



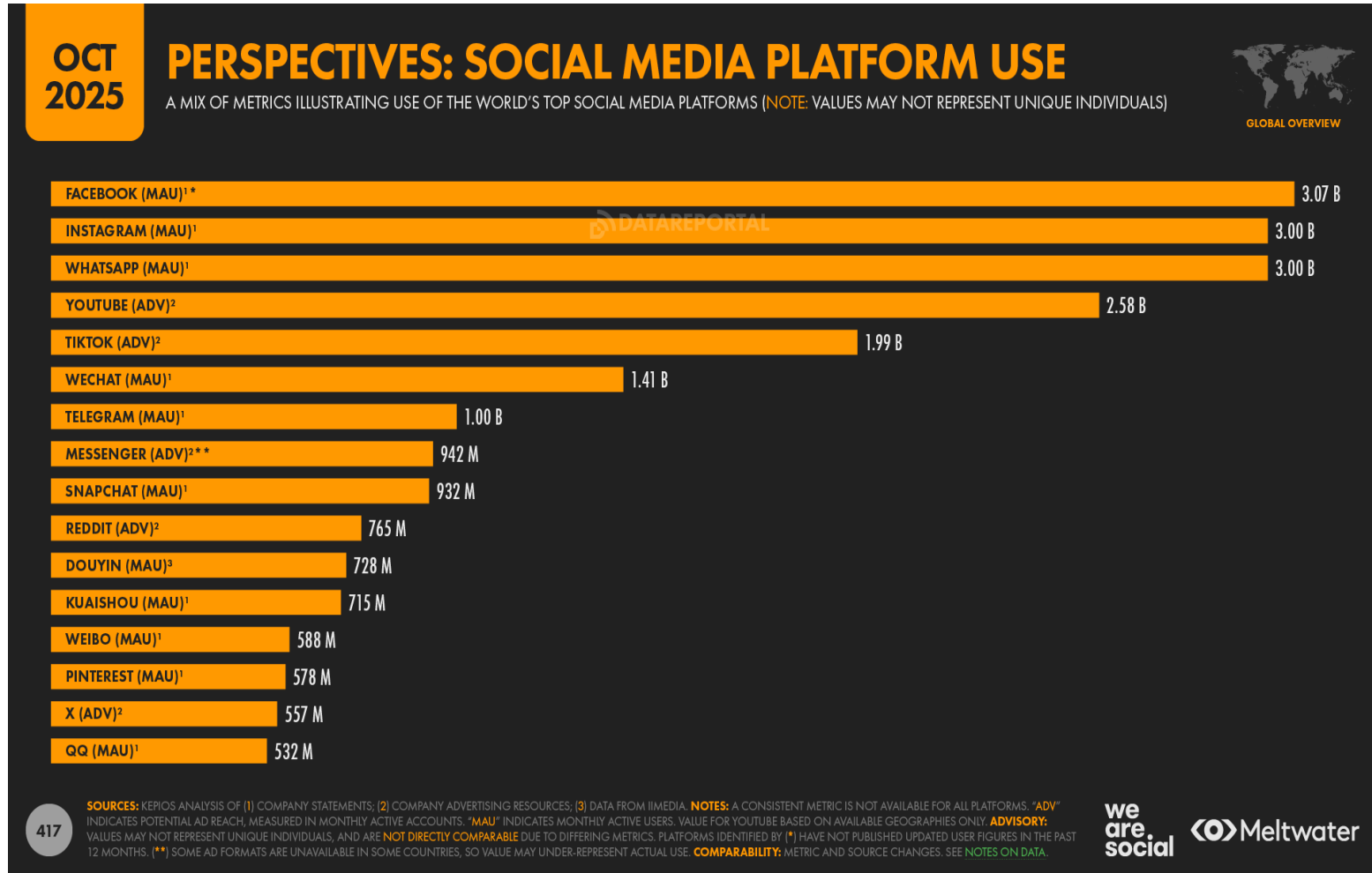
Uncovering the Hidden Market of Telegram Usernames: The Case of the Fragment Marketplace

Francesco Maggio, Andrea Michienzi, Barbara Guidi

University of Pisa - Department of Computer Science

21° International Conference on the Economics of Grids, Clouds, Systems, and Services
(GECON 2026)

Social Media in This Age



- Massive user growth;
- Quick diffusion of content;
- Limited control;
- Legal risks (copyright, privacy).

Username

- Defined as a unique string of characters used to identify a person on a digital system.
- In social platforms, usernames are central elements of digital identity;
- Digital identities problematics:
 - Impersonation;
 - Identity theft;
 - Brand infringement;
 - Cybersquatting.

Example of
Typosquatting:

Instagram



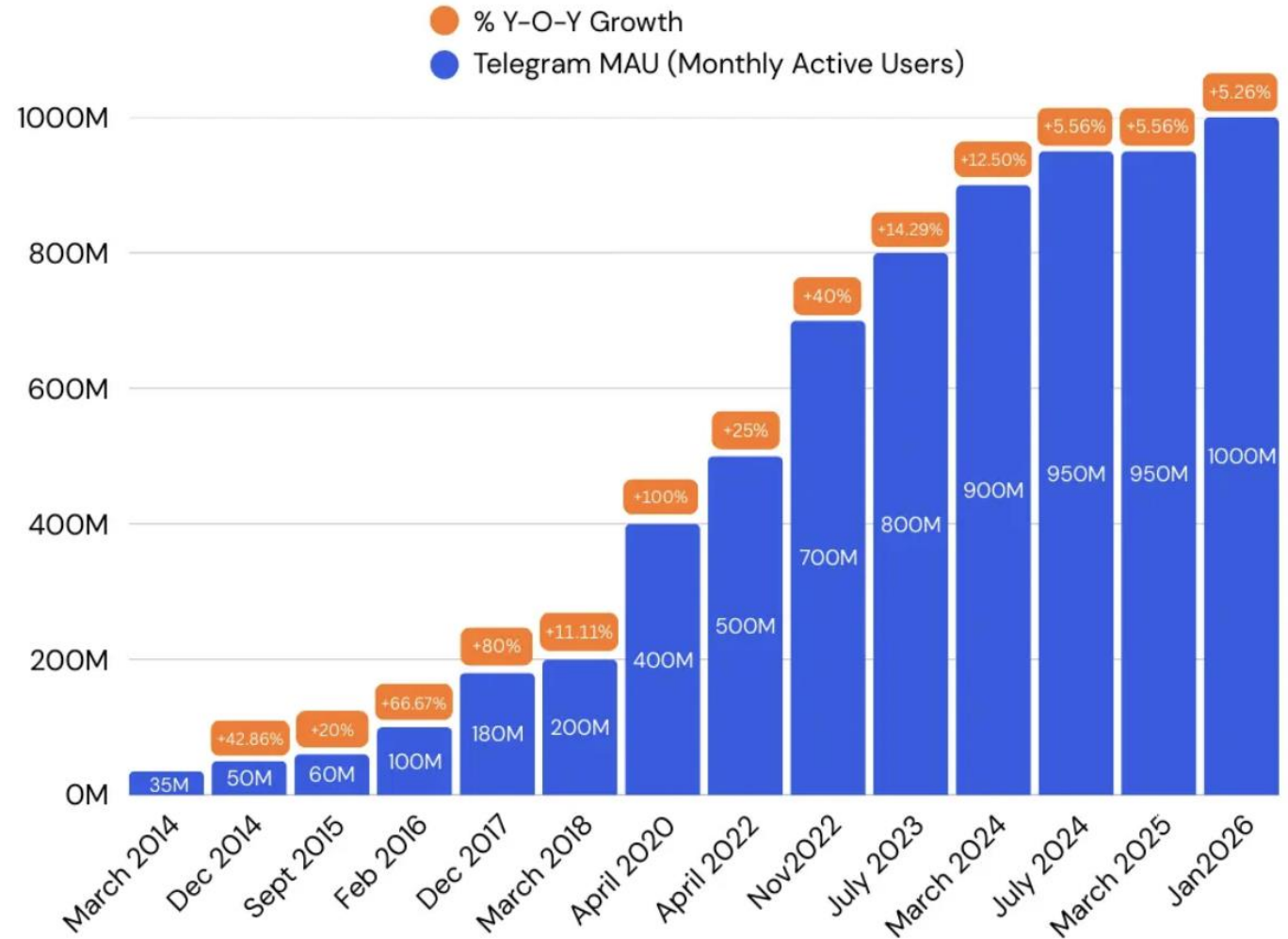
Instagram



Lowercase «L» instead of uppercase «I»

Telegram

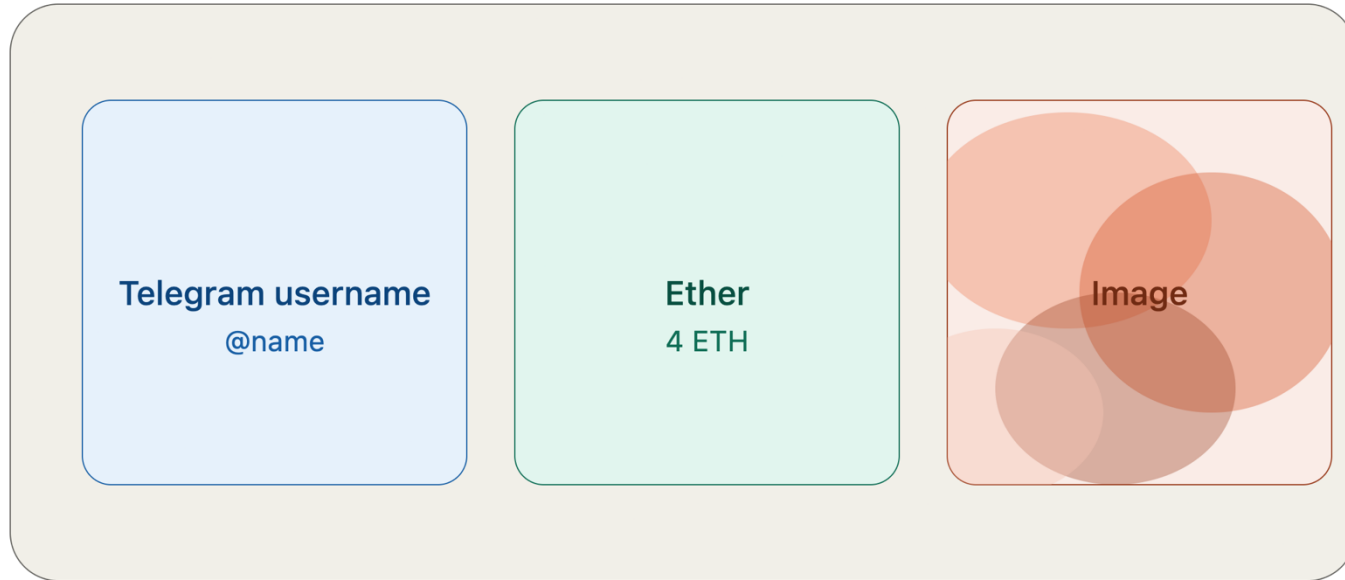
- 1 billion users reached in 2026;
- Textual and vocal chats and channels;
- Malicious activities, including revenge porn, terrorism, and racist movements.



Username as NFTs

- In 2022, Telegram introduced the possibility of converting usernames as **Non-Fungible Tokens (NFTs)**;
- **NFTs** are digital certificates registered on a **blockchain** that refer to a digital or physical artifact;
- A **blockchain** is a digital shared database constructed by chronologically linked blocks, **immutable** and **auditable** at all times.

Username as NFTs

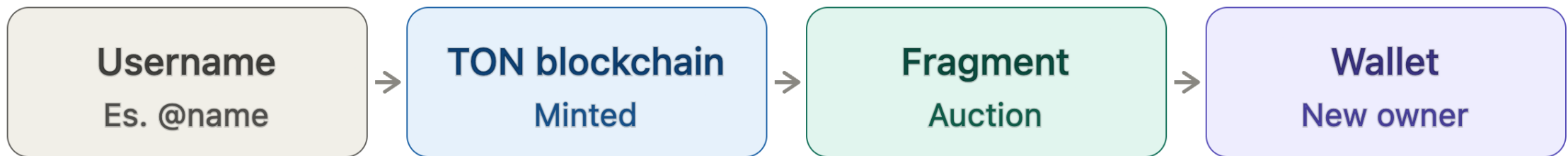


- Usernames are **minted** (recorded on the blockchain) and they can be **stored** into a user-owned digital **wallet**;
- Their histories are public; they can be bought/sold/traded.

The Fragment Phenomenon

Fragment: marketplace for collectible digital assets, based on the TON blockchain. Includes Telegram usernames.

Username have a **double use** as collectibles and identifiers. They are assigned through **auctions**.



Objectives

- Analyzing username-related digital identity issues (**impersonation, identity theft, brand infringement, cybersquatting**);
- Performing a cross-dimensional analysis on usernames, considering linguistic and structural perspectives.
- Building and sharing an open source dataset of Telegram usernames minted as NFTs;

Dataset

- Obtained through scraping between 23/05/2025 and 17/01/2026 and extended with TON API;
- **110,683** total assets including histories;
- Transfer dates from 02/11/2022 to 01/01/2026.

Asset	Scrape date	Trade history
@ghostpsycho	2026-01-16	[{'Owner': ..., ...}, ...]
@slowsunset	2025-11-16	[{'Owner': ..., ...}, ...]

Data Modeling and Graph

First step:

- Bipartite **asset-wallet** graph – 161,491 nodes (110,683 assets and 50,808 wallets), 281,654 edges;
- Giant component – 150,421 nodes (102,649 assets and 47,772 wallets), 275,330.

Data Modeling and Graph

Second step:

- Projection into **username-username** graph;
- Edges indicate shared wallets, weights indicate the number of common wallets;
- 3,609,071 edges with mean weight = 1.11.

Structural Communities

Community detection on the username-username graph through **Louvain** and **Leiden** algorithms.

	Number of communities	Modularity score
Louvain	113	0,74
Leiden	138	0,75

Normalized Mutual Information (NMI)	Adjusted Mutual Information (AMI)
0,846	0,844

Linguistic Clustering

- Combination of surface-level properties and BGE-M3 embeddings with 1024 dimensions;
- Clustering through **HDBSCAN** to group the usernames by linguistic characteristics;
- 20 clusters of homogeneous size total, leaving out 25,582 noise assets.

Linguistic Clustering

The 20 clusters were characterized based on linguistic patterns and content.

ID	Size	%	Name	Example
4	8,435	8.22	5L Phonetic	@anime
17	5,158	5.02	Repetition Patterns	@xxxx
5	4,974	4.85	Finance & Betting	@betting
1	4,842	4.72	Chinese Mix	@gongxu
19	4,357	4.24	Crypto Utility	@cryptowallet
13	2,797	2.72	Money & Lifestyle	@forextrader
15	2,413	2.35	Big Tech	@microsoft
18	2,179	2.12	Numerical Names	@a16z

Cross-dimensional Analysis

- Heatmaps considering all 20 linguistic clusters and the 15 largest structural communities;
- Both Louvain and Leiden algorithms included;
- Dimensions:
 - Distribution;
 - Average last sale price (in TON);
 - Free transfers volume;
 - Average wallet holding time.

Distribution analysis - Louvain



0	9%	7%	8%	9%	8%	5%	1%	7%	7%	5%	2%	3%	5%	4%	4%	5%	1%	3%	2%	4%
1	8%	6%	6%	6%	5%	10%	8%	7%	5%	5%	4%	3%	4%	3%	4%	4%	3%	3%	2%	2%
11	18%	11%	8%	7%	8%	7%	4%	4%	4%	4%	2%	5%	3%	2%	3%	3%	2%	1%	3%	2%
5	15%	7%	5%	4%	4%	5%	7%	5%	6%	3%	6%	9%	2%	3%	3%	3%	4%	3%	2%	3%
6	12%	9%	6%	6%	6%	6%	7%	5%	6%	4%	6%	4%	3%	3%	3%	3%	4%	3%	2%	4%
1	10%	8%	5%	6%	5%	7%	5%	5%	7%	4%	6%	4%	3%	4%	4%	3%	5%	4%	2%	4%
9	7%	3%	8%	7%	6%	5%	2%	7%	7%	6%	8%	2%	4%	6%	5%	5%	1%	4%	2%	4%
10	7%	5%	5%	4%	4%	8%	11%	5%	6%	3%	11%	4%	3%	4%	4%	3%	6%	3%	2%	3%
12	14%	9%	9%	7%	9%	6%	7%	4%	5%	3%	1%	7%	5%	3%	3%	3%	1%	1%	3%	2%
13	9%	5%	9%	10%	9%	8%	10%	6%	4%	5%	1%	3%	5%	2%	3%	4%	2%	1%	3%	1%
14	8%	7%	5%	5%	6%	13%	14%	3%	2%	3%	6%	3%	2%	1%	2%	2%	6%	1%	9%	1%
18	15%	10%	8%	8%	9%	4%	4%	5%	4%	6%	4%	4%	3%	3%	3%	2%	3%	2%	2%	2%
21	11%	7%	8%	9%	8%	7%	12%	5%	3%	5%	1%	4%	3%	2%	2%	3%	2%	1%	4%	1%
33	11%	8%	9%	8%	11%	4%	1%	7%	5%	4%	1%	3%	5%	5%	3%	3%	1%	5%	3%	2%
37	8%	6%	6%	6%	6%	5%	3%	7%	9%	4%	5%	3%	4%	6%	5%	4%	2%	6%	2%	5%
5L Phonetic																				
4/5L Semantic																				
Tech Finance																				
Finance & Betting																				
Chinese Mix																				
Repetition Patterns																				
VIP & Gambling																				
Generic Services																				
Corporate & Enterprise																				
Brands & Entertainment																				
Crypto Utility																				
Crypto & Memes																				
Asian Business 8L																				
Money & Lifestyle																				
Big Tech																				
Sports & Generic																				
Numerical Names																				
Brands & Foundations																				
Generic Phonetic																				
Media & Streaming																				

Distribution analysis - Leiden

0	10%	7%	5%	9%	1%	9%	8%	7%	7%	2%	5%	3%	1%	5%	4%	4%	5%	4%	2%	3%
1	18%	8%	6%	5%	7%	5%	4%	5%	5%	5%	4%	5%	4%	3%	3%	2%	2%	3%	3%	2%
2	8%	7%	9%	4%	6%	4%	3%	7%	7%	7%	4%	2%	7%	4%	6%	4%	4%	4%	1%	2%
3	24%	18%	2%	8%	1%	6%	9%	4%	3%	2%	4%	5%	1%	2%	2%	1%	2%	1%	2%	2%
4	7%	7%	6%	6%	11%	4%	5%	4%	5%	11%	3%	9%	5%	2%	2%	3%	2%	3%	2%	3%
6	8%	7%	14%	5%	14%	5%	6%	3%	2%	6%	3%	3%	6%	2%	2%	1%	2%	1%	10%	1%
8	9%	6%	5%	6%	4%	7%	6%	7%	8%	5%	5%	3%	2%	4%	4%	5%	4%	5%	1%	5%
10	9%	6%	7%	5%	5%	6%	5%	5%	7%	6%	4%	3%	5%	3%	4%	5%	4%	4%	2%	4%
12	17%	11%	7%	8%	4%	7%	8%	4%	4%	2%	4%	5%	2%	3%	3%	2%	3%	2%	3%	1%
16	7%	3%	4%	8%	2%	7%	6%	6%	7%	9%	6%	2%	1%	4%	6%	5%	5%	4%	2%	4%
18	8%	6%	12%	5%	10%	6%	4%	6%	5%	4%	4%	4%	5%	3%	3%	3%	3%	2%	2%	4%
21	11%	7%	7%	8%	12%	9%	8%	5%	3%	1%	5%	4%	2%	3%	2%	2%	3%	1%	4%	1%
22	13%	8%	5%	7%	6%	6%	6%	5%	6%	5%	4%	4%	3%	3%	4%	4%	3%	4%	1%	3%
55	9%	5%	8%	9%	10%	10%	9%	6%	4%	1%	5%	3%	2%	5%	3%	2%	4%	1%	3%	1%
57	15%	9%	5%	8%	4%	8%	9%	5%	4%	4%	6%	4%	3%	3%	3%	3%	2%	2%	2%	3%
5L Phonetic																				
4/5L Semantic																				
Repetition Patterns																				
Tech Finance																				
VIP & Gambling																				
Finance & Betting																				
Chinese Mix																				
Generic Services																				
Corporate & Enterprise																				
Brands & Entertainment																				
Crypto Utility																				
Crypto & Memes																				
Numerical Names																				
Asian Business 8L																				
Big Tech																				
Sports & Generic																				
Media & Streaming																				
Generic Phonetic																				
Brands & Foundations																				

Average last sale price (in TON) - Louvain

0	771	1.3k	334	462	328	340	493	898	366	461	504	285	165	193	107	311	65	81	167	48
1	128	390	103	131	214	55	79	98	130	111	59	73	57	55	68	71	87	87	56	57
2	30	28	31	30	26	27	27	32	27	32	34	30	26	35	25	41	30	26	22	23
3	694	144	66	45	24	30	47	29	76	74	47	198	16	20	30	35	33	24	12	10
4	131	191	150	117	186	77	230	37	63	84	157	75	51	46	52	46	28	38	20	18
5	95	93	79	76	158	48	104	48	266	87	293	77	48	50	58	61	27	31	20	24
6	185	346	166	183	312	102	707	88	170	194	133	109	214	95	86	111	108	122	85	28
7	39	45	34	30	58	15	36	37	17	38	21	27	29	26	53	27	41	25	21	17
8	14	16	15	17	14	14	18	16	13	14	12	15	11	11	15	15	11	10	11	13
9	191	418	135	103	513	157	99	269	97	163	77	109	152	47	37	79	71	45	39	80
10	47	145	34	55	183	46	27	60	33	27	19	22	48	29	21	18	15	15	34	45
11	38	80	87	65	36	14	66	22	29	57	29	34	47	25	22	67	14	36	11	26
12	147	395	95	87	668	807	69	266	85	331	47	104	79	31	40	41	38	42	82	26
13	3.6k	1.6k	3.1k	2.9k	572	3.0k	1.2k	1.5k	1.2k	544	823	673	950	936	542	200	97	231	38	38
14	2.2k	3.0k	1.8k	1.7k	2.3k	343	880	521	1.3k	594	425	576	381	336	480	402	449	261	257	133
15	5L Phonetic	4/5L Semantic	Brands & Entertainment	Tech Finance	Crypto & Memes	Repetition Patterns	Finance & Betting	VIP & Gambling	Asian Business 8L	Chinese Mix	Sports & Generic	Generic Services	Generic Phonetic	Media & Streaming	Corporate & Enterprise	Big Tech	Brands & Foundations	Money & Lifestyle	Numerical Names	Crypto Utility

Average last sale price (in TON) - Leiden

0	1.6k	1.5k	414	485	363	941	728	540	1.4k	505	605	368	376	278	156	193	77	188	124	30
1	73	34	35	35	57	48	37	21	28	47	46	33	22	42	28	20	55	14	17	12
2	22	32	22	35	15	35	17	15	77	53	41	21	22	26	59	17	25	17	16	14
3	111	38	25	233	23	28	25	20	31	29	24	40	22	22	25	21	20	23	25	23
4	112	1.4k	34	51	86	38	83	27	26	65	41	233	29	27	28	21	20	17	28	21
5	190	72	203	37	36	104	46	70	62	33	26	29	49	21	27	40	21	37	20	47
6	2.9k	1.8k	2.1k	1.5k	2.2k	1.6k	1.6k	1.2k	462	740	462	561	374	583	492	331	469	302	267	140
7	207	124	261	158	289	98	99	55	43	165	287	141	76	98	79	194	30	26	36	31
8	28	31	25	29	26	30	33	28	32	32	32	33	25	41	26	35	17	21	26	22
9	371	186	322	743	144	183	159	108	85	200	136	101	239	110	90	94	110	90	122	23
10	474	105	236	72	152	88	138	54	102	151	51	66	70	98	64	58	105	50	77	57
11	457	146	623	69	88	104	95	772	261	332	47	98	81	40	39	31	38	83	42	24
12	182	147	229	257	71	121	171	95	48	95	171	81	59	48	48	53	30	26	41	19
13	423	192	519	96	97	104	141	158	270	163	77	111	155	79	37	47	71	39	45	80
14	83	38	30	57	30	67	89	14	20	56	27	32	46	69	22	20	14	11	35	14
15	4/5L Semantic	5L Phonetic	Crypto & Memes	Finance & Betting	Asian Business 8L	Tech Finance	Brands & Entertainment	Repetition Patterns	VIP & Gambling	Chinese Mix	Sports & Generic	Generic Services	Generic Phonetic	Big Tech	Corporate & Enterprise	Media & Streaming	Brands & Foundations	Numerical Names	Money & Lifestyle	Crypto Utility



Pepe the frog

Free transfers volume - Louvain

0	1479	1151	662	1395	1395	1303	88	120	987	673	315	743	527	693	76	480	336	594	350	205
2	125	115	173	96	119	111	197	80	122	47	43	63	64	65	57	64	42	53	31	63
4	1393	932	429	726	640	764	283	150	329	289	258	371	348	274	61	207	207	230	140	50
6	2150	1140	865	888	808	614	730	879	677	866	288	493	525	351	416	493	440	394	473	587
8	3054	1888	1241	1339	1244	1293	975	1139	985	1284	355	918	965	650	808	753	810	584	687	529
10	1098	817	754	693	706	579	558	727	622	841	238	568	479	293	584	441	568	424	389	397
12	180	56	120	342	159	169	60	202	247	210	32	189	48	165	36	178	130	118	131	57
14	354	320	369	179	194	251	497	490	214	269	65	169	169	110	287	145	136	110	147	92
16	114	71	67	69	52	73	34	6	33	27	19	25	56	37	5	29	25	29	10	8
18	582	362	419	603	688	642	393	48	359	178	194	313	179	314	94	148	99	191	81	37
20	622	681	1906	395	399	420	1164	1457	206	96	2848	273	430	142	767	130	68	145	57	30
22	273	142	45	176	97	128	24	56	57	59	15	59	62	37	32	57	45	33	21	25
24	579	551	606	569	461	451	806	43	272	162	312	334	263	217	108	66	106	142	63	67
26	244	174	70	182	128	278	20	15	141	93	68	69	57	103	9	44	112	65	29	11
28	111	124	54	64	68	78	8	10	56	115	12	69	50	41	13	56	45	42	53	81
30	5L Phonetic	4/5L Semantic	Repetition Patterns	Tech Finance	Finance & Betting	Chinese Mix	VIP & Gambling	Crypto Utility	Generic Services	Corporate & Enterprise	Generic Phonetic	Brands & Entertainment	Crypto & Memes	Asian Business 8L	Numerical Names	Big Tech	Money & Lifestyle	Sports & Generic	Media & Streaming	Brands & Foundations

Free transfers volume - Leiden

0	1483	1090	1435	550	1283	1281	70	115	953	635	277	756	435	661	460	60	579	332	338	187
2	1826	1001	717	722	716	527	677	686	579	629	275	426	472	302	304	371	296	291	357	251
4	578	444	250	384	289	212	235	299	367	420	60	301	119	236	307	299	224	228	222	167
6	702	635	245	112	230	362	44	25	110	48	108	86	204	91	36	8	59	13	17	15
8	518	351	295	347	208	304	440	550	226	334	53	152	290	88	209	234	106	180	243	317
10	600	630	391	1931	385	396	1128	1446	192	95	2863	264	415	133	113	715	131	64	36	34
12	114	117	60	54	73	73	6	19	56	118	12	75	48	42	58	14	40	45	57	85
14	990	678	652	732	679	543	586	749	582	801	245	487	398	336	461	549	373	581	393	393
16	1386	946	711	467	636	809	285	138	337	314	269	360	385	279	219	69	251	183	119	68
18	161	69	349	126	162	156	62	239	240	197	28	197	47	167	179	36	117	135	129	58
20	108	76	72	160	90	58	185	51	74	44	38	37	53	31	49	56	36	37	26	60
22	585	556	559	572	456	465	821	52	272	154	303	336	250	214	66	108	140	106	62	67
24	2844	1751	1260	1082	1192	1222	830	959	919	1218	313	835	890	587	680	699	546	766	603	496
26	582	359	600	419	688	642	389	48	356	178	194	328	179	314	148	94	191	99	81	37
28	262	138	168	53	109	123	28	60	48	63	18	71	82	37	55	32	33	55	21	25
30	5L Phonetic	4/5L Semantic	Tech Finance	Repetition Patterns	Finance & Betting	Chinese Mix	VIP & Gambling	Crypto Utility	Generic Services	Corporate & Enterprise	Generic Phonetic	Brands & Entertainment	Crypto & Memes	Asian Business 8L	Big Tech	Numerical Names	Sports & Generic	Money & Lifestyle	Media & Streaming	Brands & Foundations

Average wallet holding time - Louvain



0	290	462	265	235	242	251	257	257	295	270	250	268	224	244	250	230	240	248	268	243
1	425	372	350	335	424	269	258	355	205	241	308	326	295	287	203	241	278	254	207	227
2	130	145	100	133	119	110	92	98	96	127	112	111	115	102	66	119	121	83	111	115
3	65	105	82	76	66	70	47	59	126	88	54	46	73	92	70	65	85	99	49	39
4	91	80	77	95	88	86	82	84	84	76	80	63	59	92	81	69	80	77	52	89
5	319	215	262	264	266	239	184	225	179	182	206	197	228	205	214	188	239	246	134	189
6	483	547	455	389	317	338	387	269	387	335	301	337	364	327	456	287	317	354	394	382
7	245	181	243	226	269	262	284	231	178	117	209	164	200	217	149	208	183	189	175	139
8	1	1	1	1	3	6	4	1	6	2	10	4	1	6	1	4	9	6	5	11
9	303	339	200	58	181	135	166	152	154	341	130	174	145	126	180	110	155	100	121	142
10	328	152	350	270	311	191	136	247	190	161	184	172	246	178	169	222	214	221	174	153
11	306	99	295	247	361	326	337	186	110	289	282	220	249	327	177	275	281	276	183	182
12	453	352	489	518	450	390	373	481	458	352	402	450	432	421	392	389	374	406	429	371
13	145	246	189	447	274	450	298	307	243	332	450	440	336	348	390	501	330	254	446	349
14	705	929	727	757	676	639	856	748	980	756	630	634	637	582	750	623	599	659	541	586
15	Brands & Foundations	Crypto Utility	Money & Lifestyle	Media & Streaming	Corporate & Enterprise	Finance & Betting	Generic Phonetic	Sports & Generic	VIP & Gambling	Numerical Names	Tech Finance	5L Phonetic	Big Tech	Chinese Mix	Repetition Patterns	Brands & Entertainment	Generic Services	Asian Business 8L	4/5L Semantic	Crypto & Memes

Average wallet holding time - Leiden

0	314	277	505	252	268	249	308	263	282	237	254	269	281	287	285	280	299	254	279	303
1	85	138	119	92	94	87	137	76	44	102	94	90	108	96	108	73	48	85	91	47
2	98	51	52	58	39	70	64	71	84	55	55	73	101	59	51	82	61	60	83	47
3	611	687	386	576	645	452	537	333	311	439	428	366	277	357	288	351	296	378	272	245
4	178	179	183	227	146	191	207	198	316	175	178	217	229	217	182	192	162	203	75	133
5	263	290	146	262	162	230	174	240	134	207	177	180	173	145	191	173	170	189	142	181
6	695	707	842	671	785	733	865	752	859	668	625	656	643	736	650	648	665	616	667	610
7	300	279	230	295	209	293	198	280	218	257	276	283	266	235	273	274	240	252	206	191
8	155	115	158	131	131	152	122	103	102	128	139	117	105	87	95	118	113	132	116	115
9	477	421	492	282	337	345	371	274	377	341	321	321	325	422	355	288	348	289	382	362
10	480	378	421	444	255	368	183	321	186	259	284	226	226	186	292	270	284	170	197	183
11	431	495	335	465	342	523	460	483	380	442	379	392	422	406	411	409	450	390	380	427
12	93	72	94	88	77	113	67	90	63	59	77	70	80	70	66	74	61	59	79	47
13	303	200	339	181	341	58	154	152	155	145	150	134	126	172	100	130	174	109	143	112
14	306	288	84	362	289	244	111	188	307	256	280	317	298	169	289	273	218	272	186	186
15	Brands & Foundations	Money & Lifestyle	Crypto Utility	Corporate & Enterprise	Numerical Names	Media & Streaming	VIP & Gambling	Sports & Generic	Generic Phonetic	Big Tech	Chinese Mix	Repetition Patterns	Asian Business 8L	Tech Finance	5L Phonetic	Brands & Entertainment	Crypto & Memes	4/5L Semantic		

Conclusion

- Proven reproduction of dynamics already observed in decentralized naming systems, especially for short and recognizable names;
- Cybersquatting dynamics emerge;
- Structural position and community affiliation appear to be strong predictors of economic value.

Limitations and Future Works

- Linguistic analysis is constrained by the brevity and irregular structure of usernames;
- Structural communities derived from co-ownership patterns, and not direct social interaction.
- Possible future works include investigation of in-depth temporal dynamics and study of feature importance.



Thank you for your attention!

For questions and inquiries:

- Francesco Maggio f.maggio8@studenti.unipi.it
- Andrea Michienzi andrea.michienzi@unipi.it
- Barbara Guidi barbara.guidi@unipi.it