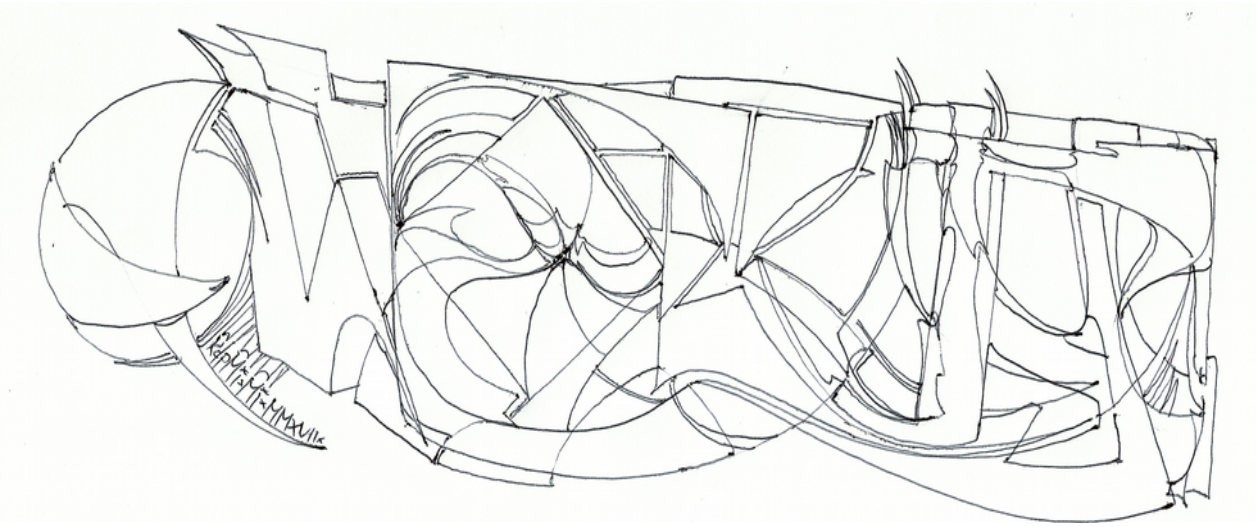
Salvaguardia Arti e Mestieri, in collaborazione con **CRXM** CrossMedia, presenta la:

Ricerchina Della Domenica - I/II-XVIII

SUBSTITUTIO PRICE CRISTHY ET EPHIPANIA DOMINI MMXVII-MMXVIII

Sull'Arte della difesa della privacy nell'era dell'informatizzazione



1. Intro
2. Situazione attuale
3. Resoconto dell'esperimento CLSS
4. Rete di Spionaggio distribuita
5. Interessi Cartacei
6. Brevissime sulle più interessanti tecniche moderne
    1. Incognito Tracking
    2. Browser Fingerprint
    3. Filter Bubble
    4. Super Cookie

Appunti e illustrazioni di KdPM
Testi e consulenza a cura di CRXM CrossMedia
Pubblicazione f.i.p Salvaguardia Arti e Mestieri

#1. **Intro** – dove vengono brevemente esposti gli argomenti della trattazione successiva e suggeriti dei link di approfondimento. E provo anche di fare il simpatico.

Fin da prima della diffusione degli *smartphone*, in molti già conosciamo - più o meno propriamente - la così detta "pubblicità basata sugli interessi", così come la "modalità di navigazione anonima".

Nel secondo semestre 2018, continuano gli aggiornamenti agli accordi per la protezione di dati personali da parte delle *Big Data Companies* internazionali, costrette ad adeguarsi all'entrata in vigore del Regolamento UE 2016/679 in materia di trattamento dei dati personali - perlopiù producendo documenti noiosamente incomprensibili. Qualcuno avrà anche sentito del quasi miliardo di Euro della sanzione inflitta a Telecom e altri provvedimenti del garante per la Privacy.

Al Museo Diocesano di Milano è in corso una mostra fotografica dedicata al gruppo MAGNUM, "*L'ITALIA DI MAGNUM. Da Cartier-Bresson a Paolo Pellegrin*". Interessante anche se di modeste dimensioni, è comunque illuminante nel ricordare come questi Big dell'immagine abbiano collaborato al PWB bellico e post. E poi inquietanti dichiarazioni di pratiche di controllo sociale di massa come il Bubble Filtering o lo scandalo Cambridge Analityca, travisate a incrostare l'opinione comune di leggende metropolitane alla Blue Whale.

I *Call Center* intanto continuano a mitragliare l'intimità personale e familiare, mentre le *Big Data Companies* non sono ancora percepite come aziende di promozione pubblicitaria. *Tracciamenti*, *Fake News*, manipolazioni, ad arrivare alla definizione di **informatica** stessa: "*la scienza di informare*"... Mentre qualcuno rimane disinteressato agli argomenti di sociopolitica, altri sono molto preoccupati. Ma il Mondo è bello perché è vario :)

L'intento di questo excursus è quello di radunare e analizzare le informazioni riguardanti questi fenomeni per capire cosa sia in atto, orientandomi dal punto di vista dell'utente finale internet, quale mi considero: ingenuamente *libero cittadino* che richieda accesso a *libere informazioni*; non vuole annoiare con un approccio tecnico, tanto meno infangarsi nei neandri *lobbystici* delle motivazioni. Anche se, al solito, lo farà.

Scoprirò che il detto "*l'informazione accelera*" rispetta la teoria della relatività, vale a dire che prima o poi la coscienza implode.

Illustrazione 1: "L'ITALIA DI MAGNUM. Da Cartier-Bresson a Paolo Pellegrin", 2018, Museo Diocesano, Milano

#2. **Situazione attuale** – in cui lagno le rivelatesi aberrazioni tecnologiche moderne

Fatto: se comunichi oggi, molto probabilmente utilizzi uno strumento elettronico digitale per farlo.

Altri fatti riscontrabili da chiunque durante l'utilizzo delle tecnologie digitali odierne sono:

- Visitando il sito X dal device Y, compaiono pubblicità di X sui device Y e Z.
- Inviando dal device Y un SMS contenente X, compaiono pubblicità di X sul device Z.
- Utilizzando il servizio X, vieni contattato da Y e Z con proposte di servizi affini a X.

Questi tipi di "nuova pubblicità" vengono definiti annunci personalizzati, o annunci basasti sugli interessi, o anche pubblicità personalizzate etc. Al fine di limitarli o eliminarli paiono inutili precauzioni quali non utilizzare *account social*, navigare *in incognito*, utilizzare *software open* o *negare il trattamento* dei propri dati, seguire guide sull'utilizzo dei *cookie*, e quant'altro - ad arrivare inutilmente fino a creare *account honeypot* o *virtual airgap* (non spaventarti per la sventagliata di terminologie tecniche, tutto verrà spiegato facilmente a seguito; basti capire che qualunque livello di competenze tecniche si possieda lo sforzo è inutile! vedi #3,#4,#6)

La domanda è come sia possibile e perché avviene questo. La risposta plausibile è quella errata, la stessa data dagli operatori di *call center*: "*avrà cliccato su qualche sito dando l'autorizzazione*".
Ma che la realtà dei fatti sia più perversa viene presto suggerito anche da altri inquietanti fenomeni, quali ad esempio:

- Utilizzi il servizio X in *incognito*, ti vengono comunque proposti servizi affini a X.
- Nel privato di casa parli di X, vieni contattato con proposte di investimenti su X.
- Proponi accordi a Y, Y viene contattato da Z con proposte concorrenziali alla tua.

Le risposte plausibili diminuiscono in maniera proporzionalmente inversa all'ansia: qualcuno ti spia. *Cortona* e *Siri* ti ascoltano? Intercettazioni ambientali? Illeciti criminosi? Caso? Psicopatia paranoide? [spoiler] A parte il fatto che, comunque la si metta, la normativa sulla privacy viene ampiamente violata…

Illustrazione 2: Kdpm, "Corrispondenza violata"- inchiostri su busta da lettere – 2016

#3. **Resoconto dell'esperimento CLSS** – Ciò La Sim Sconnessa

Per cercare di chiarire attuo il semplice esperimento descritto a seguito, battezzato ***CLSS - Ciò La SIM Sconnessa***. Ciò perché mi piace.

Per lo svolgimento la cosa più difficile da recuperare è un essere umano cavia (**K**) disposto a utilizzare continuativamente per le sue comunicazioni un telefono cellulare (**HW**) retrodatato di almeno 15 anni - volendo imbarazzante, volendo radical chick.
Scegliamo un modello che supporti solo i protocolli di comunicazione *GSM* e *SMS*.
- è molto importante che il costo dell' **HW** scelto sia inferiore ai 20€ se si vuole accentuare il fattore drammatico di questo esperimento.

Ora utilizziamo con il suddetto **HW** una *SIM* "*solo voce*" (NON abilitata al traffico dati), facendola attivare a un benevolente *prestanome* – nel nostro caso (**V**). Per il contratto, scegliamo un operatore telefonico che abbia un chiaro e NON precompilato accordo sul metodo di trattamento dei dati personali. [REFUSO]: è in realtà questa la cosa più difficile da recuperare per lo svolgimento di questo esperimento.

Per esagerare, apriamo anche alcuni account adhock per i principali servizi internet - quali ad esempio Gmail e Facebook - in cui ci registriamo con nomi e numeri telefonici differenti, comunque senza MAI rendere noti il reale numero telefonico di **HW** o dati reali su **K**; li chiameremo *account honeypot*.

Infine lasciamo che **K** continui la sua Vita utilizzando internet e i suoi account, ed eccoci qui: con questa configurazione si sarebbe portai a pensare che **K** debba poter usare **HW** senza che nessuno - organi Statali e tecnici impiccioni a parte - possa lecitamente accedere alle sue comunicazioni, o alla sua identità, quindi sia impossibile un vero tracciamento utente… [ingenuo].

Invece, anche con questa configurazione avviene che:
> se **K** invia da **HW** *SMS* contenente X, riceve su **HW** *SMS* inerenti X da sconosciuti Y e Z;
> se **K** si iscrive ad un servizio X, riceve su **HW** *SMS* inerenti X da sconosciuti Y e Z;
> [… etc.]

Inoltre, compaiono allarmanti anomalie inaspettate, quali:
> **K** viene identificato spesso come **V**;
> **K** viene spesso associato ai dati dei suoi *accounts honeypot*;
> Le telefonate e contatti non richiesti ricevuti su **HW** sono eccessivi, indagatori, aggressivi o palesemente sarcastici.

in chiusura:

> Nessun **K** è stato maltrattato nel corso di quest'esperimento - se non dagli operatori di *call center.*
>**V** è felicemente al mare e pensa sicuramente ad altro.
> **HW** è ancora funzionante, continuando a ricevere fantastiche promozioni giornaliere indesiderate.
> Siamo speranzosi che questo testo chiarificatore possa anche aiutarci ad uscire dai tracciamenti "*antiterroristici*" e "*antifrode*" sicuramente in atto sugli *account honeypot ;)*

#4. **Rete di spionaggio distribuita** – quando ogni strumento si tramuta in una cimice e ogni essere umano in un informatore.

Per giustificare quanto emerso dall'esperimento CLSS, schematizziamo la semplice situazione in cui la cavia in esame (**K**) voglia comunicare con un' altro essere umano (**Y**), e rappresentiamo con → la catena di "*passaggi di mano*" che compie il messaggio (♥), segnalando con [∞] quelli che implicano più passaggi interni (come ad esempio le procedure di smistamento nelle spedizioni postali o nella transazione dei pacchetti TCP-IP).

Se **K** *parla direttamente* con **Y**, possiamo riassumere:
K → etere → **Y**

Quando **K** decide di *spedire una lettera*:
K → posta[∞] → **Y**

Optando per una *e-mail*:
K → Application Software[∞] → Sistema Operativo[∞] → Hardware[∞] → NET[∞] ... per poi risalire la catena sino a **Y.**

Generalizzando dunque ad ogni comunicazione digitale otteniamo una catena del genere:
K→ AS[∞] → SO[∞] → HW[∞] → NET[∞] → HW[∞] → SO[∞] → AS[∞] → **Y**

Considerando ogni passaggio [∞] come un "*layer informatico*", ogniuno di questi *layer* ha accesso di fatto al messaggio ♥, e in ogni passaggio → e [∞] la comunicazione risulta esposta contemporaneamente ad almeno 2 delle seguenti classi di manipolazioni:

1. ***Tracking*** quali cookie, *Browser Fingerprint* et similia, in grado di *riconoscere* univocamente e nel tempo il device utilizzato al fine di creare una *profilazione utente*.
2. ***Exploit*** - intendendo in generale *Virus, Trojan* - *altri termini sinistri* - AS che fanno da SO, SO che fanno da AS; e in genere le tecniche che sfruttano i *layer* tra AS e HW.
3. ***MITM*** a.k.a *Man In The Midlle* - che chiameremo così anche quando effettuate con software automatizzato - per definire qualsiasi metodo sfrutti i *layer* tra HW e NET.

Si palesa chiaramente come la possibilità di accedere alla comunicazione tra **K** e **Y** aumenti esponenzialmente alla complessità del sistema di scambio scelto (16*X>=2 è maggiore di 2 ;))

Nell'esperimento CLSS, **K** limita unicamente i passaggi nella catena di sua competenza, mentre suo malgrado probabilmente **Y** riceverà ♥ tramite uno *smartphone*; mentre nei primi tre passaggi della catena il messaggio si può considerare "protetto" - illeciti a parte – il tracciamento dell'*SMS* inviato da **K** tramite **HW** rimane possibile sfruttando i *layer* SO o AS messi a disposizione da **Y**: lo *smartphone* utilizzato per ricevere ♥ traccia la comunicazione andando ad aggiungerla sia nel profilo di **Y**, sia in quello di **K**, i cui dati per il riconoscimento vengono ricavati direttamente da quelli in possesso di **Y**. Ecco che - come nella leggenda metropolitana sugli aggeggi cinesi che un giorno verranno attivati da remoto per conquistare ogni casa del globo :)) - lo *smartphone* di **Y** è tramutato in una "*cimice*" mentre **Y** a sua insaputa in un "*informatore*".

Dall'esperimento si riscontra inoltre come anche lo stesso operatore telefonico venda i dati dei propri utenti, pur senza alcun avviso o consenso; è infatti l'unico a conoscenza dell'associazione di **HW** con **V**, eppure come esposto **K** viene chiamato **V** dai *call center* che lo contattano su **HW.**

#5. **Interessi Cartacei** – di come il fine ultimo dell'era digitale rimanga sarcasticamente analogico.

Per arrivare ad un dunque non politico tocca tralasciare appositamente i decisivi interessi Governativi, NSA, CIA, KGB, così come anche evitare di scomodare i demoni che agiscono a valle dei protocolli di comunicazione – reminiscenze di Echelon, Carnivore, Zero-Day et altri miti – evitando infine di cadere nelle piaghe degli illeciti perpetrati direttamente dalle proprie compagnie telefoniche, nostrane e non… Prenderemo in considerazione inizialmente la questione monetaria, tramite l'esempio principe: *Google*.

Chiedendo cosa sia Google, la risposta ricevuta sarà perlopiù: "un *motore di ricerca*".
[cit.] *È per questo che la specie umana è un fallimento e mi fa incazzare* [/cit.]
Il fatturato della società di Mountain View denominata *Google LLC* (ossequiosamente **G**) per il 2017 ammonta a 110,86 miliardi di dollari, con un utile netto di 12,6 miliardi. Questi 110.860.000.000$ non provengono certo dal *motore di ricerca* - gratuito. Argutamente si può puntualizzare come **G** fornisca una miriade di altri prodotti, dal S.O. Android a Chrome, alle piattaforme Youtube, Google+, Gmail, Maps, Heart, StreetView, Analytics, più quasi un'altra cinquantina di servizi. Ancora, tutti, gratuitamente. Quali sono dunque i servizi da cui **G** incassa 12.600.000.000$ di utile annuo? Risulta che AdWords, il servizio di advertising, sia la principale fonte di remunerazione, che nel 2013 ha permesso di guadagnare più di 50 miliardi di dollari.
Per farla breve, dovrebbe semplicemente essere illuminante la seguente *rivelazione*:
G è una società di advertising - **promozione pubblicitaria**, in Italiano.

Il *WebTAP privacy project* ha rilevato che il codice **G** è installato nel 75% del milione di siti più usati - seguito da Facebook al 25%. Significa che **G** non vende pubblicità solo sul proprio sito, bensì su più di 2.2 milioni di altri siti web, e oltre 1 milione di app, in costante aumento. Significa che **G** registra informazioni ogni volta che utilizziamo uno di questi quasi quattro milioni di "servizi" o "informazioni" traccianti - seppur non di proprietà della società **G** – o che contattiamo un altro utente - perlopiù all'insaputa degli utenti stessi.

In questo contesto i servizi forniti gratuitamente da **G** - così come dalle altre Big Companies quali Facebook, Amazon etc- sfuggono alla filantropia, concretizzandosi come ulteriori affinamenti del sistema di riconoscimento del target, così come succede anche con una registrazione ad un portale web o l'utilizzo di social network ad esempio: i nostri dati sensibili forniti volontariamente – ma inconsapevolmente – se divulgati riescono infine ad etichettare e radunare le cronistorie registrate, fornendo quella che si può definire una schedatura completa.

La promozione mirata genera miliardi di dollari in utili perché evidentemente non si basa solo su una semplice *estrapolazioni momentanea* del nostro utilizzo internet - come portati a credere - bensì sulla compravendita di una schedatura non autorizzata ma reale, più dettagliata di quanto noi stessi possiamo immaginare, di noi stessi – la nostra locazione, i nostri contatti, le nostre usanze , frequentazioni, reddito, stato d'animo, fisico, inclinazione all'acquisto e quant'altro i complessi algoritmi, creati anche a suon di concorsi, possano estrapolare dai nostri ingenui click.
Per intuire la portata e il dettaglio di questo sistema si consideri anche la molteplicità di società che operano il tracciamento a vari livelli, con varie tecnologie, alcune atte anche evidentemente a continuare il tracciamento contro la volontà esplicita dell'utente stesso (vedi #6.4). Il nostro profilo utente viene dunque aggiornato, integrato e condiviso continuamente tra le Big Data Societies, a creare una speciale block-chain ridondante e distribuita di profili utente univoci e dettagliati, usati *anche* per compilare quei pacchetti da miliardi di dollari da rivendere alle aziende per effettuare *promozione mirata*. E così – con i metodi a seguito esposti - fan tutte, pare.

#6. **brevissime sulle più interessanti tecniche moderne** – dove viene esposto il come

Per concludere seguono alcuni appunti – sicuramente incompleti e vagamente confusionari, ma poco tecnici e integrati - di interessanti nozioni introduttive sulle questioni di navigazione web e privacy, fornite principalmente dall'ottimo motore di ricerca DuckDuckGo.com, e dal sito Eletronic Frontier Foundation. Consiglio agli anglofoni la consultazione degli articoli dai siti originali [1] [2]

#6.1. Incognito Tracking

Potrebbe sorprenderti che quelle pubblicità continuino a starti attorno anche quando navighi in "*modalità incognito*", e rimarrai imbarazzato quando compariranno nella navigazione normale le pubblicità di quelle ricerche fatte in *incognito*. Questo succede perché la *modalità incognito*, la *navigazione anonima*, o altre tecniche di *browsing privato*, non sono mai realmente private! Se non lo sapevi, non sei solo: il 67% delle persone risulta sovrastimare le capacità del *browsing privato*.
La modalità privata perlopiù inibisce l'accettazione dei *cookie HTTP* e aiuta a pulire la cronologia del browser, ma questo non impedisce alla rete di siti di recuperare un *cookie* preesistente continuando il tracciamento utente, come ci dimostra chiaramente nothingprivate.ml. Inoltre, come descritto in seguito, vengono utilizzate anche tecniche di "*browser fingerprint*" e "*Super-Cookie*" per continuare il tracciamento dell'utente anche in modalità anonima o con una politica restrittiva impostata ai cookie HTTP.

#6.2. Browser FingerPrint

Il nome dato al metodo principale per identificare e tracciare univocamente e nel tempo gli strumenti su internet. Proprio come ogni uno di noi, ogni strumento nella rete Internet possiede infatti una sua impronta digitale unica: i siti web ad esempio nel momento in cui vengono "visitati" ricevono una stringa di versione dal browser in uso (*User Agent)* che ne indica l'ambiente operativo, quali plug-in sono installati, e dozzine di altre informazioni che vengono usate per creare un ID univoco che può essere usato per tracciarti.
"*New Cookie Technologies: Harder to See and Remove, Widely Used to Track You*" è un'esaustiva analisi tecnica sull'argomento, effettuata da Seth Schoen datata settembre 2009 (in lingua inglese).
Il sito *ClickClickClick* mostra in tempo reale come queste informazioni possano essere usate.
WebShadowAnalizer di CRXM attesta come l'assegnazione - pur tramite uno dei meno sofisticati metodi di fingerprint - sia comunque persistente nel tempo e indipendente dal metodo di connessione o precauzione applicata.

#6.3. Filter Bubble

Recentemente abbiamo assistito a sdegno giustificato sulle "*fake news*" e sul fatto che molte persone vivono in una "*camera dell'eco*" online. Le compagnie che effettuano il tracciamento usano il profilo utente ottenuto per filtrare i risultati da mostrarti, basandosi su quello che *reputano* sia più probabile che sceglierai. Questo è correntemente conosciuto come Filter Bubble. È una forma di censura che può essere usata per influenzare l'opinione pubblica – come anche nello scandalo Cambridge Analytica. Senza nulla di personale da aggiungere, lascio al lettore volonteroso le considerazioni sulle ripercussioni sociali, magari suggerendo questo TED talk di Eli Parsier.

#6.4. Super-Cookie

Come accennato, il *Browser FingerPrint* è solo uno dei modi di aggirare le preferenze utente nelle questioni di privacy. Sempre già nel 2009, un guppo di ricercatori della UC Berkeley ha rilasciato uno studio dal nome "*Flash Cookies and Privacy*", nel quale mostra come anche i *Local Shared Objects* di Adobe – leggi *plugin Flash* - siano utilizzati per tracciare gli utenti internet. Lo studio mette in luce come questo tipo di tecnica venga estensivamente utilizzata per aggirare deliberatamente la "*politica cookie HTTP*" impostata dell'utente nel proprio browser, in quanto riportanti le stesse informazioni registrate nei cookie HTTP al solo fine del recupero delle stesse in caso l'utente decida di cancellarle, rendendo dunque possibile assegnare ad un utente un cookie HTTP che "ricompare" nonostante venga cancellato. Esistono altri metodi per creare questo tipo di "*Super-Cookie*" fuori dal controllo dell'utente finale, oltre ai citati Local Shared Objects di Adobe, risulta possibile sfruttare il DOM storage dell'HTML 5, i cookie Microsoft Silverlight, la Microsoft Internet Explorer User Data Persistence, e Google Gears.

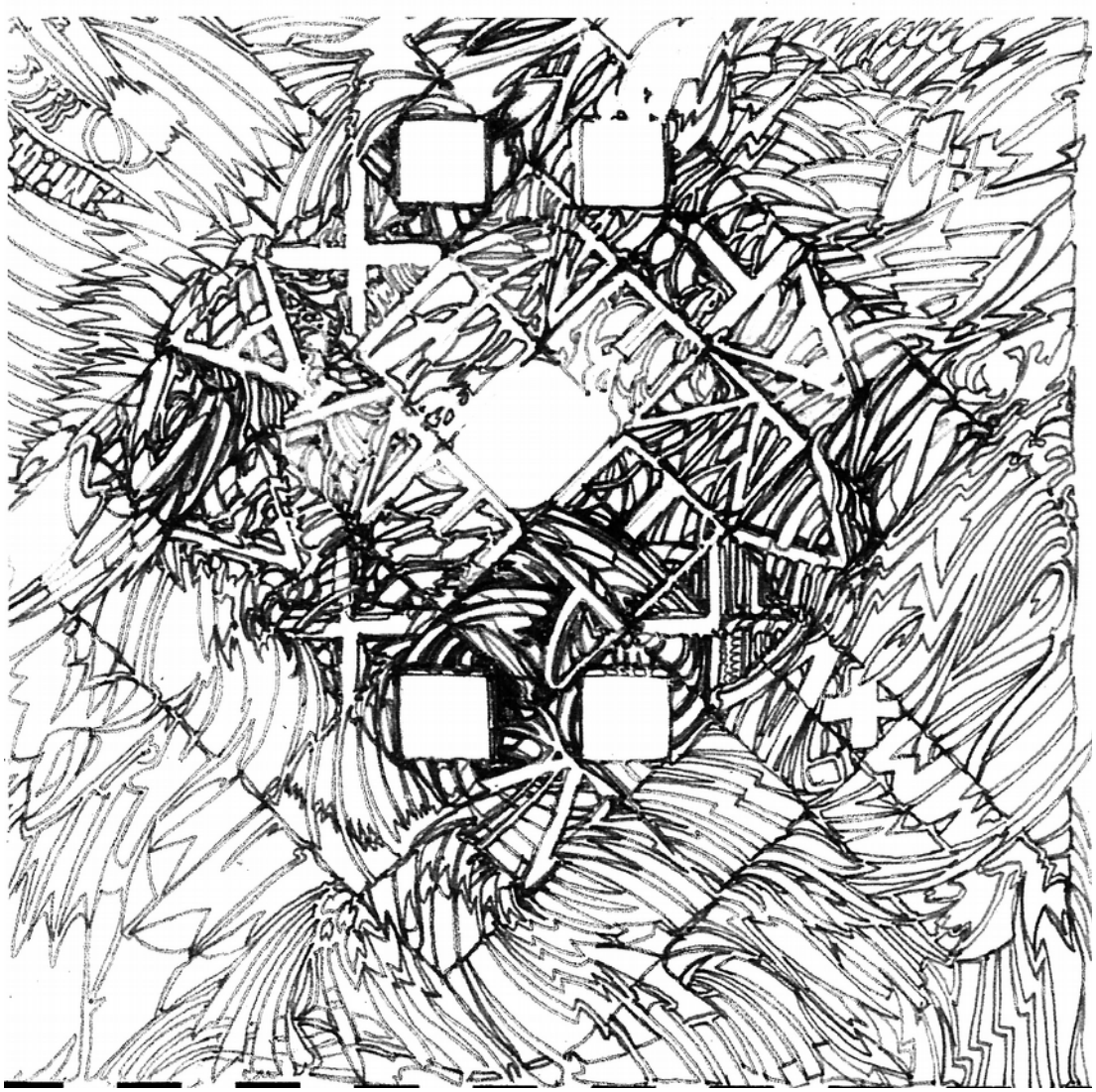


Illustrazione 3: KdPM, "tempo di salutarci. Ma non preoccuparti, potrai sempre sapere dove sono e cosa faccio. Basta seguire la traccia. Bless" - scansione digitale, 2018