     TStringInput stringInput(response);
 63
 64     TString line;
 65     TConnection connection;
 66     TString extId;
 67     TMatches matches;
 68
 69     while (stringInput.ReadLine(line)) {
 70         Split(TStringBuf(line), '\t', connection.Ip, connection.Port, connection.Timestamp, connection.Domain, extId);
 71         matches[connection] = TMatchResult{.Status = TMatchResult::EStatus::Found, .ExtId = extId};
 72     }
 73
 74     return matches;
 75 }
```

Rostelecom Matcher

```
ext_fp > matcher > lib > matchers > rostecom_matcher > rostecom_matcher.cpp
31 void TRostelecomMatcher::AddConnection(const TFpEvent& event) {
32     auto connection = MakeConnection(event);
33
34     Stats.Count->Add("events.incoming.rostecom.count");
35     Request += TStringBuilder() << connection.Ip << '\t'
36                                     << connection.Port << '\t'
37                                     << connection.Timestamp << '\t'
38                                     << connection.Domain << '\n';
39 }
40
```

Rostelecom Matcher

```
60
61 TMatches TRostelecomMatcher::ParseResponse(const TString& response) {
62     TStringInput stringInput(response);
63
64     TString line;
65     TConnection connection;
66     TString extId;
67     TMatches matches;
68
69     while (stringInput.ReadLine(line)) {
70         Split(TStringBuf(line), '\t', connection.Ip, connection.Port, connection.Timestamp, connection.Domain, extId);
71         matches[connection] = TMatchResult{.Status = TMatchResult::EStatus::Found, .ExtId = extId};
72     }
73
74     return matches;
75 }
```

Test Result Data

ext_fp > matcher > bin > test > canondata > {} result.json > []test_matcher.test_matcher > {} 1

```
1  {
2      "test_matcher.test_matcher": [
3          {
4              "duid": 16999999761000006,
5              "ext_id": "fake_ertelecom_id_for_5.3.100.0",
6              "ext_source": "ertelecom",
7              "hitlogid": 100506,
8              "ip": "5.3.100.0",
9              "log_type": "bs-watch-log",
10             "logid": 0,
11             "original_domain": "domain-6.ru",
12             "port": 5555,
13             "rtmr_timestamp": 1699999977,
14             "unixtime": 1699999970,
15             "user_agent": "Mozilla/5.0 (Windows NT [PYC])",
16             "watchid": 200000000000000006,
17             "yuid": 1006169999976
18         },
19         {
20             "duid": 16999999861000016,
21             "ext_id": "mts_id_for_160.1.2.4",
22             "ext_source": "mts",
23             "hitlogid": 100516,
24             "ip": "160.1.2.4",
25             "log_type": "bs-watch-log",
26             "logid": 0,
27             "original_domain": "domain-16.ru",
28             "port": 4444,
29             "rtmr_timestamp": 1699999987,
30             "unixtime": 1699999970,
31             "user_agent": "Mozilla/5.0 (Windows NT [PYC])",
32             "watchid": 200000000000000016,
33             "yuid": 1016169999986
34         },
35         {
36             "duid": 16999999931000003,
37             "ext_id": "fake_ertelecom_id_for_5.3.62.0",
38             "ext_source": "ertelecom",
39             "hitlogid": 100503,
40             "ip": "5.3.62.0",
41             "log_type": "bs-watch-log",
42             "logid": 0,
43             "original_domain": "domain-3.ru",
44             "port": 2222,
45             "rtmr_timestamp": 1699999994,
46             "unixtime": 1699999990,
47             "user_agent": "Mozilla/5.0 (Windows NT [PYC])",
48             "watchid": 200000000000000003,
49             "yuid": 1003169999993
50         }
51     ]
52 }
```

Roadmap

- Background on Yandex Leak
- Dive into code:
 - What data Yandex is collecting
 - What Yandex is doing with that data
 - Who Yandex is sharing that data with
- Conclusions and wrap up
- Q&A

Conclusion

What they have:



WiFi SSID



Device information



Location



Search history

What they can determine:



Household composition



Age



Income



Gender



Biometric voice data



E-Mail



Phone number



Children's ages



Interests



Home location



Birth date



App usage data



Service usage data



Physical neighbors



Travel plans/
history



Switching providers



Similar users



Proximity to
local businesses



Smoker



Music
preferences



Home location



Work location



Movie/music
history



E-Mail data

Wrap Up

- Yandex has access to a broad international reach of data and it has been evasive about what it can do with that data
- A small amount of data can say a lot when it is matched to entries from a company's other data sources and analyzed
- Yandex has code to sync some of its data with a Russian-state owned entity

Takeaways

- Anonymization is very easily undone when data gets combined with pools from other sources that may contain identifying data
- Pay attention to who runs your SDKs, what data points they collect, and where they send your user data.
- Who gets access to a company's user data when its assets are sold, the geopolitical climate changes, or a government tightens its control?



Link to Write Up:
(Available 8/10)
bit.ly/455utBP

Q & A



Link to Slides:
(available 8/10)
bit.ly/3qail3p