

# DarkBERT learns language of the Dark Web

May 23 2023, by Peter Grad

---



Credit: CC0 Public Domain

Two types of people thrive in the distant hidden recesses of the Dark Web. One type consists of the good guys: whistleblowers, freedom fighters, journalists, the intelligence community and law enforcement agencies, all generally waging the good fight against power, greed and tyranny.

The other type is made up of the bad guys: criminals, drug gangs, extortionists, weapons dealers, terrorists.

The Dark Web is an active mall where criminals offer a laundry list of

criminal digital services providing passwords to [bank accounts](#), Social Security numbers and other private data for [identity theft](#), malware and cyberattack packages that can bring down a company, a town or a country.

"There's a compounding and unraveling chaos that is perpetually in motion in the Dark Web's toxic underbelly," James Scott, a senior fellow at the Institute for Critical Infrastructure Technology, once said.

Researchers at a national research university in South Korea are trying to shine a little more light on that toxic underbelly. Their report, "DarkBERT: A Language Model for the Dark Side of the Internet," appeared this week on the *arXiv* preprint server.

While the Dark Web comprises barely 5% of the entire internet, it draws roughly 3 million users daily. Cybersecurity Ventures predicts proceeds from global cybercrime will top \$10 trillion by 2025.

To help combat that menace, researchers at the Korea Advanced Institute of Science & Technology have pre-trained a large language model on documents obtained from the Dark Web. They said such an effort was needed to bring greater efficiency to efforts to navigate the Dark Web and aid those seeking to stem criminal activity.

Researcher Youngjin Jin said his team's language model, named DarkBERT, will "combat the extreme lexical and structural diversity of the Dark Web that may be detrimental to building a proper representation of the domain."

Jin said pre-trained language models, such as the earlier BERT and RoBERTa projects based on Surface Web content (as opposed to Dark Web content), "are not ideal for ... extracting useful information, due to the differences in the language used in the two domains."

"Our evaluation results show that the DarkBERT-based classification model outperforms that of known pre-trained language models," Jin said.

The researchers noted three key areas in which DarkBERT proved effective: ransomware leak detection, noteworthy thread detection in which potentially malicious threads were spotted, and threat keyword inference defined as "a set of keywords that are semantically related to threats and drug sales in the Dark Web."

Jin noted that manual review of the voluminous quantities of posts on the Dark Web would require "massive human resources." Automating such analysis would "significantly reduce the workload of security experts," especially with a language model trained in the unique vocabulary of the Dark Web, Jin said.

Law enforcement has made some progress in crushing illegal activity on the Dark Web. The first modern Dark Web marketplace, Silk Road, which made more than a billion dollars in illegal drug sales, was shut down by the FBI and its creator sentenced to life in prison. AlphaBay, which sold hundreds of millions of dollars worth of drugs and hacked data, was shut down by a multinational police effort.

But those efforts were a drop in the bucket. To achieve greater success, [law enforcement](#) must better learn the language of the cybercriminals.

DarkBERT appears to be a good step in that direction.

**More information:** Youngjin Jin et al, DarkBERT: A Language Model for the Dark Side of the Internet, *arXiv* (2023). [DOI: 10.48550/arxiv.2305.08596](https://doi.org/10.48550/arxiv.2305.08596)

Citation: DarkBERT learns language of the Dark Web (2023, May 23) retrieved 29 April 2024 from <https://techxplore.com/news/2023-05-darkbert-language-dark-web.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.