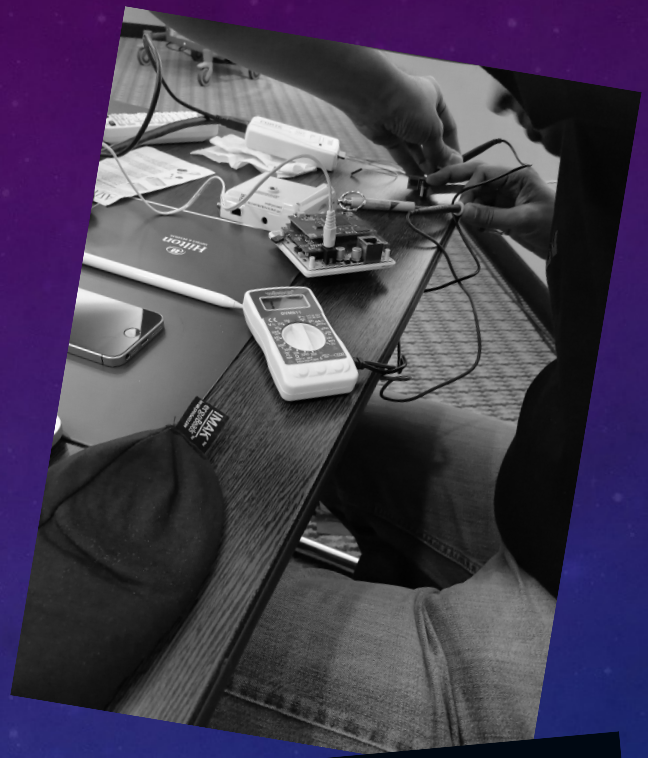
NATURAL LANGUAGE PROCESSING WORKSHOP

HACKIVITY 2023

PAULINE BOURMEAU

WHO AM I

- Linguist
- Hardware hacking
- Threat Intel, criminology
- Opensource
- And climber ^^.



Pauline bourmeau

CODE ENVIRONMENT

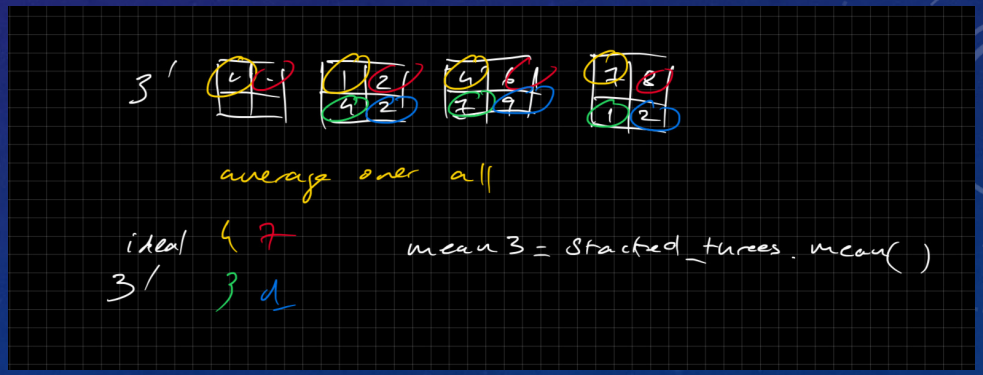
- Source: github
  - https://github.com/C00kie-/nlp-workshops
- Lab: Google Colab or Kaggle
  - Colab, upload the notebook1
  - Kaggle, create and account and go to this notebook:
    - (https://www.kaggle.com/code/paulinepeps/workshop-notebook-1)

MACHINE LEARNING

Stable diffusion

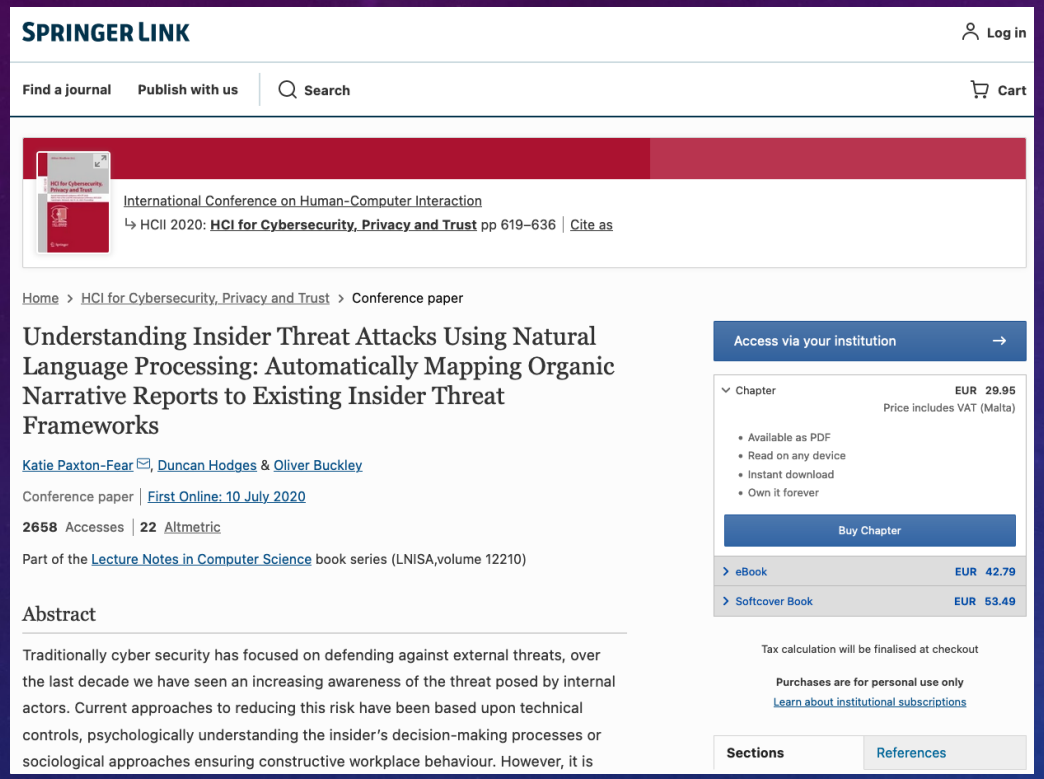
“What would a squirrel in a Iron Man suit looks like?”

“And a robot-squirrel?”



MNIST Dataset for image recognition

WHAT FOR?



SPRINGER LINK

Log in

Find a journal | Publish with us | Search | Cart

International Conference on Human-Computer Interaction

↳ HCII 2020: **HCI for Cybersecurity, Privacy and Trust** pp 619–636 | Cite as

Home › HCI for Cybersecurity, Privacy and Trust › Conference paper

Understanding Insider Threat Attacks Using Natural Language Processing: Automatically Mapping Organic Narrative Reports to Existing Insider Threat Frameworks

Katie Paxton-Fear, Duncan Hodges & Oliver Buckley

Conference paper | First Online: 10 July 2020

2658 Accesses | **22** Altmetric

Part of the Lecture Notes in Computer Science book series (LNISA,volume 12210)

Abstract

Traditionally cyber security has focused on defending against external threats, over the last decade we have seen an increasing awareness of the threat posed by internal actors. Current approaches to reducing this risk have been based upon technical controls, psychologically understanding the insider's decision-making processes or sociological approaches ensuring constructive workplace behaviour. However, it is

Access via your institution →

Chapter EUR 29.95

Price includes VAT (Malta)

- Available as PDF
- Read on any device
- Instant download
- Own it forever

Buy Chapter

eBook EUR 42.79

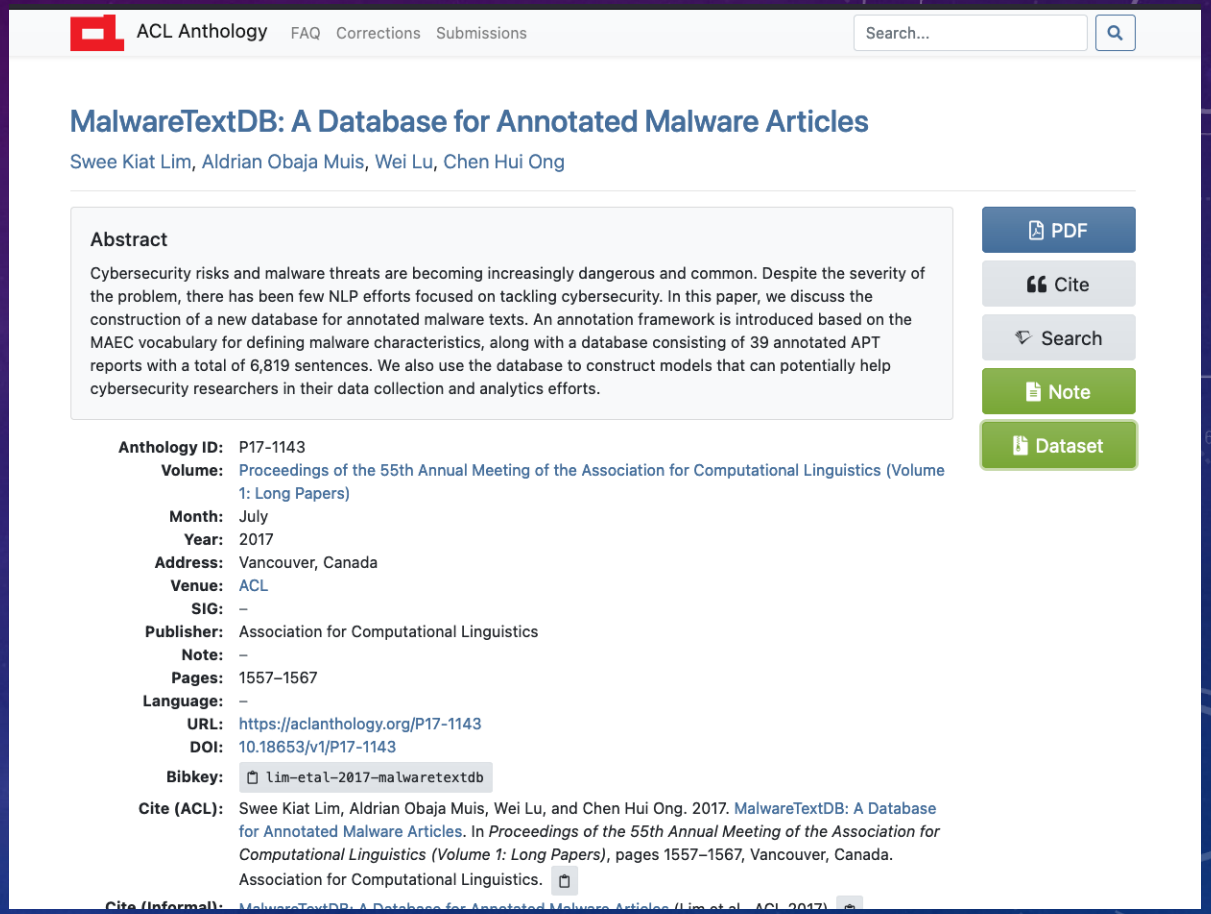
Softcover Book EUR 53.49

Tax calculation will be finalised at checkout

Purchases are for personal use only

Learn about institutional subscriptions

Sections | References



ACL Anthology FAQ Corrections Submissions Search...

MalwareTextDB: A Database for Annotated Malware Articles

Swee Kiat Lim, Aldrian Obaja Muis, Wei Lu, Chen Hui Ong

Abstract

Cybersecurity risks and malware threats are becoming increasingly dangerous and common. Despite the severity of the problem, there has been few NLP efforts focused on tackling cybersecurity. In this paper, we discuss the construction of a new database for annotated malware texts. An annotation framework is introduced based on the MAEC vocabulary for defining malware characteristics, along with a database consisting of 39 annotated APT reports with a total of 6,819 sentences. We also use the database to construct models that can potentially help cybersecurity researchers in their data collection and analytics efforts.

PDF | Cite | Search | Note | Dataset

Anthology ID: P17-1143
Volume: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)
Month: July
Year: 2017
Address: Vancouver, Canada
Venue: ACL
SIG: –
Publisher: Association for Computational Linguistics
Note: –
Pages: 1557–1567
Language: –
URL: https://aclanthology.org/P17-1143
DOI: 10.18653/v1/P17-1143
Bibkey: lim-etal-2017-malwaretextdb
Cite (ACL): Swee Kiat Lim, Aldrian Obaja Muis, Wei Lu, and Chen Hui Ong. 2017. MalwareTextDB: A Database for Annotated Malware Articles. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1557–1567, Vancouver, Canada. Association for Computational Linguistics.

WHAT'S NEXT

- Pull request on Github your notebooks!
- Updated regularly.
- Participate to Kaggle Competitions =)

Thanks you!