

WORLD-CHECK REFERENCE GUIDE

Data File Field Guide

For formatting differences between CSV fields and XML elements, see Appendix 6.

Field	Description	CSV Field Position	Type	Max Length* (varies)	Value Count	Separator
UID	Each record is identified by the Unique Identified (UID). The UID value is an integer and it is not reused for any other record. Mandatory field	1	Integer	7	1	N/A
LAST NAME	For Individuals, value will be the primary last name of the individual. For entities, this will be the full entity name. Names will always appear in all capital letters. Mandatory field	2	Text	230	1	N/A
FIRST NAME	For Individuals, value will be the primary (most reported) first & middle name(s). Names appear in mixed case. For Entities, this field will be empty. For individual with a single name (no distinction between first and last name), this field will contain a dash (-). This applies to individuals in Myanmar, Indonesia, Timor-Leste & India.	3	Text	125	0-1	N/A
ALIASES	Field will contain all AKAs and FKAs for both individuals and entities. Entities will appear in all capital letters. For individuals, may contain last names, first names or both. Contains names which differ partially or completely from name contained in primary name fields, or contains names in addition to name contains in primary name fields.	4	Text	18,546	0+	;
LOW QUALITY ALIASES <i>(Only Available in Premium Datafile)</i>	Single names, abbreviated names, descriptive names, nicknames, and other noms de guerre; or other aliases not useful for screening, filtering, or searching. Aliases in sanction records will only be considered LQA if reflected as such by all major sanction listings.	5	Text	363	0+	;
ALTERNATIVE SPELLING	Alternative spellings for primary name. May contain last names, first names, or both. Does not contain any names in addition to those listed in primary name fields (LAST NAME / FIRST NAME).	6	Text	251	0+	;
CATEGORY	One or two word description of the main justification for entry being listed within World-Check. See WCDF Reference Section for full list of categories. Mandatory population field.	7	List	21	1	N/A
TITLE	Royal, military, or academic title. For example, General, Baroness, Commander,	8	Text	34	0-1	N/A

	Doctor, Judge, Professor					
SUB-CATEGORY	Specifies whether the subject holds or held a Politically Exposed Person (PEP) position. This also includes State Owned Entities (SOE), immediate relatives and close associates. See WCDF References section for list of sub-categories.	9	List	9	0-1	N/A
POSITION	Social or political, position(s) (not occupation), such as Ambassador, Alderman, Cabinet Minister, Director, etc. Only applicable to PEP and sanction records.	10	Text	255	0+	;
AGE	If DOB is not available when this field provides the age as reported from source, where available. Used in conjunction with AGE DATE (AS OF DATE) field. If DOB is available, then AGE field will be empty.	11	Integer	3	0-1	N/A
DOB	Date of birth, specified in the format YYYY/MM/DD in CSV file and YYYY-MM-DD in XML field. Month and/or day can be zeros ('00') if that information is unavailable.	12	Date	10	0-1	N/A
DOBS	Multiple dates of birth, specified in the format YYYY/MM/DD in CSV file and YYYY-MM-DD in XML file.	13	Text	175	0+	;
PLACE OF BIRTH	All reported place(s) of birth for the individual (city, state/province, country where available). Blank for entities.	14	Text	255	0+	;
DECEASED	Where applicable, the date of death of the individual. Specified in YYYY or YYYY/MM/DD format in XML file. Month and/or day can be zeros ('00') if that information is unavailable. Field is blank for entities.	15	Date	10	0-1	N/A
PASSPORTS	Passport number(s), were unknown, for individuals. In CSV file the country of issuance is denoted in parentheses. In XML file the country of issuance is an attribute of the passport element.	16	Text	649	0+	;
SSN	Social Security Number or Tax ID number, where known. For example, US Social Security Number (SSN) or Brazilian Taxpayer Identification Number (CPF).	17	Text	23	0+	;
IDENTIFICATION NUMBER	Identification Numbers from national international authorities. Each ID number is tagged as {XXX-YYYYY}, where XXX is the two-character ISO 3166 country code, or INT or UNK, and YYYYY is the abbreviation for the ID type. See WCDF Reference section for ID Type List.	18	Text	74868	0+	;
LOCATIONS	Location(s) (city, state/province, country where available) where the individual or entity is active. For example, residence company HQ, base of the operations. If no locations are known this field will contain a comma surrounded by tildes (~,~). See Appendix 3 for Location field formatting.	19	Text	6,977	0+	;
COUNTRIES	For individuals, primary country of nationality or citizenship, extracted from the	20	List	44	1	N/A

	citizenship field. For entities, the country where entity is registered. Countries are listed in uppercase and (mostly) follow the ISO 3166 English short name standard. SEE WCDF Standards and References document for World Check country list and ISO 3166 list. Mandatory field.					
CITIZENSHIP	For individuals, country/countries of nationality or citizenship. For entities, the country where entity is registered. See Standards and References document for country list. Mandatory field.	21	List	59	1+	N/A
COMPANIES	Companies, organizations, and other entities to which the record is understood to be related. Companies can be listed as either the UID of the related record or the name of the company if it is not profiled in World-Check.	22	Text	22,982	0+	;
E/I	Specifies if the record is an Entity (E) or an Individual (M,F or U). Individuals are specified as Male (M), Female (F) or Unknown (U). 'I' is deprecated value. Mandatory field.	23	List	1	1	N/A
LINKED TO	Individual(s) to whom the record is understood to be related, listed as the UID(s) of the related record(s).	24	Integer	8,441	0+	;
FUTHER INFORMATION	A semi-structured text field containing biographical and further identifying information on profiled entry, relationships, negative or risk reports, as well as full text of all sanction entries (where applicable), as well as background supporting, relevant information. Mandatory field. See Appendix 4 for FI field formatting.	25	Text	110,395	1	N/A
KEYWORDS	List of all sanctions, regulatory and law enforcement lists on which the record has been listed, as well World-Check defined lists. Keywords are represented as abbreviations. The abbreviation mapping can be found in the WCDF Reference section.	26	List	507	0+	~
EXTERNAL SOURCES	List of all public sources, including both government and media sources, used to build the full World-Check record. Mandatory field.	27	URL	102,936	1+	<space>
UPDATE CATEGORY <i>(Only available in Premium Datafile)</i>	Categorization of most recent update. Updates are categorized into 5 groups: C1-C4 & C6. See Reference section for category definitions.	28	List	2	0-1	N/A
ENTERED	Date the record was initially created and entered into World-Check specified in YYYY/MM/DD format in CSV file and YYYY-MM-DD in XML file. Mandatory field.	29	Date	10	1	N/A
UPDATED	Date the record was last updated, specified in YYYY/MM/DD format in CSV file and YYYY-MM-DD format in XML file. For records that have not been updated since creation, this field will be blank.	30	Date	10	0-1	N/A

EDITOR	Deprecated; always blank.	31		0	1	N/A
AGE DATE (AS OF DATE)	Provides the date when the AGE value was valid, specified in YYYY/MM/DD format in CSV file and YYYY-MM-DD format in XML file.	32	Date	10	0-1	N/A
PEP ROLES (Only Available in Premium Datafile)	<p>A set of subfields that provides structured information about the current and former PEP roles held by the PEP. See Appendix 5 for PEP ROLES field formatting.</p> <p><u>PEP ROLES Subfield Definitions:</u> PEP Role – the social or political job/role held by the PEP per country or position naming conventions. PEP Role Level – the PEP hierarchy per level or sphere of impact within political or government structures of the PEP job/role. PEP Position – the standard World-Check PEP position held by the PEP. PEP Role Start Date – the start date of the term or office in the role. An additional item is also provided to indicate whether the start date is Reported or Confirmed. If term start date is not known field will be populated with 'UNKNOWN.' PEP Role End Date – the end date of the term of office in the role. An additional item is also provided to indicate whether the end date is Reported or Confirmed. Field will be empty until the term end date occurs, at which time the end date will be populated if it is available. PEP Role Status – indicates whether role is Current, Former, or Unknown. If term end date has not occurred and term start date & length of term are known, then PEP Role Status will be derived as 'Current' or 'Former'. If term end date has not occurred and term start date is known but length of term is known, then PEP Role Status will be 'Unknown'. PEP Role Bio – denotes the specific title of the job role and includes a narrative about the position held.</p>	33	Text	10,554	0+	~
PEP STATUS (Only Available in Premium Datafile)	<p>The overall status of the PEP, which will be active, Inactive or Unknown. The below conditions for PEP status can be manually overridden by researchers for edge cases; e/g. royalty who don't have start/end dates.</p> <p><u>For ALL PEPS</u> Active: If atleast one of the PEP Roles Status is Current</p> <p><u>For Primary PEPS</u> Inactive: If none of the PEP Role Status is Current or Unknown, and atleast one of the PEP Roles Status is Former</p>	34	List	8	0-1	N/A

	<p>Unknown: If none of the PEP Roles Status is Current, and atleast one of the PEP Roles Status is Unknown</p> <p><u>For PEP Associates and Relatives</u> Active: If atleast one of the linked Primary PEP's PEP Role Status is Active Inactive: If none of the linked Primary PEP's PEP Role Status is Active or Unknown, and atleast one of the Primary PEP's PEP Role Status is Inactive Unknown: If none of the linked Primary PEP's PEP Role Status is Active, and atleast one of the Primary PEP's PEP Role Status is Unknown</p>					
SPECIAL INTEREST CATEGORY <i>(Only Available in Premium Datafile)</i>	Records in relation to export controls and trade restrictions; sanctions related advisory notices; sanctions seizure/freezing orders; sanctions/embargo related vessels, aircraft, seaports, and airports; North Korea and Iran related date within World-Check	35	List	531	0+	;

*Max Length is current as-of document date and is subject to change; no advance notice is given when length changes.

APPENDIX 1 - VALUE COUNT DEFINITIONS

Value Count	Definition
0	Deprecated field. No value will appear.
1	One and Only one value.
0-1	Zero or one value
0+	Zero or more values
1+	One or more values

APPENDIX 2 - TYPE DEFINITIONS

Type	Definition
Unique	Unique value. Guaranteed to be unique within the entire database
Text	Any set of characters that can be expressed in ISO/IEC 8859-15 encoding
List	A selection from a set of values. See Reference Section for set.
Integer	Positive Integer. Can be assumed to be less than 2^{31} .
URL	URL that complies with the RFC 3986 standard
Date	Date expressed as YYYY/MM/DD. In certain cases, MM and/or DD may be zeros ('00').

APPENDIX 3 - LOCATIONS FIELD FORMATTING

<Street Address><SPACE>~<SPACE><City>,<SPACE><State/Region><SPACE>~<SPACE><COUNTRY>

Examples:

Full Address is known	64 Honeybadger Road ~ Borrowdale, Harare ~ ZIMBABWE
City, State, and Country are known	~ New York City, New York ~ USA
City and Country are known	~ St John's,~ ANTIGUA & BARBUDA
State and Country are known	~, State of Missouri ~ USA
Only Country is known	~,~ MOLDOVA

APPENDIX 4 – FURTHER INFORMATION FIELD FORMATTING

[TEXT TAG]<SPACE><Text><SPACE> *optionally*: [TEXT TAG]<SPACE>[TEXT TAG]<SPACE><Text><SPACE>

Examples:

[BIOGRAPHY] African Union peacekeeping mission established by the United Nations Security Council (Feb 2009 -). [IDENTIFICATION] OfficeTitle: LocationName (Oct 2017 -). PositionTitle: LocationName (Mar 2019 -). [REPORTS] Oct 2013 - reportedly listed in an United Nation report to be using child soldiers in its ranks. Investigations opened. Jul 2015 - reportedly using child soldiers during operations and in its ranks. Jun 2016 - listed in the United Nation report to be using child soldiers in its ranks. Nov 2019 - no further information reported.

[STATE CONTROLLED] [BIOGRAPHY] Organisation established under the Act on the Establishment of Financial Supervisory Organisations to examine and supervise financial institutions (1999-). [IDENTIFICATION] OfficeTitle: CityName (May 2018 -). [REPORTS] Nov 2019 - no further information reported.

APPENDIX 5 – PEP ROLES FIELD FORMATTING

<PEP Role>|<PEP ROLE LEVEL>|<PEP Position>|<PEP Role Start Date>|<PEP Role Start Date Reported>|<PEP Role End Date>|<PEP Role End Date Reported>|<PEP ROLE STATUS>|<PEP Role Bio>

Examples:

Role End Date has not occurred (and length of term is known)	Counsellor PEP N Senator 2016/00/00 REPORTED CURRENT Counsellor
Role Start Date is not known	Mayor PEP L Mayor UNKNOWN 2006/00/00 CONFIRMED FORMER Mayor of Burlington
All information is available	Cabinet Minister PEP SN Member of Executive 2015/10/00 CONFIRMED 2017/10/00 CONFIRMED FORMER Provincial Minister of Children and Youth Services and Minister responsible for Women's Issues

APPENDIX 6 – CSV AND XML FORMATTING DIFFERENCES

Field/Element	Difference	Examples
ALIASES	CSV alias names are separated by semi-colons. XML alias names are listed as separate elements within aliases parent element.	CSV: LASTNAME1,Firstname1;LASTNAME2,Firstname2 XML: <aliases> <alias>LASTNAME1,Firstname1</alias> <alias>LASTNAME2,Firstname2</alias> </aliases>
LOW QUALITY ALIASES	CSV low quality alias names are separated by semi-colons. XML low quality alias names are listed as separate elements within low quality aliases parent element.	CSV: NAME1;Name2;Name3 XML: <low_quality_aliases> <alias>NAME1</alias> <alias>Name2</alias> <alias>Name3</alias> </low_quality_aliases>
DOB	CSV date of birth is formatted as YYYY/MM/DD. XML date of birth is formatted as YYYY-MM-DD.	CSV: 1960/10/26 XML: 1960-10-26
DOBS	CSV date(s) of birth are formatted as YYYY/MM/DD and separated by semi-colons. XML date(s) of birth are formatted as YYYY-MM-DD and are listed as separate elements within dobs parent element.	CSV: 1980/06/13;1949/00/00 XML: <dobs> <dob>1980-06-13</dob> <dob>1949-00-00</dob> </dobs>
DECEASED	CSV deceased date is formatted as YYYY or YYYY/MM/DD. XML deceased date is formatted as YYYY or YYYY-MM-DD.	CSV: 2016/06/02 XML: 2016-06-02

PASSPORTS	<p>CSV passports are separated by semi-colons and the country name, if available, appears in parenthesis after the passport value. XML passports are listed as separate elements within passports parent element and the country name, if available, is an attribute of the passport element.</p>	<p>CSV: 00999999 (TUNISIA);K1111111;K111111 (TUNISIA)</p> <p>XML: <passports> <passport country="TUNISIA">00999999</passport> <passport country="">K1111111</passport> <passport country="TUNISIA">K111111</passport> </passports></p>
IDENTIFICATION NUMBERS	<p>CSV identification numbers are separated by semi-colons with the country code and identification code displayed in curly braces before the value. XML identification numbers are listed as separate elements within id_numbers parent element, with country code and identification code displayed as attributes of the id</p>	<p>CSV: {BR-CPF}312.933.957-49;{BR-RG}362985 MB/RJ</p> <p>XML: <id_numbers> <id loc="BR" type="CPF">312.933.957-49</id> <id loc="BR" type="RG">362985 MB/RJ</id> </id_numbers></p>
LOCATIONS	<p>Locations are addresses comprised of street & city, state & country, in whole or in part.</p> <p>CSV address components are separated with tildes, with each new address separated by a semi-colon. XML addresses are displayed as separate elements within locations parent element, with country, city, and state displayed as attributes and the street address displayed as the value of the location element.</p>	<p>CSV: ~ Cotabato City, Soccsksargen ~ PHILIPPINES;Street n100, Wazir Akbar Khan ~ Kabul, Kabul ~ AFGHANISTAN</p> <p>XML: <locations> <location country="PHILIPPINES" city="Cotabato City" state="Soccsksargen" xsi:nil="true"/> <location country="AFGHANISTAN" city="Kabul" state="Kabul">Street n100, Wazir Akbar Khan</location> </locations></p>
CITIZENSHIP	<p>CSV citizenships are separated by semi-colons. XML citizenships are listed as separate elements within citizenships parent element.</p>	<p>CSV: RUSSIAN FEDERATION;UKRAINE;UNKNOWN</p> <p>XML: <citizenships> <citizenship>RUSSIAN FEDERATION</citizenship> <citizenship>UKRAINE</citizenship> <citizenship>UNKNOWN</citizenship> </citizenships></p>
COMPANIES	<p>CSV linked company names/UIDs are separated by semi-colons. XML linked company names/UIDs are listed as separate elements within citizenships parent element.</p>	<p>CSV: 785008;9603</p> <p>XML: <companies> <company>785008</company> <company>9603</company> </companies></p>
LINKED TO	<p>CSV linked individual UIDs are separated by semi-colons. XML linked individual UIDs are listed as separate elements within linked_to parent element.</p>	<p>CSV: 1125226;4964757</p> <p>XML: <linked_to> <uid>1125226</uid> <uid>4964757</uid> </linked_to></p>

KEYWORDS	CSV keywords are separated by tildes. XML keywords are listed as separate elements within keywords parent element.	CSV: DFAT~DFAT-UN~EU XML: <keywords> <keyword>DFAT</keyword> <keyword>DFAT-UN</keyword> <keyword>EU</keyword> </keywords>
EXTERNAL SOURCES	CSV source links are separated by spaces. XML source links are listed as separate elements within external_sources parent element.	CSV: http://sourcelink1.htm http://sourcelink2.html XML: <external_sources> <uri>http://sourcelink1.htm</uri> <uri>http://sourcelink2.htm</uri> </external_sources>
ENTERED	CSV entered date is formatted as YYYY/MM/DD. XML entered date is formatted as YYYY-MM-DD.	CSV: 2013/02/01 XML: 2013-02-01
UPDATED	CSV updated date is formatted as YYYY/MM/DD. XML updated date is formatted as YYYY-MM-DD.	CSV: 2019/11/25 XML: 2019-11-25
AGE DATE (AS OF DATE)	CSV age date is formatted as YYYY/MM/DD. XML age date is formatted as YYYY-MM-DD.	CSV: 2003/07/29 XML: 2003-07-29
PEP ROLES	Each CSV PEP role is comprised of the sub-fields identified in Appendix 5. Each sub-field is separated by a pipe, with a new PEP role indicated by a semi-colon. The PEP Status is a separate field. Each XML PEP role is displayed as a set of elements within the pep_role_detail parent element. The pep_role_details parent element contains one pep_status element and one or more pep_role_detail children. CSV: President PEP N Head of Government 2000/07/00 CONFIRMED 2007/07/00 CONFIRMED FORMER President of Syria~President PEP N Head of Government 2007/07/00 CONFIRMED 2014/06/00 CONFIRMED FORMER President of Syria XML: <pep_role_details> <pep_status>active</pep_status> <pep_role_detail> <pep_role>President</pep_role> <pep_role_level>PEP N</pep_role_level> <pep_position>Head of Government</pep_position> <pep_role_status>former</pep_role_status> <pep_role_term_start_date term_start_date_reported="confirmed" term_start_date_unknown="false">2000-07-00</pep_role_term_start_date> <pep_role_term_end_date term_end_date_reported="confirmed">2007-07-00</pep_role_term_end_date> <pep_role_bio>President of Syria</pep_role_bio> </pep_role_detail> <pep_role_detail> <pep_role>President</pep_role> <pep_role_level>PEP N</pep_role_level> <pep_position>Head of Government</pep_position> <pep_role_status>former</pep_role_status> <pep_role_term_start_date term_start_date_reported="confirmed" term_start_date_unknown="false">2007-07-00</pep_role_term_start_date> <pep_role_term_end_date term_end_date_reported="confirmed">2014-06-00</pep_role_term_end_date> <pep_role_bio>President of Syria</pep_role_bio> </pep_role_detail> </pep_role_details>	