

LOCARD: R&D for detecting deviant online behaviour

Constantinos Patsakis, March 25th 2021 i-LEAD Research Day











#whoam



```
"name": "Constantinos",
 "surname": "Patsakis".
 "organisation": [{
  "affiliation": "University of Piraeus",
  "role": "Associate professor"
  "affiliation": "Athena Research Center",
  "role": "Adjunct researcher"
 "expertise": ["security", "privacy", "cryptography",
"cybercrime"],
 "email": "kpatsak@gmail.com",
 "webpage": "https://www.cs.unipi.gr/kpatsak"
```



Agenda



- Research and tools for detecting deviant online behaviour from the LOCARD project
- Mainly focused on NLP
- Results



Challenge



- NLP methods have made huge steps forward and present very good results when handling large chuncks of data
- What happens with short messages?





• When we exchange messages, is it always verbal?





Grooming detection



- We collected a large dataset from online chats where grooming was performed and synthesized a filter to detect such behaviours
- Dataset available for researchers and LEAs: https://zenodo.org/record/3560365
- Analysis and results:
 - Nikolaos Lykousas and Constantinos Patsakis. Large-scale analysis of grooming in modern social networks. Expert Systems With Applications, pages 1–1, 2021



Grooming detection (topic detection)



A predator is a person who asks kids to undress in front of the camara

And his bio said he likes meeting young girls

Block foot fetish creep

Don't show the **creeps** anything

Everyone **report** chat **police**

Leave her alone creep

Pervs. This kid is like 12

Report that creep too the police

Report the users asking kids to undress; to authorities not LiveMe

Show your kids

So if they didn't ban people for nudity you would show?

This needs to be **reported** what sort of sick people are ye. She is only 11 YOUR MOM WILL NOW GET A CALL TO KNOW YOU TALK TO 40 years old **creeps**

You **pervs** are **nasty** as f***

block & report nasty stuff

creeps make kids do nasty stuff

he's following lots of **young girls**

pervs stop asking her to undress

report these pedos to police mate

she is a **child** stop asking that

she not leting you **creeps** or sick **perv** seeing her dress or undress ok

she's a **kid**. ...**perv**

they can't ban you if you delete your video after you show

too young this is illegal nd worng lol

try not to undress on stream, it will draw in a lot of creeps

you have **creeps** who made you do **nasty** stuff

you look very ${\color{red}\mathbf{young}}. {\color{blue}\mathbf{there}}$ are lots of pedos on here. be careful



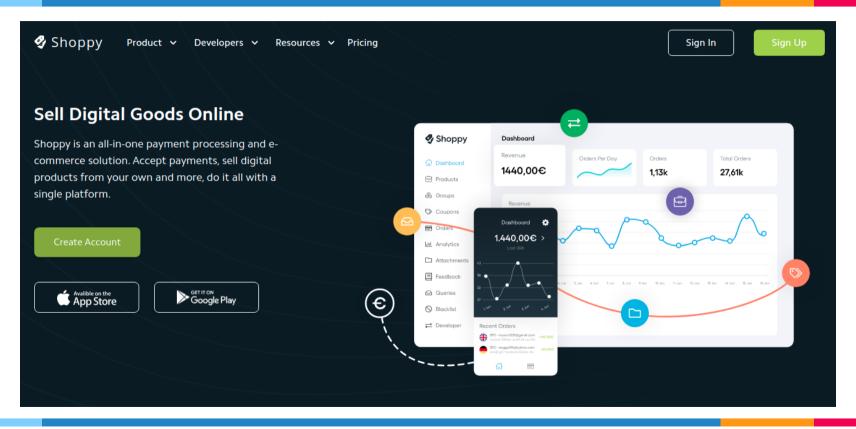
Online marketplaces



- Darkweb marketplaces are known for what is being sold
- What about the surface web?
- What if there is no search option in the marketplace?

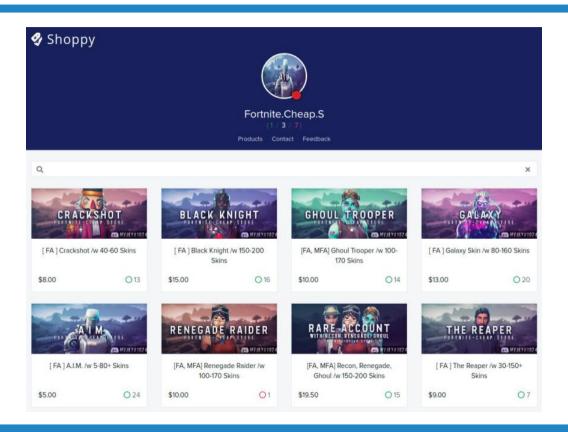






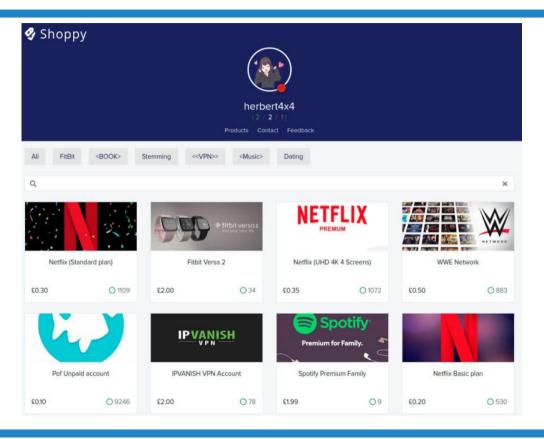
















| Combo List — 528M Yahoo.com | 528,000,000 | 400 |
|--|-------------|-----|
| Combo List 376M Hotmail.com | 376,000,000 | 200 |
| Facebook - 267 Million Records Breach [FULL DB] | 267,000,000 | 500 |
| Combo List 258M Gmail.com | 258,000,000 | 250 |
| Zynga - 213 Million Records | 213,000,000 | 250 |
| Dubsmash Full Database — 162 million — hashed | 162,000,000 | 120 |
| DubSmash - 162 Million Records (FULL SQL DB) | 162,000,000 | 35 |
| MyFitness Pal - 4.62GB (144 Million Records) | 144,000,000 | 25 |
| Xiaomi - 144 Million Records | 144,000,000 | 25 |
| Canva - 137 Million Records | 137,000,000 | 80 |
| 111 MILLION-RECORD PEMIBLANC USA DATABASE COMBO LIST | 111,000,000 | 30 |
| MyHeritage - 92.2 Million Records | 92,200,000 | 75 |
| Houzz - 49 Million Records | 49,000,000 | 300 |
| Facebook DB - 45 Million | 45,000,000 | 265 |
| Chegg - 29 Million Records (Dehashed) | 29,000,000 | 50 |
| Evony - Multiplayer Game : 28.7 Million + 13.8 Million Records | 28,700,000 | 24 |
| Hautelook - 28 Million Records (Full DB) | 28,000,000 | 100 |
| 24 Million LUMINATI PROXY DATABASE HQ [EmailPass] Combo | 24,000,000 | 40 |
| YouNow - 18.2 Million Records | 18,200,000 | 50 |
| 8tracks - 18 Million Records | 18,000,000 | 41 |
| 500PX - Full DB [14.9 million records] | 14,900,000 | 250 |
| Dubsmash — 12 million lines — Private Combo | 12,000,000 | 25 |
| CouponMom / Armor Games - 11 Million Records | 11,000,000 | 41 |
| Cafepress *NEW* 11 Million Records | 11,000,000 | 41 |



Short texts



```
"price": 10,
   "currency": "USD",
   "stock warning": 0,
   "type": "service",
   "stock": 9223372036854775807
   "id": "bdkYlhF".
   "title": "5.7M Facebook Profiles w\/Email".
   "image": {
    "url": "https:\/\s3-storage.shoppy.gg\/shoppy-
gg\product images\8ew6yoHYJqti4pnH8nxyl1Hf23LKwaRAlC8kmQFQ.png",
    "path": "product images\/8ew6yoHYJqti4pnH8nxyl1Hf23LKwaRAIC8kmQFQ.png"
```



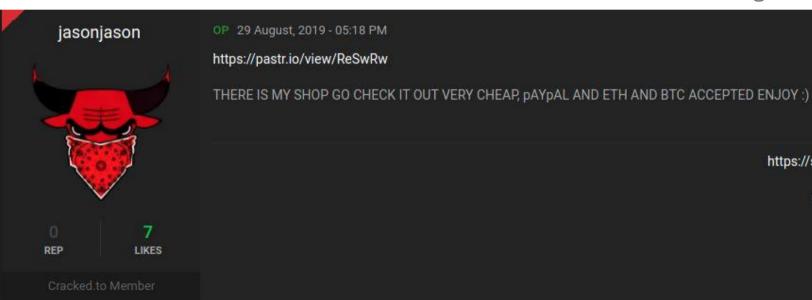
Find users



https://shoppy.gg/@rora1g

slime#0629

• Correlate blackhatworld & cracked.to usernames and signatures





Extract topics



| skin, fortnite, trooper, black, knight, renegade, galaxy, raider, ghoul, recon, expert, ikonik, skull, psn, linkable, aerial, random, assault, season, stack | 466 |
|--|-----|
| spotify, premium, netflix, nordvpn, vpn, family, hulu, pornhub, upgrade, nord, crunchyroll, owner, expire, bulk, uhd, disney, grammarly, screen, student, support | 344 |
| discord, month, paypal, service, btc, day, buy, week, server, term, read, bot, youtube, twitch, contact, apple, cheap, token, rust, music | 219 |
| method, amazon, free, balance, paypal, store, acc, guide, money, carding, make, google, ebook, check, work, ebay, refund, script, bank, link | 205 |
| access, full, minecraft, nfa, minecon, cape, sfa, hypixel, acces, optifine, semi, rank, roblox, unmigrated, vip, mvp, ufa, letter, elite, ninja | 197 |
| key, lifetime, steam, origin, game, uplay, window, license, edition, of- fice, pro, battlefield, standard, global, year, crypto, fifa, deluxe, mi- crosoft, home | 178 |
| config, capture, openbullet, checker, fast, cpm, full, ultra, package, proxy, api, onetap, instagram, aimware, master, mycanal, view, cfg, update, real | 153 |
| premium, hulu, disney, hbo, live, monthly, pass, plan, yearly, commercial, showtime, ad, nba, gold, annual, starz, tidal, fitbit, directv, espn | 135 |
| private, day, combo, pack, email, mail, pass, combolist, usa, fresh, access, valid, hit, list, user, domain, mixed, shopping, guarantee, database | 130 |
| point, dominos, balance, wing, wild, buffalo, free, reward, pizza, usa, subscription, jersey, sonic, drive, mike, payment, amc, grubhub, pts, shein | 128 |
| account, random, gta, shop, dork, crack, site, move, box, source, red, shoppy, mixed, high, target, game, million, keyword, turbo | 119 |
| premium, year, subscription, vpn, auto, lifetime, renewal, adult, security, site, brazzers, monthly, device, avast, pro, membership, pornportal, unlimited, kaspersky, renew | 96 |





Thanks for your time and patience









