Best Practices to Improve Data Quality of UN/LOCODE for the UN/LOCODE Focal Points

Drafted by Stefano Sabatini

In this presentation:

- **1. Error detection**
- 2. Metadata improvements

Task 1: Preparing the file for working

Approach 1

- 1. Download CSV package from the UNECE Website
 - a. Go to http://www.unece.org/cefact/codesfortrade/codes index.html
 - b. Click [cvs]
 - c. Open http://www.unece.org/fileadmin/DAM/cefact/locode/loc161csv.zip
- 2. Select the Code List part with the country you're dealing with in the unzipped package.
 - (i.e. 2016-1 UN/LOCODE CodeListPart1.csv)
- 3. Open it in a spreadsheet software (like Microsoft Excel or Libreoffice Calc)
- 4. When importing it, you need to:
 - a. Choose 'Windows-1252' as character encoding (when you notice garbled characters caused by diacritics)
 - b. Add the header in the first row
 - Copy the following texts in the first cell

Change, Country, Location, Name, NameWoDiacritics, Subdivision, Function, Status, Date, IATA, Coordinates, Remarks

- Split the rows in columns ("Text to columns" function) using comma (,) as separator

c. When being asked to choose the Data format for each column, select the Data column and explicitly set it to 'Text' (this solves the disappearance of the leading zeroes in the column); do this also for Subdivision and Status (if not for all of them)

About UNIECE Our work	Thernes W	hare we work	Open UNECE	Events	Publications	Media
	UN CENTRE FOR TRADE	FACILITATION AN	D C-BUSINESS			
Trade Programme	Codes for T	rade				
About us	Aprenteal	instant		Title		Deservices
Projects	UNICOCODE	10.01 LIN	COLORIN De Libertor en	Camp (015)-1		Land a
Outputs						6
Meetings and Events						Del TO I
Publications		10%	LOCODE Marsuel			Last - 1
ION UNICEFACT	150 Country an	d Currency Codes				
Policies, Proceduces and		190	3100-1 sourcey codes a	etal surrendes as	eF150 4217	
Terms of Reference	Coded for Units	of Measurement	used to the Internatio	eral Trade		
- National Trade Facilitation		Cov	les with narrow. Rednes Acators	of committees as	Potration loss description	les 1
Marka Chaines						
and capped a						

Name	Size	Packed Size
2016-2 SubdivisionCodes.csv	173 482	44 289
2016-2 UNLOCODE CodeListPart1.csv	4 055 677	860 661
🗟 2016-2 UNLOCODE CodeListPart2.csv	2 079 738	451 584
🕼 2016-2 UNLOCODE CodeListPart3.csv	2 449 795	506 542
2016-2 UNLOCODE SecretariatNotes.pdf	275 849	242 242

Change	Country	Location	Name	NameWol Subdivisi	icFunction		Status	Date IATA	Coordinat Remarks
	CLA.		ANDORRA						
	AD	ALV	Andorra la Vella	Andorra la Vella	34-0		AL	601	4230N 00131E
	AD	CAN	Canillo	Canillo	3		RL .	107	4234N 00135E
	AD	ENC	Encemp	Encamp	3		RL.	307	4232N 00134E
	AD GA	ESC	Escaldes-Engorda	Escaldes-Engordani	y toBring		R3,	807	4231N 00133E
	AD	EAC	Escès	Escas	4 3		RL.	1407	4233N 00131E
	AD	FMO	La Farga de Moter	La Farga de Moles	· · · · · ·	#NAME?	RCL	307	
	AD	LMA	La Massana	La Massena	3		R1,	307	4234N 00129E
	AD.	ORD	Ordino	Ordino	an Barres		HL.	307	4233N 00131E
	AD	PAS	Pas de la Casa	Pas de la Casa		#NAME?	RL	307	4233N 00144E
	AD	SJL	Sant Julià de Lôni	Sant Julià de Loria			83.	1101	4225N 00130E

Task 1: Preparing the file for working

Approach 2:

1. Use the package from

http://data.okfn.org/data/core/un-locode :

- 2. Select Code list and click Download
- 3. When opening it, the character set is UTF-8, the separator is the comma

Note: This package has removed

- the country rows (example: ,"AD",,".ANDORRA",,,,,,,)
- the name equivalence rows

(example: "=","AE","","Ruwais = Ar Ruways","Ruwais = Ar Ruways","","","","")

4. Filter the data (select the first row and apply the Filter function) to select a country in question

() () интелеризование нападать на падать на п Падать на падать на пада На падать на падать на падать на падать на Падать на падать на па		A B	С	D		E	F	G	н	1	1
ANNO, LUNALINE, RUMAL RUMANALIZUMETING, KUMUNUTUN, KUNUNUTUN, KUNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUNUN	1	Y AD Y	-	.ANDC) – IC	Δ 👻	-	-	-	-	
7. 584. Regime, Regime, Tet, 875, 8767, 1957F. Addit.	81	Sort A to Z			1.12	Andorra la	Vella	34-6	ΔI	601	
2. Box Bankard, Bankard, J. J. (1997). A straight of the st	2.*	gontratio				Contillo	· · ciid	2		207	
7, 577, 5, 5, 57,	∕A↓	Sort Z to A				Canillo		3	RL	307	
s can chang a burga chang ang ang ang ang ang ang ang ang ang		Sort by Color				Encamp		3	RL	307	
n, new Aray, the farmer's Aray, the farmer's strategy to a second strategy for the Aray and the Aray and the second strategy to a		Sol <u>e</u> by color			5-6	Escaldes-E	ingordany	3	RL	307	
a Charles Anna a' an anna Charles Anna Anna Anna Anna Anna Anna Anna Ann	T	Clear Filter From "AD"				Escas	4		RI	1407	
7, CTV, Cheylo Deceo, Cheylo Deceo, 57, eeedeeee, 1567, 54318, 665577,		Filter by Color				Lo Cours du			20	2107	
r den de la presenta de la contra de la contra El presenta de la contra de la co		Finter by Color			a	La Farga de	eiviores	#INAIVIE F	кų	307	
starotarium deiremtarium deiremtarit s		Text <u>Filters</u>		Þ	an	n La Massan	a	3	RL	307	
1. PHD: INVERSION PARAMETER 1 1 (1997).		-		-		Ordino		3	RL	307	
statutave seeneekke side seeneekke väätave at toovat sasye veeree.		Search		2	a	Pas de la C	asa	#NAME?	RI	307	
C. WILL STREE THEORY AND ADDRESS AND ADDRESS AND ADDRESS.		(Select All)		^	1.0	Sant Julia	do Loria	2	DI	1101	
- WILL CAREER, MERRER, MA. V		AD			Id	Sancouna	ue cona		nL.	1101	
S de Distribui de la contra de la		/				Santa Colo	oma	3	RL	307	
Sall Palle and so the sector all the) A	ARAB EMIRA	TES				
r, MTA, The extension of the second		- AG			- au	Abu al Buk	choosh	1	RL	307	
219906_05.001_2001012_241203152015200012_22						Abu Dhab	47	1 245	A1.	1101	
, NUL, REAL NEEDER, NEEDER, LOU, JOIN, JOI					10	Abu bhab	AL	1-345	AI	1101	
P. MIT. MILLER, MILLER, MILLER, MILLER, C.		\			sa	a Abu Musa		1	RL	201	
n, and a second a s		AO /			Bi	i Ahmed Bi	AZ	1	RL	1407	
F, WH, Kicashan, Kicashan, AT, sasaanaa, ABIS,		AQ AQ				Aiman		1-3	RL	103 (LAO
r vale valende andre en regel a state en				+		AL Aim		4	A1.	9601	
a a sa a						ALAIII			AI	5001	
CHERTERENTATION CONTRACTOR CONTRACT		OK	C	ancel	a	AI Unafra		34	AI	507	
					.: ra	a ΔI Fuiavra	FU	1	Δ1	212	

dataokfn.org	Tools Data	Standards	Docs	Contribute	Roadmap	Vision	Contact	OF SHILL CHARGE
UN-LOCODE Cod	delist							
The United Nations Code for Trade and Trae UNLCL, United Nations agency, to had litate Download Data	rsport Locations is a s trade. Data — Read n	ode list mantain Iore	od by	Ť	internality Source UNECE	taru Ndouwe cur fu Ta	ert com Abrane	usher tu
Data Files								
Codellist		Download		1	Local: CSV	ISON]		
Country codes		Download	1	1	Local: CSV	ISON J		
Function classifiers		Download		1	Local: CSV -	ISON 1		
Status Indicators		Download		1	Local: CSV	ISON1		
Subdivision codes		Download			Local: CSV -	ISON 1		

Task 2: Checking Coordinates

Note:

- In this step, we are going to check for common errors, not for coordinates precision.
- For your reference, Excel can change the function names based on your language.

This website can help search the function translations: <u>http://dolf.trieschnigg.nl/excel/index.php</u>

\leftrightarrow \rightarrow \circlearrowright dolf.trieschnigg.nl/excl	l/index.php					□ ☆ =	- Z & -
Microsoft Excel function	n translations	Enter function name				in Share 90	Abou
Displayed languages	▶ Info	Microsoft excel Excel	formulas Excel vlockup function				
All None	English/English	French/Irançais	Dutch/Nederlands	German/Deutsch	Spanish/español	Portuguese/português	Russian/pycc
English / English	ARS	ARS	ARS	ARS	ARS	ARS	ARS
French / français	100	100	Hoo	100	100	100	100
Dutch / Nederlands	ACCRINT	INTERET.ACC	SAMENG.RENTE	AUFGELZINS	INT.ACUM	JUROSACUM	накопдохо
German / Deutsch	ACCRINTM	INTERET ACC MAT	SAMENG RENTE V	AUEGELZINSE	INT ACUM V	UROSACI MV	накоплохо
Spanish / español	HOOKITTII	and encertained	Source of the second se	NOT OCCUPION	1113100m.1	Jon Control Internet	101001040700
Portuguese / portugués	ACOS	ACOS	BOOGCOS	ARCCOS	ACOS	ACOS	ACOS
Russian / русский	ACOSH	ACOSH	BOOGCOSH	ARCCOSHVP	ACOSH	ACOSH	ACOSH
Italian (Italiano	HOUGHT	Roodi	50000000	Alcobolin	NOCON	NOCON	NOCON
Danish / dansk	ACOT	ACOT	BOOGCOT	ARCCOT	ACOT	ACOT	ACOT
Polish / polski	ACOTH	ACOTH	BOOGCOTH	ARCCOTHYP	ACOTH	ACOTH	ACOTH
Finnish / suomi							
Norwegian Bokmål / norsk bokmål	ADDRESS	ADRESSE	ADRES	ADRESSE	DIRECCION	ENDEREÇO	АДРЕС
Czech / čeština	AGGREGATE	AGREGAT	AGGREGAAT	AGGREGAT	AGREGAR	AGREGAR	ΑΓΡΕΓΑΤ
🕑 Hungarian / magyar							
Turidsh / Turkçe	AMORDEGRC	AMORDEGRC	AMORDEGRC	AMORDEGRK	AMORTIZ PROGRE	AMORDEGRC	АМОРУМ
Advertentie gesloten	AMORLING	AMORLINC	AMORLING	AMORLINEARK	AMORTIZ LIN	AMORLING	АМОРУВ
door Google	AND	ET	EN	UND	Y	E	И
Weergevon	ARABIC	CHIFFRE ARABE	ARABISCH	ARABISCH	NUMERO ARABE	ÁRABE	АРАБСКОЕ
Waarom deze advertentie? (>	AREAS	ZONES	BEREIKEN	BEREICHE	AREAS	ÁREAS	ОБЛАСТИ
	ASC	ASC	ASC	ASC	ASC	ASC	ASC
	ASIN	ASIN	BOOGSIN	ARCSIN	ASENO	ASEN	ASIN
	ASINH	ASINH	BOOGSINH	ARCSINHYP	ASENOH	ASENH	ASINH

Task 2: Checking Coordinates

Step 1. Conversion from sexagesimal format into decimal one

Suppose you have the first row at position 2, the Coordinates column at position K.

	1	А	В	С	D	E	F	G	Н	1	J	K	L	М	N	0	Р	Q	R	S
1			AD		.ANDORR	Δ														
2			AD	ALV	Andorra la	Andorra l	a Vella	34-6	AI	601		4230N 001	31E							
3			AD	CAN	Canillo	Canillo		3	RL	307		4234N 001	.35E							

1. Position yourself in cell M2 from here we are going to split the coordinates in their components:

example: 4230N 00131E

Latitude_degrees (M2): =MID(\$K2,1,2)	42
Latitude_minutes (N2): =MID(\$K2,3,2)	30
Latitude_direction (O2): =MID(\$K2,5,1)	Ν
Longitude_degrees (P2): =MID(\$K2,7,3)	001
Longitude_minutes (Q2): =MID(\$K2,10,2)	31
Longitude_direction (R2): =MID(\$K2,12,1)	Е
	Latitude_degrees (M2): =MID(\$K2,1,2) Latitude_minutes (N2): =MID(\$K2,3,2) Latitude_direction (O2): =MID(\$K2,5,1) Longitude_degrees (P2): =MID(\$K2,7,3) Longitude_minutes (Q2): =MID(\$K2,10,2) Longitude_direction (R2): =MID(\$K2,12,1)

- 2. In case you'd want to have a column with the decimal conversion of the coordinates (GPS coordinates, used in web mapping applications) you can add
- a column for latitude [S2] with this formula =\$M2+(\$N2/60)*(IF(\$O2="S",-1,1)) and
- another for longitude [T2] with this formula =\$P2+(\$Q2/60)*(IF(\$R2="W",-1,1))

	А	В	С	D	E	F	G	Н	1	J	К	L	М	N	0	Р	Q	R	S	Т
1		AD		.ANDORRA	4															
2		AD	ALV	Andorra la	Andorra la	Vella	34-6	AI	601		4230N 001	31E	42	30	N	001	31	E	42.5	1.516667
3		AD	CAN	Canillo	Canillo		3	RL	307		4234N 001	35E								

Task 2: Checking Coordinates

Step 2: Checking

1. Minutes errors detection:

For this step you should check the two minutes column (N2 and Q2): If greater than 59, they need to be fixed (because in sexagesimal notation numbers go from 0 to 60, and 60 equals 0)

- 2. Bounding box test:
- Create a rectangle enclosing your area of interest : bounding box.

You can use this tool <u>http://boundingbox.klokantech.com/</u> and select "CSV" format in the left-bottom corner, which gives you a string containing "western latitude, southern longitude, eastern latitude, northern longitude". i.e. the bounding box for Switzerland is with the parameters: 5.96,45.82,10.49,47.81.

Search via the Filtering function of the spreadsheet for values outside this rectangle.

Example: Ireland has a rough bounding box -10.3,51.7,-5.7,55.5;

on the spreadsheet if we filter for latitude less than 51 or more

than 55 we find one match, i.e. IE GOW, which should have 52

instead of 57 as latitude degree.

Note: It could be possible to find also direction problems (West instead of Ea



Task 3: Location name checking

Order the spreadsheet by Name or NameWoDiacritics.

For each name, check

- a. for spelling correctness (abbreviations, apostrophes, diacritics)
- b. Check based on section 3.3 of the UN/LOCODE manual (<u>http://www.unece.org/fileadmin/DAM/cefact/locode/unlocode_manual.pdf</u>)
- c. check if there could be possible duplications comparing similar names with the assistance of Subdivision and Coordinates metadata

Α	B	C	D	E	F	G	н	- I	J	K	L
:hange	Country	Location	Name	NameWoDiacritics	Subdivision	Function	Status	Date	IATA	Coordinat	t Remarks
	AD		.ANDORRA								
	AD	ALV	Andorra la Vella	Andorra la Vella		34-6	AI	601		4230N 001	131E
	AD	CAN	Canillo	Canillo		3	RL	307		4234N 001	135E
	AD	ENC	Encamp	Encamp		3	RL	307		4232N 001	134E
	AD	ESC	Escaldes-Engordany	Escaldes-Engordany		3	RL	307		4231N 001	133E
	AD	EAC	Escàs	Escas	4	3	RL	1407		4233N 001	131E
	AD	FMO	La Farga de Moles	La Farga de Moles		#NAME?	RQ	307			
	AD	LMA	La Massana	La Massana		3	RL	307		4234N 001	129E
	AD	ORD	Ordino	Ordino		3	RL	307		4233N 001	131E
	AD	PAS	Pas de la Casa	Pas de la Casa		#NAME?	RL	307		4233N 001	144E
	AD	SJL	Sant Julià de Lòria	Sant Julia de Loria		3	RL	1101		4228N 001	130E
	AD	SCO	Santa Coloma	Santa Coloma		3	RL	307		4230N 001	130E
	AE		.UNITED ARAB EMIRA	ATES							
	AE	ABU	Abu al Bukhoosh	Abu al Bukhoosh		1	RL	307		2529N 053	308E
	AE	AUH	Abu Dhabi	Abu Dhabi	AZ	1-345	AI	1101		2428N 054	422E
	AE	AMU	Abu Musa	Abu Musa		1	RL	201		2552N 055	501E
	AE	ARP	Ahmed Bin Rashid P	Ahmed Bin Rashid Po <mark>t</mark>	AZ	1	RL	1407		2532N 055	533E
	AE	MLA	Ajman	Ajman		1-3	RL	103	QAJ		
	AE	AAN	Al Ain	Al Ain		4	AI	9601			
	AE	DHF	Al Dhafra	Al Dhafra		34	AI	507		2414N 054	432E

Status (Refresh older entries)

The objective of the Metadata improvements is to ensure identification of the locations, each one of them should have the correct metadata.

Priority in (re)checking entries should be given to:

- entries where Date column is empty or where Status is not yet Approved (start focusing on QQ and UR)
- entries where Date column is older than 20 years

Subdivision

Subdivision is based on ISO-3166-2 standard (<u>https://www.iso.org/obp/ui/#search/code</u>).

ISO A	Online Bro	wsing Platform (OBP) Search results ×				₩ Sign in ► Language ► Help Search
249 results	for			Sort by: Eng	jish short name	a-3 code Numeric code Results per page: 25 🔻
Country codes	Unicially assigned	L OUES				
Standards				1 2 3 4 5 6 7 8 9 1	0 Next	
Collections Publications		For definition of the different code types, please click $\underline{\text{here}}$				Follow
Graphical symbols		English short name	▲ French short name	Alpha-2 code	Alpha-3 code	Numeric
Terms & Definitions		Afghanistan	Afghanistan (l')	AF	AFG	004
Country codes	249	Åland Islands	Åland(les Īles)	AX	ALA	248
		Albania	Albanie (l')	AL	ALB	008
Code type	•	Algeria	Algérie (l')	DZ	DZA	012
Officially assigne	249	American Samoa	Samoa américaines (les)	AS	ASM	016
other codes		Andorra	Andorre (l')	AD	AND	020
		Angola	Angola (l')	AO	AGO	024
		Anguilla	Anguilla	AI	AIA	660

Function

Function	Description	Remark
0	Specifies that the functional use of a location is not known	
0	and is to be specified.	
1	Specifies that the location is a Port, as defined in UNECE	
L	Recommendation 16.	
2	Specifies that the location is a Rail terminal.	
3	Specifies that the location is a Road terminal.	
4	Specifies that the location is an Airport.	Authorized by IATA
5	Specifies that the location is a Postal exchange office.	Authorized by UPU
6	Value reserved for multimodal functions, ICDs etc.	
7	Value reserved for fixed transport functions (e.g. oil	
/	platform).	
В	Specifies that the location is Border crossing.	

- Function '0' indicates "function not known". Entries with function '0' should be updated with the correct function(s).
- To assist in the classification, each function should map to a general criterion, which could be established at country level or could be discussed to provide uniformity.

Example questions to which you would answer:

- What is a postal exchange office in your country? A "mail sorting center"?
- Is there any official registry for this function to use as reference?
- Should a location have the function "airport" only when a location has an IATA-classified airport?

Coordinates

When feasible, you should fill in the coordinates in the correct format: DDMMx DDDMMx

If you're given a coordinate, you should ask if it is in the correct projection (data should be in Web Mercator projection, EPSG:900913), especially when asking to people working with geographic systems, because in different parts of the world are using different coordinates (more precise) projection.

To convert from a decimal coordinate, you could use this simple algorithm (example: 41.894100, 12.502441):

- The whole part becomes the degree (41)
- Remove the whole part and multiply per 60 (41.894100 - 41 = 0.8941; 0.8941*60 = 53.646)
- The whole part becomes the minute (53)
- The desired coordinate is 4153, since the starting number is positive, we would use direction N, so the final number is 4153N
- Repeating the steps for the second coordinate, gives 1230E.
- The conversion result is '4153N 01230E'

When checking locations or processing DMRs, you could ask for confirmation to the responsible agency in your country.

- AA Approved by competent national government agency
- AC Approved by Customs Authority
- **AF** Approved by national facilitation body
- AI Code adopted by international organisation (IATA or ECLAC)
- AS Approved by national standardisation body
- **RL** Recognised location
- **RN** Request from credible national sources for locations in their own country
- RQ Request under consideration
- **RR** Request rejected
- QQ Original entry not verified since date indicated
- **XX** Entry that will be removed from the next issue of UN/LOCODE

Entries with function(s) authorized by governmental bodies should be with status starting with 'A'.

Free data sources:

http://www.geonames.org/v3/ http://www.openstreetmap.org/

https://en.wikipedia.org/wiki/Main Page





				Shotlogged in Talk Contributions Greate account	
Welcome to Wikipedia, the two encyclopelin the argume care off. 5.30.11 underse a right	• Ads • Biograp • Geogra	ly the state of th	History Mathematics Science	Society Society Technology All potats	
From today's featured article	In the news				
The Elder Samb // Oblinet is a lack on the space state part we despect by litheted care Subce and packated by Beresda on and the Two Final Instantion descent 2.4 Caress. Oblion was researed on Nach 24, 2006, for Microsoft Windows and Subce Xia, and lack mode process of Byolikawa in the main structure. By days cardiant the Instantial Multice Samb 2016 and the seared racial Collina. The space care to make anywhere in the grave work of the Instantial Multice Samb 2016 and the main structure benchmarks. The partice care structure of the anywhere the space work of any hyper lack structure structure careful gramentals to be spaces the sam. The Eddard AL space manages the complex descents of the sampler datacets. (PGO) part when the new careful and complex data moves of lock on the same structure benchmarks and the sampler datacets. (PGO) part heating the careful and complex data moves of lock on the same structure benchmark and the sampler datacets. (PGO) part heating the the same the Eddard AL space benchmarks and the sampler datacets. (PGO) part heating the fact and endard monogeneous structures. (Include the same structure benchmark and the same structure benchmar	Arretical guitarist, si So, So, So, A colossal ancient T depicing phranch Ps The P The Vious for a Turk prenal election, bat Classities over a Turk Turkey and the Neth South Karrats Consili Onogene Table of Mose	Anextain guards, target, and rick and of porters Check Barry Sockred des at the ape of A colorated ancient Egyption tables, discovered in Callo, a preiminarity dentified as dentify maximum Damark I. The VVC. Leik y consort Prime Matter Lake fulls, viso Primor stables in the Dathe penetrie detection, bit this growing collisition tables in traphy. Biguines on a Tarkits conditioned Table (Sale Ley dependie Incident tables) Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and the Methematin Hogger & deplendie Incident tables as Things and tables are as a stables and tables as The Methematines are as a stable tables are as a stable tables as Things are as a stable tables as a stable tables are as a stable tables as			
Recently featured. Ser/ Rambal - Resetta Stone - George Bernard Archive - Bri email - More featured art	Shaw Recent deaths: Derek W	alcott - Hasyim Muzi	adi - André Tosel - I	Henri Cueco	
Did you know	On this day				
Institute Rest can be Flight tals Egypti a popular sciencific für Ortstan at, but the cartest known example provined only attrib to about 1377 Institute Cartest Marka Cartest Statist a scoppe of the sension in LACI22 mode humans subceptive to devices such as torowing with Internations Institute This Central Human Marka Cartest Statist and Marka Statistics Institute This Central Human Marka Cartest Statist and Marka Statistics Institute This Central Human Marka Cartest Statist and Marka Statistics	March 20 March equin (2017) Independence Di 205 – Maximinus Th emperor who gained 1 600 – Five advisors effectively ending his 1 810 – After escaping	x (10:28 UTC, 2017) y in Tunisia (1956) ax succeeded to the ower by virtue of his o Polish-Swedish id reign in Sweden from exile in Elba, h), Nownuz in Iran, C throne of the Rom command of the a ng Sigismund III Va Japoleon Bonaparte	Iental Asia, he Bahál Faih, and Zoroastranism an Empire, a so-caled barracks mmy: Isa wete publicly executed, o (potured) enterol Paris,	

WIKIP

OpenStreetMap and Whosonfirst

• OpenStreetMap (<u>http://www.openstreetmap.org/</u>)

OpenStreetMap is a global geographic database built by the people on the Wikipedia model (crowdsourcing): everyone can log in and edit the data by themselves, and the data is available to download and process. Many services are built using it as source of data for map displays.

The main website provides services to the users, but can be used also as a browsing interface to look for places and check their descriptions (OSM uses a key=value data model for attributes).

Whosonfirst (<u>https://whosonfirst.mapzen.com/spelunker/</u>)

Whosonfirst is a global gazetteer , essentially a big list of places with their geometry. It can be used to look up places and get information about their administrative structure (getting their subdivision for example). A geocoding service using this project is Mapzen search (<u>https://mapzen.com/projects/search/</u>).



Wikipedia / Wikidata

• Wikipedia

Wikipedia, the free encyclopedia, is a reliable source of information and if is often one of the first results when searching on Google.

On Wikipedia there are different experiments of using UN/LOCODE:

- a proposal of displaying into the Infobox

(https://en.wikipedia.org/wiki/Template_talk:Infobox_settlement/Archive_7#UN.2FLOCODE)

- a redirect template (<u>https://en.wikipedia.org/wiki/Template:R_from_UN/LOCODE</u>) that redirects a UN/LOCODE to the city (or location) page.

Wikipedia could be used to get information about infrastructure, location (articles could have coordinates in the top right corner of the page), and administrative subdivisions.

• Wikidata

The sister project Wikidata is a database of information editable like Wikipedia: each object has an identifier (Q plus a number) and has different 'predicates' (property –P plus a number- and value). The property "UN/LOCODE" exists as P1937.

Example: Genoa, Italy (<u>https://www.wikidata.org/wiki/Q1449</u>) has UN/LOCODE (<u>https://www.wikidata.org/wiki/Property:P1937</u>) ITGOA.

Each predicate can have 'qualifiers': one could add a start date (P580) or an end date (P582) to the predicate when it is added or deprecated.

The advantage of Wikidata is that in the future it could be used as a data source for Wikipedia, and also can be interrogated as a database to gain information with the Wikidata query service (<u>https://query.wikidata.org/</u>) using a language called SPARQL.

Wikipedia / Wikidata

As an example, this query lists all the objects in Wikidata having a UN/LOCODE property, returning a list with the name and the coordinates.

Applying this to a single country is simple, we change the country predicate to return all the locations in San Marino (<u>https://www.wikidata.org/wiki/Q238</u>).