



SUCCESSING
TOGETHER

Effective use of AEOI data—CRS Tool

Summary

The Common Reporting Standard (CRS), developed in response to the G20 request and approved by the OECD Council on 15 July 2014, calls on jurisdictions to obtain information from their financial institutions and automatically exchange that information with other jurisdictions on an annual basis.

The completeness and correctness of data relies on financial institutions. While financial institutions need to be compliant with CRS requirements and due diligence procedures, but the quality and truthfulness of information heavily relies on the account holders' self-certifications. Since the CRS was a rather new standard, financial institutions often did not have the opportunity to access all the information related to the client to be reported or to prove its accuracy. Those factors result in practical problems associated with automatic matching of data received with the available taxpayer data in internal databases. For example, issues arise when tax identification numbers (TINs) are unavailable, names and addresses cannot be matched, or jurisdiction uses numbers other than TINs.

The FTS of Russia has developed a special software (CRS Tool) that ease the resource burden and facilitating review the CRS information provided by foreign counterparts and match it with the taxpayer's data in our internal databases.

The CRS Tools provides:

1. Automatic identification by means of verification, normalization, and saturation techniques
2. Risk-Profiling of identified CRS records by matching CRS data with Personal Income Tax Returns and CFC databases
3. Risk processing to local tax authorities to conduct tax audits.

A. Background

Since 2018 the FTS of Russia has been participating in the Common Reporting Standard (hereinafter – CRS). On an annual basis, the FTS of Russia receives financial data on the foreign financial institutions’ accounts of Russian tax residents.

	2017	2018	2019	2020
Number of exchange partners from which information was received	59	76	79	73
Number of Financial Accounts reported	556 266	836 557	950 052	1 239 699
Identification rate	84.9%	81.2%	78.5%	83.6%

Table 1. The FTS of Russia CRS background

The FTS of Russia has faced several issues considering identification of raw data received from other jurisdictions:

- low percentage of Tax Identification Numbers (TIN) regarding accounts of individuals (only 50% of accounts are reported with TIN);
- fake TIN data (around 28% of TIN are not valid);
- some reported accounts do not contain dates of birth;
- CRS data is received in the Latin alphabet while all internal databases in Russia are in the Cyrillic alphabet;
- first/last/middle name mixed up;
- wrong dates of birth or date/month mixed up in the reported dates of birth;
- middle names are not reported in CRS while it is an important part of the name in Russia;
- data technical errors (multiplication of accounts or overestimation of generated income amounts).

Dedicated software has been designed to address data quality issues and automate risk assessment. The FTS of Russia has developed a digital solution (CRS Tool) aimed at the automatic identification of CRS data and its risk profiling in 2019.

According to Russian law, all Russian tax residents are subject to income tax on their worldwide income. Generally, personal income is taxed at 13% tax rate that could be adjusted based on the relevant Double Tax Treaty. Besides, Russian tax residents must report controlled foreign companies (CFCs) and pay taxes on the undistributed income of CFCs.

The CRS Tool was meant to identify taxpayers avoiding paying taxes on foreign-sourced and CFC income.

B. Solution

The process of CRS data analysis consists of the following stages:

1. Identification of CRS Records
2. Risk-Profiling
3. Risk-processing to local tax offices

Stages of the business process:

1. Identification Process

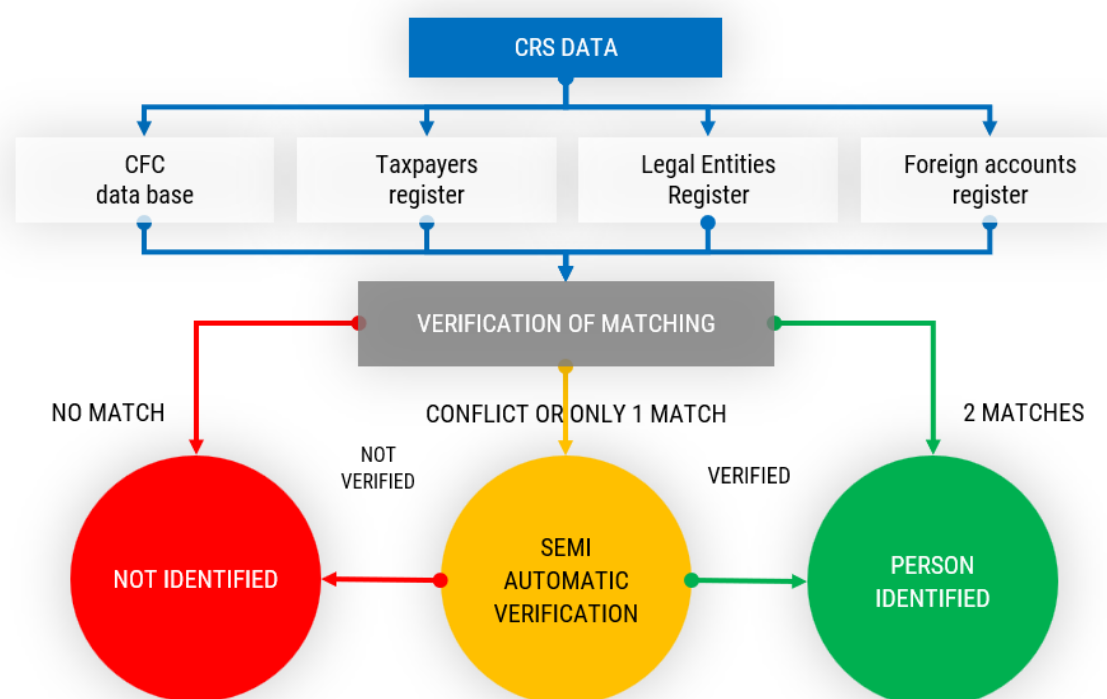


Figure 1. Identification stage

We use the following datasets for identification purposes:

- CFC database
- Individual Taxpayers Register
- Legal Entities Register
- Foreign Account Register

Automatic Matching Facility is the application of the CRS Tool which is operating the following sub-stages:

- (i) transmitting the CRS data and internal dataset to the Data Lake;
- (ii) applying normalization (cleansing) techniques to prepare CRS data for

subsequent matching for identification purposes;

(iii) launching of 16 matching ID algorithms in hard tune (at least two ID elements of the CRS record and dataset record have to be fully matched).

The detailed comments on these stages are provided below.

The CRS Tool assumes that the Account Holder has been identified in case at least two of Account Holder’s ID elements are fully matched. Based on our experience, “Name + BirthDate” combination appears to be the most effective one.

However, if initial CRS data contains valid TIN, the CRS Tool captures it to identify the CRS record.

TIN identification is to be achieved with the following techniques:

- First, we use standard basic normalization techniques which represents simple deletion of letters and symbols in TIN records.




Received TIN	Normalized TIN	
INN 665 901 499 163	665901499163	 Valid TIN after normalization
04 21 456 563 115 C	0421456563115	 Unknown numbers
Passport No: 71 6289698	716289698	 Not a TIN, it’s number of foreign passport

Table 2. Basic normalization

- Secondly, Intelligent normalization techniques are devoted to data mining analysis of non-identified CRS records with TIN element filled in.

Intelligent normalization deals with patterns like replacing letters with a zero if such letters are located between digits or adding a zero to the beginning / end of TIN received.

Received TIN	Basic normalization		Intelligent normalization		
3608136691RU13 08193601	360813669113 08193601		3608136691 1308193601		Not a TIN, it's number of internal passport
51 3346352 21/03/1957	513346352 21031957		513346352		
7727DD748430	7727748430		772700748430		Valid
99C9189335	9999189335		9909189335		Valid IN
24504945308	24504945308		024504945308		Valid

Table 3. Intelligent normalization

- Thirdly, algorithm saturates ID elements of the CRS record for the previous period with subsequent period data, in case other elements (e.g., AcctNumber, FI's name, and IN, Account Holder name) are fully matched.

Reporting Period	Transmitting Country	Account Number	Reporting FI Name	Reporting FI IN	Account Holder Name	TIN/IN	BirthDate
31.12.2018	AB	AB987654321	ABCBank	340.201	Ivanova Tatjana	Copied	
	=	=	=	=	=		
31.12.2019	AB	AB987654321	ABCBank	340.201	Ivanova Tatjana	770871060520	

Table 4. Saturation algorithm

The technique was developed while analyzing the tax risks of a particular taxpayer. There were risks identified for the 2019 year in the CRS Tool but not detected in the 2018 year.

So, the initial CRS data was re-checked and it turned out that unlike 2019, the 2018 CRS record did not contain both TIN and Birth Date.

Identification of Individuals and passive non-financial entities' (NFE) names is to be achieved with the following techniques:

As CRS data is presented in the Latin alphabet, we use a complex combination of transliteration methods (three Russian official standards and one customized) and

matching techniques:

- data normalisation rules

Deletion of default words

“NFN”, “NONE”, “Middle Name”, “Mr”, “Mrs” etc

- first/last/middle name recombinations

Name data received	Last name after separation	First name after separation
IVANOVSTEPAN NFN	IVANOV	STEPAN

- transliteration

Three basic Russian standards	One customized standard by the FTS
ГОСТ 7.79-2000 (ISO 9-95)	<pre> normal: nikolay variants: - nikolai - nikolaj - nicolai - nicolay - nicolaj - nicola - nicolas </pre>
Transliteration standard by the Russian Ministry of Internal Affairs	
Transliteration standard by the Russian Ministry of Internal Affairs	

- typo exclusion rules
- separation rule

Last Name + First Name ± Middle Name	First Name + Last Name ± Middle Name
Last Name + Middle Name + First Name	First Name + Middle Name + Last Name etc

Once all the automatic stages of identification process are accomplished, each CRS record is marked with one of the following statuses:

- Identified
- Non-identified (If there is no match at all, then CRS record is recognized as non-identified)
- To be verified

In case of (i) a conflict between IDs or (ii) only one ID element match, the Account Holder data is taken for semi-automatic verification to **Master of verification mode**.

The Master of verification is the application of the CRS Tool that displays potentially identified taxpayers from our internal dataset that have to be manually verified by the tax inspector.

For example, the Individuals Taxpayer Register contains two taxpayers with same BirthDate, Surname, and First Name, while TIN was not reported in CRS.

This is quite a typical case for widespread Russian Surnames.

The algorithm cannot identify the correct taxpayer in a fully automatic way and thus sends this record to the Master of verification where tax inspector identifies the

correct taxpayer based on the address reported in the CRS record.

Automatic Matching Facility allows us to achieve an average matching rate of **82%** for CRS records received.

2. Risk-Profiling

At this stage, basically the similar technology as at the Identification stage is applied.

Identified CRS records are transmitted to the Data Lake and matched with Personal Income Tax Return data and CFC databases.

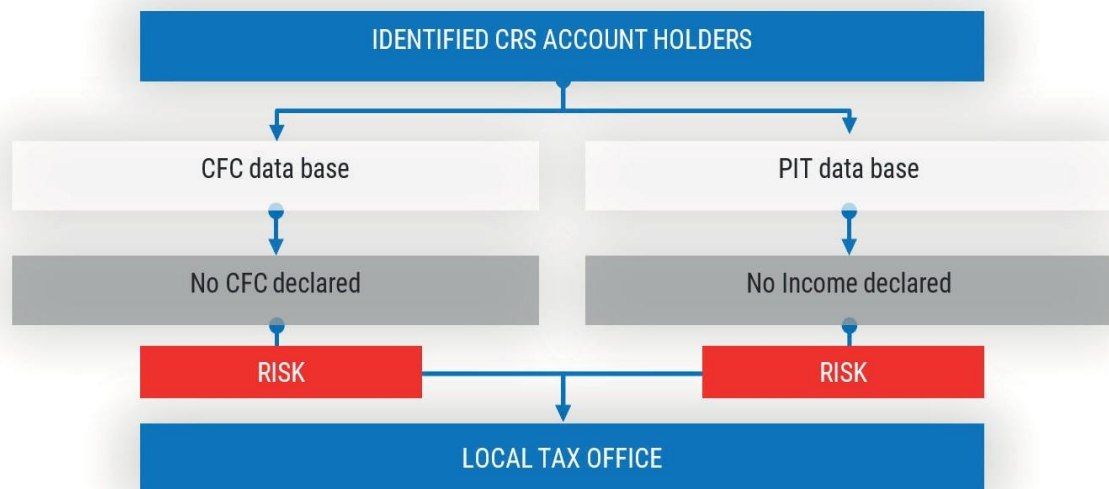


Figure 2. Risk-profiling stage

When the risk is identified, it is delivered to responsible local tax offices for double verification and tax audit measures.

For data protection purposes, local tax officer is granted access to CRS data limited to high risk taxpayers that are registered in the relevant local tax office.

3. Risk-processing

Risk processing is the last stage of the CRS Tool process, where calculated tax risks are being delivered as user tasks to particular local tax offices and monitored by upper-level tax authorities (see Figure 3).

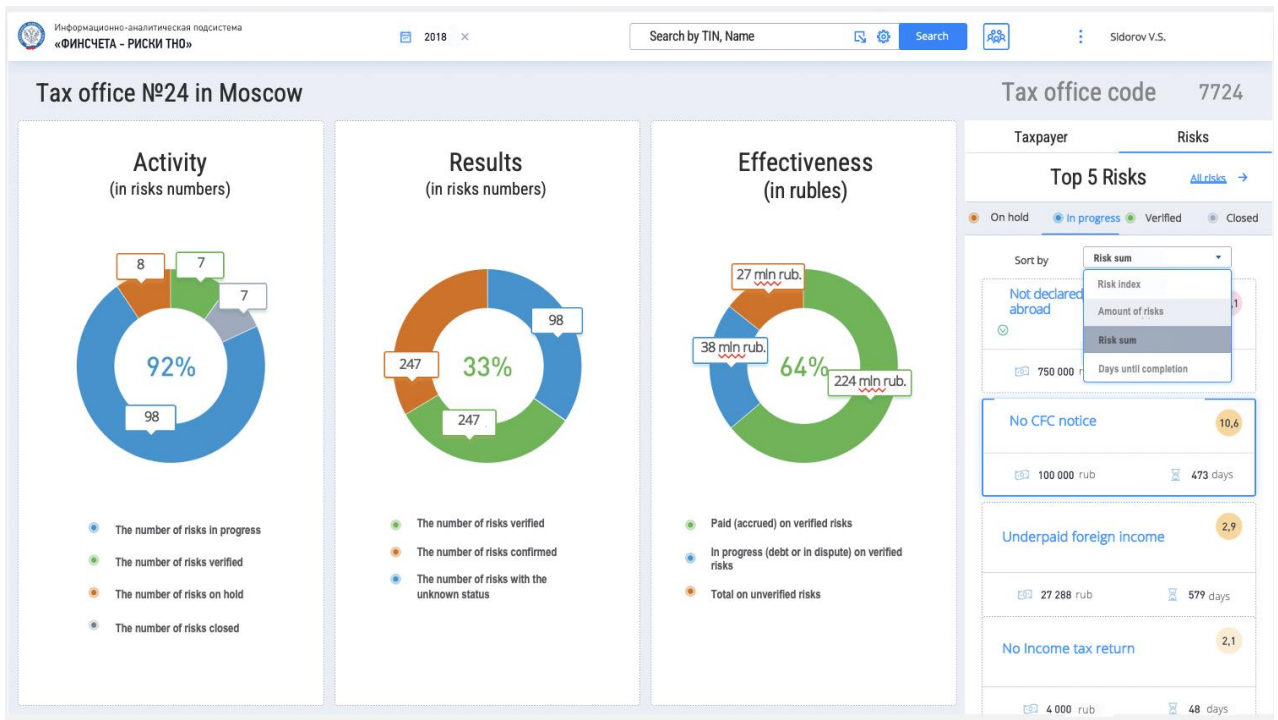


Figure 3. Dashboard of risks life cycle in the tax office

The left block presents the widgets containing information on the activity performed by local office staff, results of automated risk profiling, and effectiveness which is calculated as amounts of taxes recovered (incl. penalties).

The right block presents a chart of TOP-5 taxpayers or TOP-5 risks identified.

The same visualization is used for upper-level officials' dashboards allowing them to monitor the local tax offices' performance.

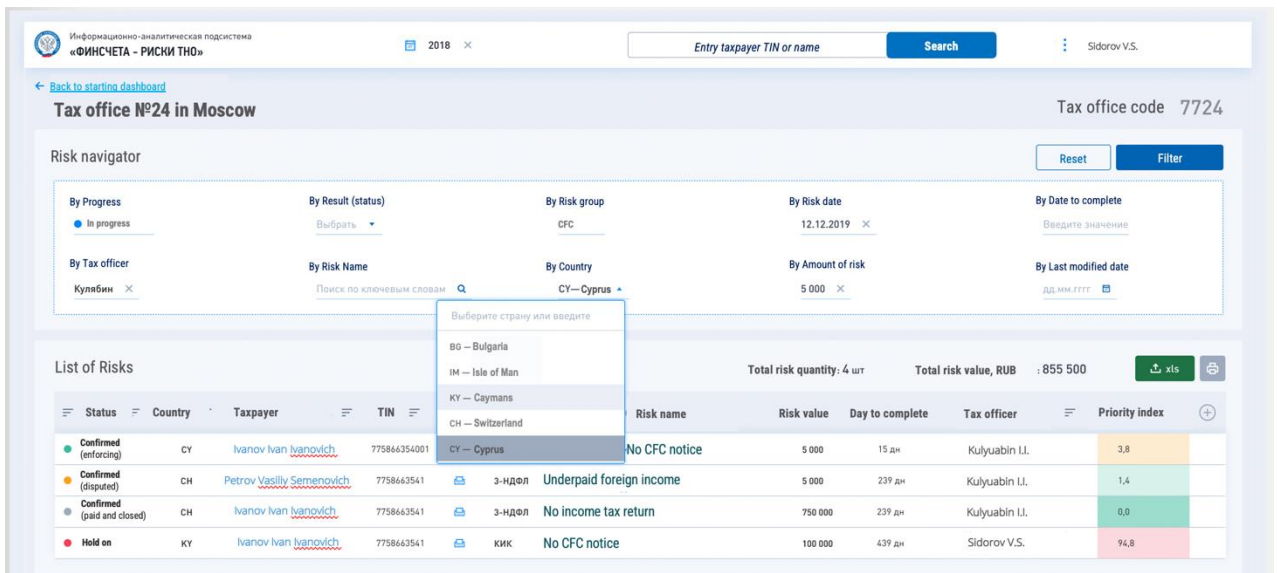


Figure 4. The interface of the Tax Official CRS Tool

The CRS Tool dashboards allows to navigate through identified risks using various filters: by work progress, risk type, risk value, or tax officer who is in charge of the risk, etc.

Risks can also be ranked by the above-mentioned filters and extracted to Microsoft Office Excel.

The risk profile is an analysis of all information on a risky taxpayer that contains in the CRS Tool and internal databases.

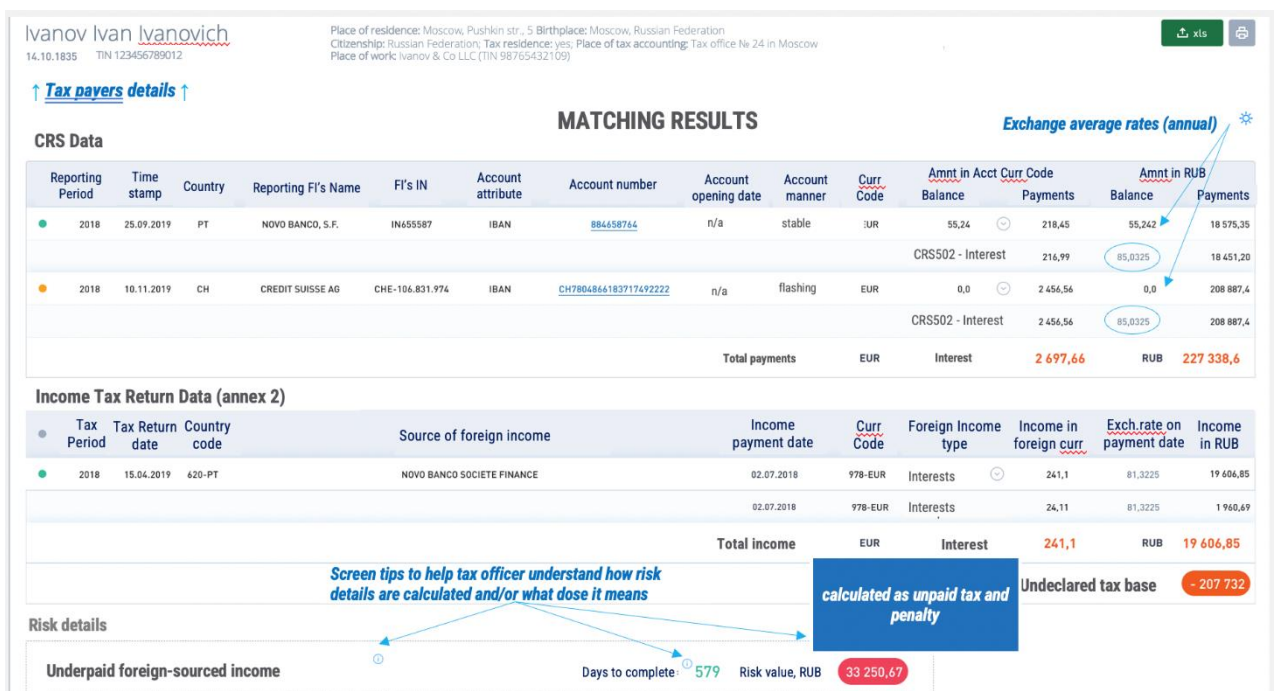


Figure 5. Taxpayer's risk profile

The CRS Tool has a special dashboard interface that presents CRS data and data from the Income Tax Return and allows the tax officer to visually compare data and extract it to Microsoft Office Excel.

The dashboard is divided into several sections:

- Left Panel (Taxpayer Details):**
 - Back to risk list** (link)
 - Ivanov Ivan Ivanovich** (Name)
 - TIN 123456789012
 - Controlling person** (toggle)
 - CC type:** CRS809-account holder
 - Account holder:** CASPER GHOST & PARTNERS BVI
 - TIN:** 1238K
 - Address:** ..., VG, ..., Tortolla, Main Street, 12
 - Account details:**
 - Reporting period:** 31.12.2018
 - Number:** ICFEL-12
 - Acct manner:** stable
 - FI's name:** SNS (Cyprus) limited Кипр
 - FI's IN:** 1238569Q
 - FI's address:** Trust, ..., VG, ..., Tortolla, Main Street, 12
- Central Panel (Risk Details):**
 - Risk details:**
 - No CRC notice:** Code 123
 - Day to complete:** 65
 - Risk value:** 50 000
 - Progress:** Closed (dropdown)
 - Result:** Not verified
 - Risk accuracy:** Low
 - Risk Group:** CRC
 - Risk date:** 12.12.2019
 - Last comment:** Risk is confirmed
 - Last modified date:** 01.02.2020
 - Tax officer:** Kulyuabin I.I.
 - Risk grounds (matching results)** (button)
 - Change progress** (button)
- Bottom Panel (Risk Progress Chronology):**

Date	Tax officer	Action or updates	Comments
12.02.2019	Sidorov V.S.	Responsible tax officer is Kulyuabin I.I.	
12.02.2019	Kulyuabin I.I.	Document is uploaded	
12.02.2019	Kulyuabin I.I.	Risk value is 50 000	
23.02.2019	Kulyuabin I.I.	Risk is confirmed	Documents are uploaded
- Uploaded documents:**
 - Объяснительная записка.pdf
 - Объяснительная записка.pdf
 - Объяснительная записка.pdf
 - Объяснительная записка.pdf

Figure 6. Taxpayer's dashboard

On the left side, the ID data on taxpayers and details on his foreign account are presented.

The central part of the Interface contains risk details. The tax officer must fill in all follow-up audit activities conducted.

The CRS Tool has a function of uploading of all relevant tax audit documents to quickly assess the results of tax officer tax audit measures.



SUCCESSING
TOGETHER