

Anàlisi automàtica de tiquets de phishing

mitjançant el processament del
llenguatge natural

Jaume Casals Vilaplana

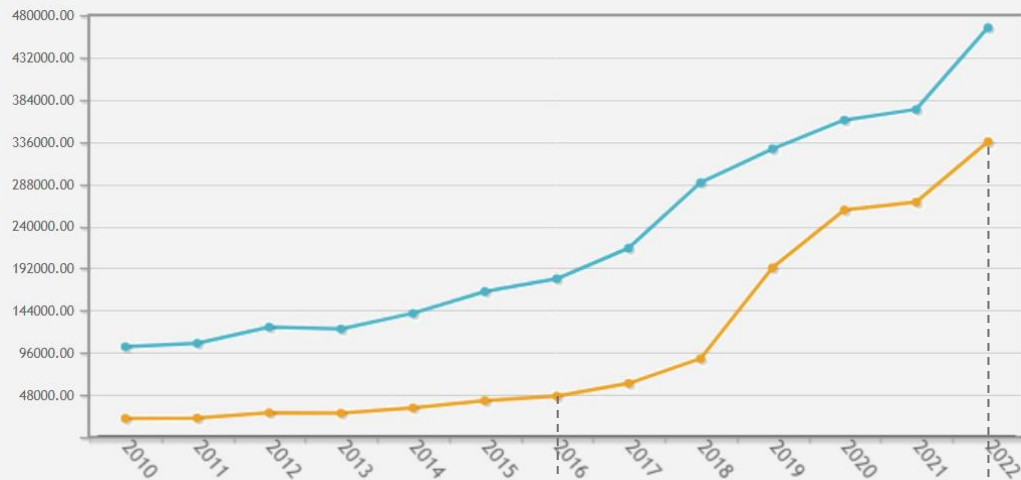




01

Context i abast

Nombre d'estafes nacionals



465.906 casos

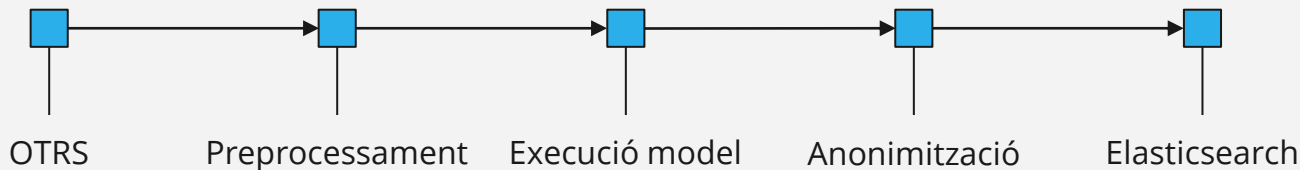
Estafes

335.995 casos

Estafes informàtiques

Augment del 370%

Abast



Objectius

Estudi de l'estat de l'art

Anonimització de la sortida

Eina descàrrega de tiquets

Emmagatzematge del resultat

Preprocessament de les dades

Implementar pipeline

Entrenament del model NLP

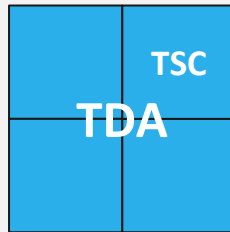
Accés mitjançant API

Actors implicats

Agència

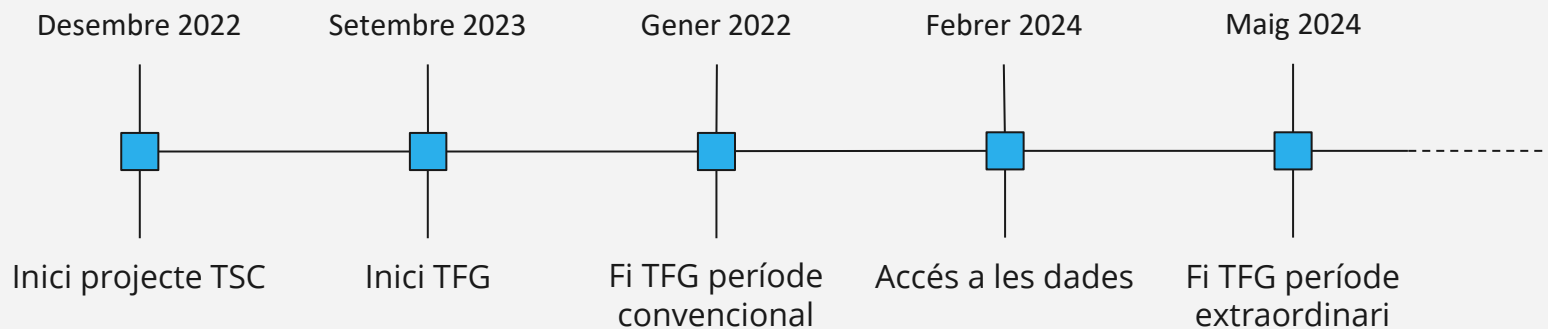


i2CAT



InLab FIB

Cronologia





02

Exploració teòrica

Anàlisi d'un tiquet

- 7. Usuaris afectats
- 8. Accions de mitigació
- 9. Accions de control
- 10. Mail de la víctima
- 11. Mail de l'atacant
- 12. Assumpte del correu
- 13. URL de l'incident

[6] Article #1

[3] From: [10] "User3 generic" <user3@gmail.com>

To: Raw

Subject: Possible correu phishing (exemple)

Created: 12/03/2023 20:03:25 (Europe/Madrid) by customer

Attachment: captura_errors.jpg (159.3 KB)

Bon dia,

Escric per informar d'un correu electrònic sospitosos que he rebut a la safata d'entrada. El missatge sembla que és un intent de phishing i em preocupen els possibles riscos de seguretat.

[11] Detalls del correu electrònic:

- [12] - Remitent: service@securemail.com
- Assumpte: Urgent: Verificació de compte requerit
- Data/Hora de recepció: 2023-12-03, 16:15 AM

El correu electrònic diu procedir d'un proveïdor de serveis legítim i em demana que verifiqui urgentment el meu compte fent clic a un enllaç que apareix al missatge. El missatge també adverteix de greus conseqüències si no actuo immediatament. El missatge inclou un enllaç que sembla sospitosos i no hi he fet clic (urlvirus.com).

[13] Adjunto una imatge del correu que m'ha arribat.

Com a mesura de precaució, m'he abstenut de fer clic a cap enllaç ni facilitar informació personal. En canvi, informo d'aquest incident al departament de ciberseguretat perquè l'investigui.

No he experimentat cap activitat inusual amb el meu compte, i aquest correu electrònic sembla sospitosos donat el to urgent i la naturalesa inesperada de la sol·licitud. Volia posar-ho en coneixement de l'equip per garantir la seguretat de la informació de la nostra organització.

Gràcies per la seva ràpida atenció a aquest assumpte.

Una cordial salutació,

John Doe
Departament de Finances

Benvolgut John Doe,

Gràcies per informar aquest incident amb promptitud. Després d'una anàlisi inicial, el correu electrònic del qual ens ha informat sembla que és un intent de phishing que només ha afectat a unes poques persones del departament de finances. Els correus als que els hi ha arribat són els següents:

- [7] - user1@gmail.com
- user2@gmail.com
- user3@gmail.com
- user4@gmail.com

Les nostres mesures de seguretat han estat activades per bloquejar qualsevol amenaça potencial associada a l'enllaç proporcionat. Estem duent a terme un examen exhaustiu per identificar la font i qualsevol impacte potencial als nostres sistemes.

A la llum d'aquest incident, recomanem que no feu clic a cap enllaç ni descarregueu cap fitxer adjunt de correus electrònics sospitosos. A més a més, a partir d'ara, sigueu previnguts i verifiqueu la legitimitat de correus electrònics inesperats posant-vos en contacte amb el supòsit remitent a través d'un canal de comunicació conegut i independent.

[8] El nostre equip continuarà la investigació i aplicarem les mesures de seguretat necessàries per minimitzar els possibles riscos. Aquestes mesures inclouen bloquejar el URL de la pàgina i bloquejar els remitents del correu.

[9] A partir d'ara, les comunicacions urgents només es faran a través del portal específic de l'empresa, evitant així dubtes sobre possibles correus de phishing. Mantingueu el portal obert per estar actualitzat de les últimes notifikacions.

Si observeu qualsevol altra activitat sospitosa o rebu correus electrònics similars, us preguem que ens ho comuniqui immediatament.

Rebeu una cordial salutació,

Departament de Ciberseguretat



OTRS i Elasticsearch

Gestió de tiquets



Znuny

Base de dades

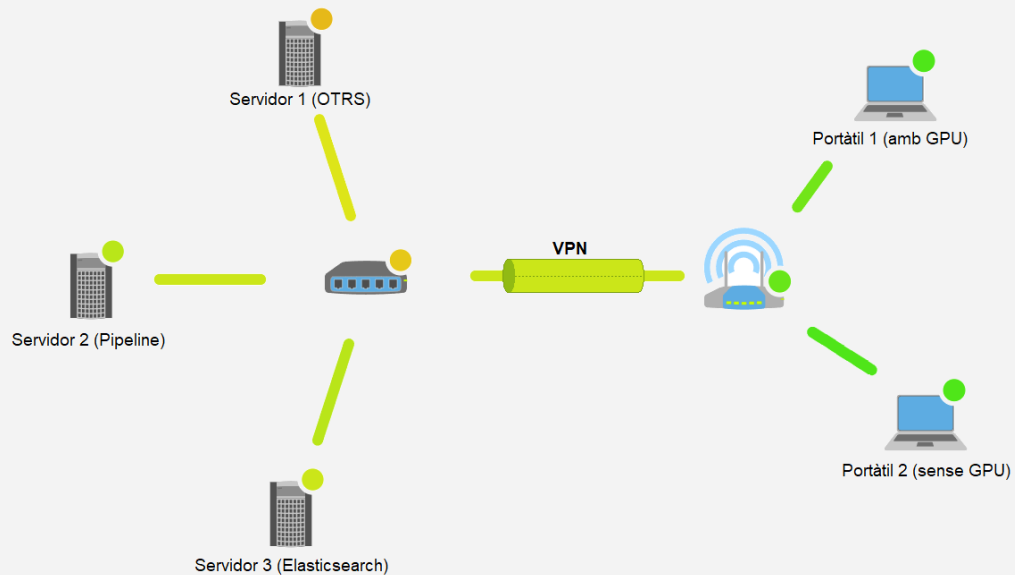




03

Execució pipeline

Diagrama de la xarxa



Extracció de dades

Eina principal: PyOTRS

- Accés API
- Tiquet en format HTML

```
</style>
</head>
<body>
  <table role="presentation" border="0" cellpadding="0" cellspacing="0" class="body">
    <tr>
      <td>&nbsp;</td>
      <td class="container">
        <div class="content">

          <!-- START CENTERED WHITE CONTAINER -->
          <span class="preheader">This is preheader text. Some clients will show this text as a
          preview.</span>
          <table role="presentation" border="0" cellpadding="0" cellspacing="0" class="main">

            <!-- START MAIN CONTENT AREA -->
            <tr>
              <td class="wrapper">
                <p>Bon dia,</p>
                <p>Escric per informar d'un correu electrònic sospitós que he rebut a la safata
                d'entrada. El missatge sembla que és un intent de phishing i em preocupen els possibles riscos de
                seguretat.</p>
                <p>Detalls del correu electrònic:</p>

                <p>Remitent: service@securemail.com</p>
                <p>Assumpte: Urgent: Verificació de compte requerit</p>
                <p>Data/Hora de recepció: 2023-12-03, 10:15 AM</p>
                <p>Adjunto una imatge del correu que m'ha arribat.</p>

                <p>El correu electrònic diu procedir d'un proveïdor de serveis legítim i em demana que verifiqui
                urgentment el meu compte fent clic a un enllaç que</p>

                <p>apareix al missatge. El missatge també adverteix de greus conseqüències si no actuo immediatament.
                El missatge inclou un enllaç que sembla sospitós i no hi he fet clic (urlvirus.com).</p>

                <p>Com a mesura de precaució, m'he abstingut de fer clic a cap enllaç ni facilitar informació
```



Preprocessament

- Eliminar repeticions
- Detectar referències

Adjunts que es llegeixen



Text



Email



PDF



Word



Word (antic)

Ticket#3500391 — Update OTRS

▼ Article Overview - 2 Article(s)

NO.	SENDER	VIA	SUBJECT	CREATED	
2	Brandie Watson	OTRS	File Preview Demo	10/01/2019 12:58 (Europe/Berlin)	3
1	David Brown	Phone	Update OTRS	10/01/2019 12:56 (Europe/Berlin)	3

▼ #2 — File Preview Demo — Eugene Krymov — 10/01/2019 12:58 (Europe/Berlin) via OTRS by Eugene Krymov

Reply to note | Mark | Print | Split

File Preview Demo

OTRS 6 - Admin Manual.docx
Unknown — 11.6 KB

otrs_admin_book.pdf
PDF — 18.6 MB

otrs_logo.png
Image — 28.6 KB

[All attachments](#)

Download all attachments as ZIP

Si observeu qualsevol altra activitat sospitosa o rebeu correus electrònics similars, us preguem que ens ho comuniqueu immediatament.

Rebeu una cordial salutació,

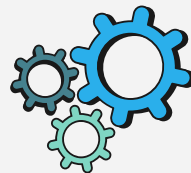
Departament de Ciberseguretat

--

Super Support - Waterford Business Park
5201 Blue Lagoon Drive - 8th Floor & 9th Floor - Miami, 33126 USA
Email: hot@example.com - Web: [1]http://www.example.com/
--

12/03/2023 20:03 (Europe/Madrid) - user3 generic wrote: > Bon dia,
> Escric per informar d'un correu electrònic sospitós que he rebut a la safata
> d'entrada. El missatge sembla que és un intent de phishing i em preocupen els
> possibles riscos de seguretat.
> Detalls del correu electrònic:
> - Remitent: service@securemail.com

Eliminació de signatures



Flan-T5-base/FLOR

...

Si observeu qualsevol altra activitat sospitosa o rebeu correus electrònics similars, us preguem que ens ho comuniqueu immediatament.

Rebeu una cordial salutació,

Departament de Ciberseguretat

SecureTech Solutions Inc.
123 Cybersecurity Way, Suite 181
TechCity, SecureZone 54321
+34 555 123 456
www.securetechsolutions.com

Aquest correu electrònic i qualsevol fitxer adjunt són confidencials i estan destinats exclusivament a l'ús de la persona o entitat a la qual s'adreça. Si heu rebut aquesta comunicació per error, aviseu immediatament al remitent i elimineu el missatge del vostre sistema. Qualsevol ús, divulgació o distribució no autoritzats està estrictament prohibit.

Si us plau, tingueu precaució quan compartiu aquest correu electrònic amb persones que no necessiten accés al seu contingut. La informació continguda aquí pot ser sensible i subjecta a restriccions legals.

En un esforç per minimitzar l'impacte ambiental i salvaguardar la informació sensible, si us plau, absteniu-vos d'imprimir aquest correu electrònic tret que sigui absolutament necessari.

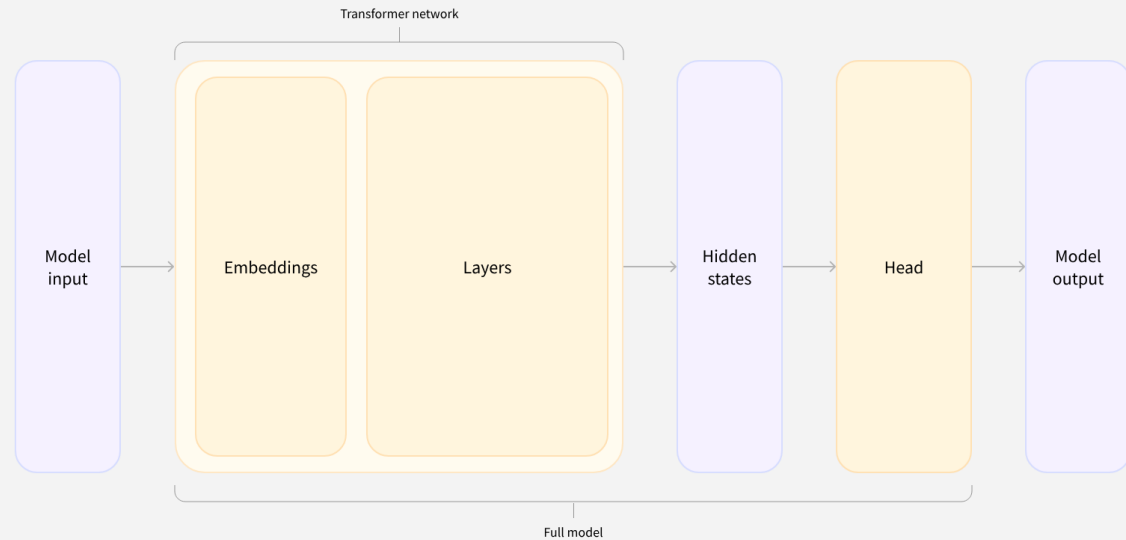
RoBERTa



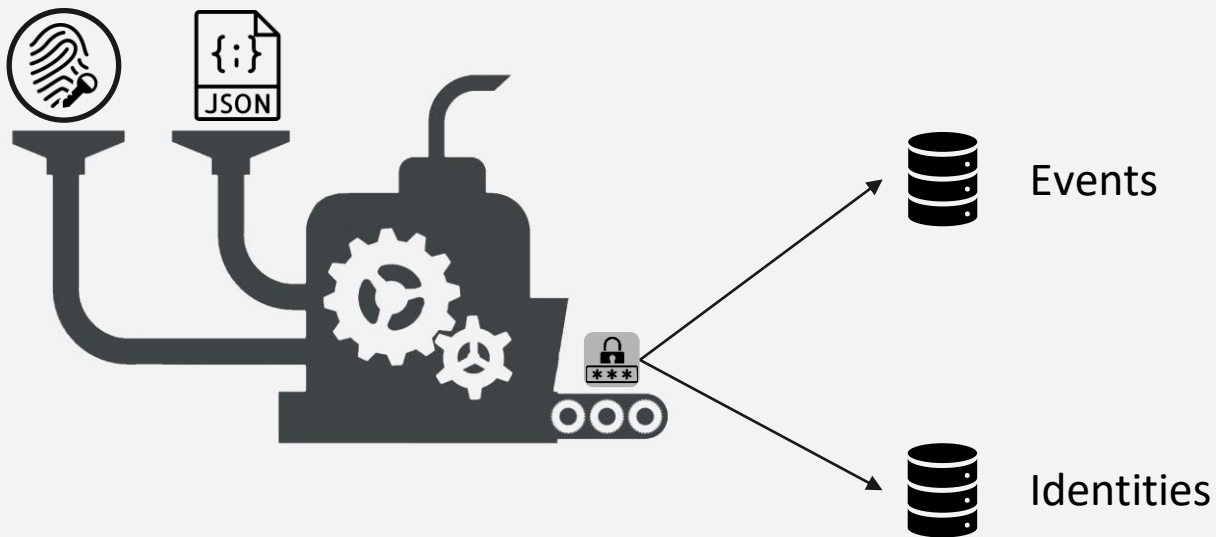
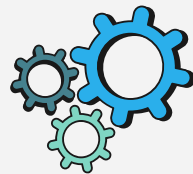
Spacy

Posició	Frase
0	Bon dia,
1	Escriu per informar d'un correu...
17	Rebeu una cordial salutació,
18	Departament de ciberseguretat
19	SecureTech Solutions Inc.
22	+34 555 123 456
23	www.securetechsolutions.com
24	Aquest correu electrònic i ...

Execució model NLP



Anonimització



Elasticsearch

Events

Hash de tots els camps

```
{
  "file": "1234567890.txt",
  "affectedUsers": [
    "73129e8a41a791aa2cbd6a76481adcc251af17f81f9426377906dd50d93eb2ca",
    "d7a63478a91895949221a1848a5afd083e2e5d7b5cf914c50da7e1b403608989"
  ],
  "mitigationActions": [
    "24eaa6c6fc94c72663822891181c91a93fd4b94d7b712492341980d416692c98",
    "1e8a5a73a110db403cca0657a3ca958b316c94f1a0633d73bedd32e17333c83a"
  ],
  "controlActions": [
```

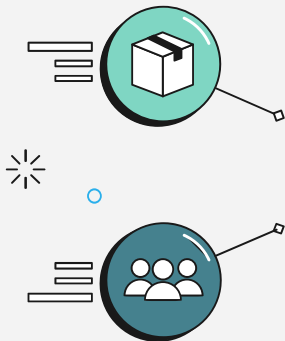
Identities

Parella clau-valor

```
{
  "value": [
    "user1@gmail.com",
    "User Lastname"
  ],
  "key": [
    "73129e8a41a791aa2cbd6a76481adcc251af17f81f9426377906dd50d93eb2ca",
    "d7a63478a91895949221a1848a5afd083e2e5d7b5cf914c50da7e1b403608989"
  ]
}
```



API



Processar un tiquet

Processar una llista de tiquets

FastAPI 0.1.0 OAS 3.1

POST `/processor_tiquet/` Process Tiquet Endpoint

POST `/processor_excel/` Process Excel Endpoint

Parameters

Name

Description

column_number ★ required

integer

(query)

column_number

Request body required

file ★ required

string(\$binary)

Responses

Code

Description

200

Successful Response

422

Validation Error



04

Entrenament models

Dades sintètiques

Entitats que depenen del context

Entitats llargues

Generació de text

Reconeixement d'Entitats a Sentències Judicials de l'Índia

The **Supreme Court of India** **COURT**
Criminal Appeal Jurisdiction
[Arising out of Special Leave Petition (Crl.) No. 7999/2010
State of Kerala **PETITIONER** ... Appellant
-versus-
Raneef **RESPONDENT** ... Respondent
Judgement
Markandey Katju **JUDGE**

1. Leave granted
2. Heard Learned counsel for the parties
3. The appellant has filled this appeal challenging the impugned order of the **Kerala High Court** **COURT** dated **17.09.2010** **DATE** granting bail to the respondent Dr. **Raneef** **OTHER_PERSON**, who is a medical practitioner (dentist) in **Ernakular** **GPE** district in **Kerala** **GPE**, and is accused in crime no. 704 of 2010 of **P.S. Muvattupuzha** **ORG** for offences under various provisions of the **I.P.C. Statute**, the **Explosive Substances Act** **Statute** and the **Unlawful Activities (Prevention) Act** **Statute**.

Preamble

Judgement Text

Models destacats



LocalGPT

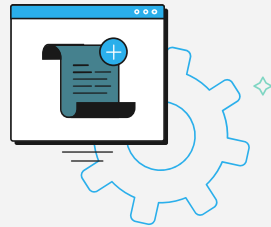


Mistral 7B



Llama 2

**No hi ha suficient
memòria**



Família Flan-T5

**Flan-T5-base
69,36%**

Dades reals

- Usuaris afectats
- Accions de mitigació
- Accions de control
- Mail de la víctima
- Mail de l'atacant
- Assumpte del correu
- URL de l'incident

Extracció d'informació en tiquets de phishing

Mètrica	Total
NotFound <i>test</i>	72.3%
NotFound <i>train</i>	67.6%
Total	68.3%

Models destacats



Família Flan-T5



Mistral 7B



Llama 2



FLOR 6.3B



Qwen 1.5

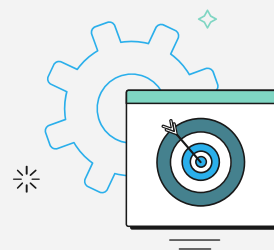




05

Avaluació

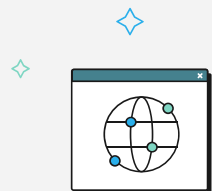
Percentatge d'encerts



FLOR 6.3B

QWEN1.5 7B

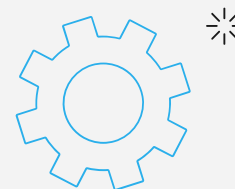
<u>Usuaris afectats</u>	0%	53,3%
<u>Accions de mitigació</u>	0%	80,0%
<u>Accions de control</u>	0%	93,3%
<u>Mail de la víctima</u>	0%	86,6%
<u>Mail de l'atacant</u>	0%	53,3%
<u>Assumpte del correu</u>	0%	60,0%
<u>URL de l'incident</u>	0%	73,3%
<u>Total</u>	0%	71,4%



Percentatge de “NotFound”

Conjunt
d'entrenament QWEN1.5 7B

<u>Usuaris afectats</u>	50,0%	80,0%
<u>Accions de mitigació</u>	63,7%	93,3%
<u>Accions de control</u>	97,5%	100,0%
<u>Mail de la víctima</u>	52,5%	80,0%
<u>Mail de l'atacant</u>	56,2%	86,6%
<u>Assumpte del correu</u>	61,2%	86,6%
<u>URL de l'incident</u>	92,5%	86,6%
<u>Total</u>	67,6%	89,5%





06

Conclusions

Desplegament temporal



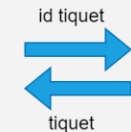
Portàtil amb GPU

1. Extracció de tiquets i preprocessament

2. Extracció d'informació del tiquet

3. Anonimització i inserció en base de dades

Màquina principal



OTRS

.49

Retorna el tiquet sol·licitat

Portàtil amb GPU



pipeline

.48

Rep el JSON via SCP i la comanda netcat mitjançant SSH.

netcat amb JSON

Port 6000



ElasticSearch

.47

Logstash anonimitza el tiquet i el guarda a l'Elasticsearch en la mateixa màquina

Anàlisi automàtica de tiquets de phishing

mitjançant el processament del
llenguatge natural

Jaume Casals Vilaplana

