

BIP0039 German Wordlist
Statement Of Requirements (SOR) related to Proposal 2019-01-15

15th of January, 2019
SebastianFloKa

- Scope of BIP0039 German wordlist / target group:

--> General creation of seeds for cryptocurrencies in Soft-, Hard-, Paper-Wallets etc.:

- Typical requirements for BIP0039 Wordlists to be fulfilled

--> "Engraving" (physical writing etc.):

- Short words to have a low average length of words

- Low number of possible characters (only capital letter to be used, not mixed with lower case letter)

- etc.

--> Verbal transmission of words: unambiguity of words:

- No homophones

- Sufficient levenshtein distance

- Only capital letters

- Description sentence with article should be easily possible: "HASE wie der Osterhase"

- etc.

--> Memorize the seed:

- Easy to remember words (avoid präposition etc.) or possibility to create cribs

- etc.

- Specific regarding German language:

- German (with its many dialects) is official language in 6 countries and partly spoken in several more: A word listed/mentioned in the German DUDEN should therefore be a good basis for the BIP0039 German Wordlist in order to meet the expectation of most German speakers.

- Nouns have as first letter always a capital letter, even in the middle of a sentence. Others have a lower case letter at the beginning of the word except at the beginning of a sentence. Some words use the same letters for verb and noun (Example: Verb: husten / Noun: Husten) or adjective and noun (fett / Fett). Therefore all letters shall be capital letters in BIP0039 German Wordlist in order to simplify and particularly to avoid misunderstanding.

Details: See next pages

- General requirements:	Example (German)	BIP0039 German WL Proposal 2019-01-15	ENGLISH	SPANISH	ITALIAN	FRENCH
Source of words		Listed in German Duden 2018				
Orthography		Along German Duden 2018				
Level of recognition		Subjectively classified as "common"				
- Regional words (austriacism, helvetism, borealism, etc.)	SEMMELE	Allowed if known in most german spoken areas				
- Foreign words (anglicism, latinism, gallizism, graecism, etc.)	FLIRT, SHIRT	Allowed if common	YES	YES	YES	YES
Matching words with BIP0039 Spanish Worlist (see Y75QMO)	GRATIS, LEER	Not allowed	NO	-	NO	NO
Matching words with other BIP0039 Wordlists except Spanish one		Allowed	YES	NO	?	YES

- Requirements on characters	Example	GER - Proposal 2019-01-15	ENG	ESP	IT	FR
Normalization to Unicode NFKD		to be done	-	-	-	-
Letters: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z / Space character (position 4, 5, 6, etc.)	KATZE	Allowed	-	-	-	-
Letters Ä, Ö, Ü, and their corresponding replacements "AE", "OE", "UE"	BÄR, BAER	Not allowed	-	-	-	-
ß (German SZ) and the corresponding replacement "SS"	FLEIß, FLEISS	Not allowed	-	-	-	-
Same Letter in a row	KUSS, AAL	Allowed	YES	YES	NO Vocal	YES
Diactrical signs	À	Not allowed	NO	YES	YES	YES
Capital letter	LACHEN, HUSTEN, BRATEN, GRABEN, FEST, FEIGE, DEFEKT	Allowed	NO	NO	NO	NO
Lower case letter	lachen, Lachen, Reich, reich husten, Husten, Fest, fest, feige, Feige	Not allowed	YES	YES	YES	YES
Min. amount of letters per word	HOF	3	3	3	4	5
Max. amount of letters per word	BANANE	6	8	8	9	8

- Requirements on social etiquette in choice of words and quality in general

GER - Proposal 2019-01-15

ENG

ESP

IT

FR

			ENG	ESP	IT	FR
Demonyms	INDER, ESKIMO, MOSLEM	Not allowed	?	?	?	NO
Embarassing or negative words, insults, major illness, etc.	ARSCH, SARG, TOD	Not allowed	NO	NO	NO	NO
Religious words	GOTT --> OK, MITRA --> NOK	Allowed if common, to be avoided if too specific	?	?	?	?
Term, technical term	HANGAR, HELIUM, KEILER	Allowed if common	?	?	?	?
Companies, Brands	PUMA --> OK, ADIDAS --> NOK	Not Allowed, except exists as standard word.	?	?	?	?
Words with high risk of not being used or known in the future	FAX	To be avoided	?	?	?	?

- Grammatical requirements:

		Example	GER - Proposal 2019-01-15	ENG	ESP	IT	FR
Nouns (Standard)	abstracta	LIEBE	Allowed	YES	YES	YES	YES
	concreta	TISCH	Allowed	YES	YES	YES	YES
	generic designation	KATZE, EICHE, BRUDER	Allowed	YES	YES	YES	YES
	name of substance	STAHL, SALZ	Allowed	YES	YES	YES	YES
	collectiva	DORF, WILD, SIPPE	Allowed	YES	YES	YES	YES
Nouns (Abbreviations)	Abbreviations (chain of letters)	PKW, GMBH	Allowed	?	?	?	?
	short words (syllable word short, acro-nym, initialisms, head words, tail words)	AZUBI, ABO, BUS	Allowed	YES	?	YES	?
proper nouns	Well known in cultural or intercultural context	ADAM, EVA, ZEUS; APOLLO, CAESAR, MERKUR	Allowed if common	?	?	?	?
	Person used as unit or award	CELSIUS, NOBEL	Allowed if common	?	?	?	?
	Countries, cities, rivers	JAPAN, WIEN, DONAU	Allowed if common	NO	?	?	?
	Current celebrities or from the near past	MERKEL, MARX, VETTEL	Not allowed	NO	?	?	NO
Nouns (special)	verb- or adjective derived noun	GEHEUL, LEBEN	Allowed	?	?	?	?
	compound noun	TIERHEIM	Allowed	?	?	?	?
Infinitive verbs		SPRINGEN	Not allowed	YES	?	YES	YES
Reflexive verbs		SICH FREUEN	Not allowed	-	?	?	NO
Adjectives		KALT	Allowed	YES	?	?	YES
Articles		DER, DIE, DAS	Not allowed	?	?	?	NO
Numeral adjectives		VIEL, WENIG	Not allowed	?	?	?	NO
Personal pronoun		ICH, SIE, WIR	Not allowed	YES	?	?	NO
Pronouns except personal pronoun		MEIN, MICH, DIESER, WAS	Not allowed	YES	?	?	NO
Numericals		EINS, ZWEI, DREI	Allowed	YES	YES	?	NO
Conjunctions		UND, ABER	Not allowed	?	?	?	NO
Prepositions		UNTER, NACH	Not allowed	YES	?	?	NO
Adverbs		QUASI, CIRCA	Not allowed	YES	?	YES	NO
Interjections		AHOI, HELAU, HALLO,	Not allowed	?	?	YES	NO
Subdefinition regarding nouns	Singular	HASE	Allowed	YES	YES	YES	YES
	Plural	HASEN	Not allowed (exc. pluraliatantum)	?	?	NO	NO
	Singulariatantum	STOLZ, OBST,	Allowed	YES	YES	YES	YES
	Pluraliatantum	ELTERN, ALPEN, FERIEN	Allowed	?	?	?	?

- Criteria on Levenshtein Distance (LS)

Same words within wordlist with Levenshtein Distance (LS) of:

	Example	GER - Proposal 2019-01-15	ENG	ESP	IT	FR
0	HOF, HOF	Not allowed (of course)	NO	NO	NO	NO
1 - Substitution: of similar sounding letters "B&P", "D&T", "G&K", "S&Z", "M&N"	RAD & RAT, BACH & DACH	Not allowed	YES	YES	NO	NO
1 - Substitution: of non similar sounding letters (see ligne above)	GRAZ & GRAF & GRAD --> OK GRAZ & GRAF & GRAD & GRAL --> NOK	Allowed but max. 3 words to be affected per position (= max. 2 collisions per position in word)	YES	YES	NO	NO
1 - Addition (= subtraction): of letters "PF&F", "H"	PFLUG & FLUG / RUHM & RUM	Not allowed	?	?	?	NO
1- Addition: except letters "PF&F", "H"	TIER & STIER / HUT & HAUT	Allowed	YES	YES	YES	NO
1- Permutation	OSLO & SOLO	Not allowed	YES	?	?	NO
2 - Addition / subtraction / permutation	HUND & HERD / HERD & PFERD	Allowed	YES	YES	YES	?
3 and more	HUND & HASE	Allowed	YES	YES	YES	YES
First 4 letters sufficient to indicate complete word		Must have	YES	YES	YES	YES
Word with 3 Letters requires Levenshtein addition of min. 2 letters		Required	NO	?	YES (3)	YES
Maximum # of collision with LS substitution on one position		2	4	4	0	0

- For Information: amount of affected words with LS of 1 - comparison between Proposal 2019-01-15 and other BIP0039 Wordlists:

LS substitution 1st letter	GUT & NUT	612	632	406	0	0
LS substitution 2nd letter	HASE & HOSE	202	148	90	0	0
LS substitution 3rd letter	FELGE & FEIGE	216	270	328	0	0
LS substitution 4th letter	FELL & FELS	216	198	116	0	0
LS substitution 5th - 8th letter		0	not allowed --> 0			
Maximum # of LS substitutions on one position		2	4*	4	0	0
LS addition 1st letter	TIER & STIER	52	109	76	9	0
LS addition 2nd letter	KASSE & KLASSE	19	36	not yet checked		
LS addition 3rd letter	BIER & BIBER	23	23			
LS addition 3rd letter	LASER & LASTER	13	30			
LS permutation 1st, 2nd, 3rd, 4th & 5th letter	FROSCH & FORSCH	0	6			

* Examples from BIP0039 English Wordlist (4 collisions = 5 words affected): wide&wife&wine&wire&wise / cave&have&pave&save&wave / eight&light&night&right&sight

- Requirements concerning the similarity in pronunciation of words:

	GER - Proposal 2019-01-15	ENG	ESP	IT	FR	
Homographs	VOLLZUG	Allowed	YES	?	?	?
Homophones (both Singular)	GRAF & GRAPH	Not Allowed	?	?	?	NO
Homophones (one Singular, one Plural)	FERSE & VERSE (FERSE --> OK)	Allowed to use the singular one	?	?	?	?
Homonyms	BANK (Geldinstitut, Sitzgarnitur)	Allowed	YES	YES	YES	YES