



# NEW PERSPECTIVES ON MEASURING CYBERSECURITY

**Molly Lesher**

Head, Digital Policy, Economics, and Measurement Unit  
OECD Directorate for Science, Technology and  
Innovation

5<sup>th</sup> International Seminar on Big Data for Official  
Statistics, 29-31 May 2024, Xiamen, China



# Measuring cybersecurity: An essential yet challenging task

- Cybersecurity **threats seem to be growing** in scale, scope, and impact.
- Evidence-based policies rely on robust data, but **cybersecurity is generally poorly measured**.
- **Data exists** from official and unofficial sources, but it **often biased and not cross-country comparable**.
- **Actors** in the cybersecurity ecosystem **often lack the skills and/or incentives** to provide key data.





## Main elements of the report

---

- Key cybersecurity **measurement considerations**.
- Cybersecurity data from **official sources**.
- Cybersecurity data from **non-official sources**.
- Two **innovative approaches** to measuring **uncertainty** in cyberspace: **news reports & online searches**.



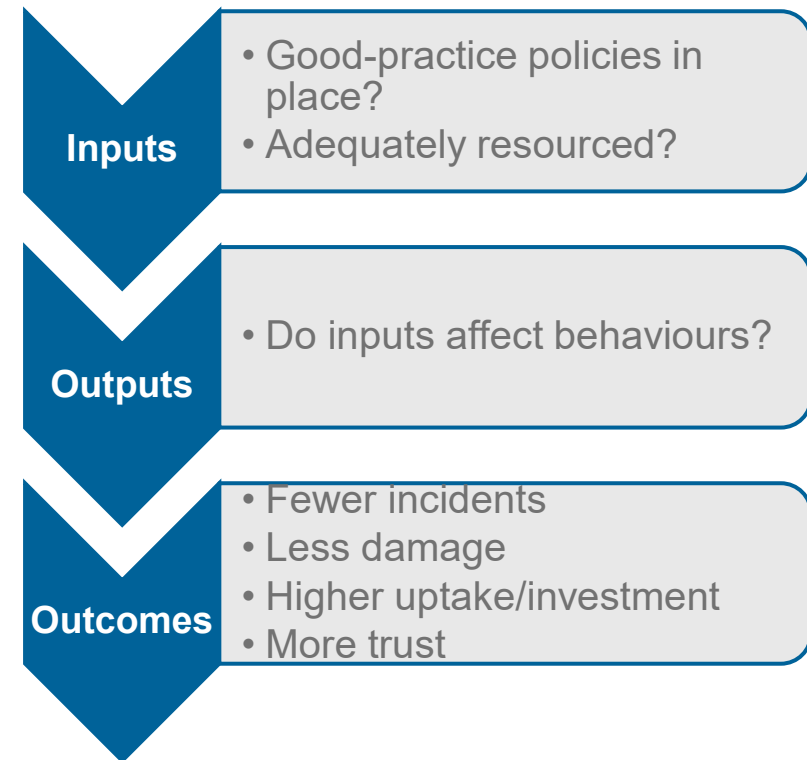
# CHECKLIST OF CYBERSECURITY MEASUREMENT CONSIDERATIONS



# Key cybersecurity measurement considerations

1. **Define** what to measure,
2. Explore different **data sources**,
3. **Assess** frequency, timeliness, and comparability,
4. Consider **interpretability, transparency,** and bias.

Figure 1. Assessing policy impact via inputs, outputs, and outcomes



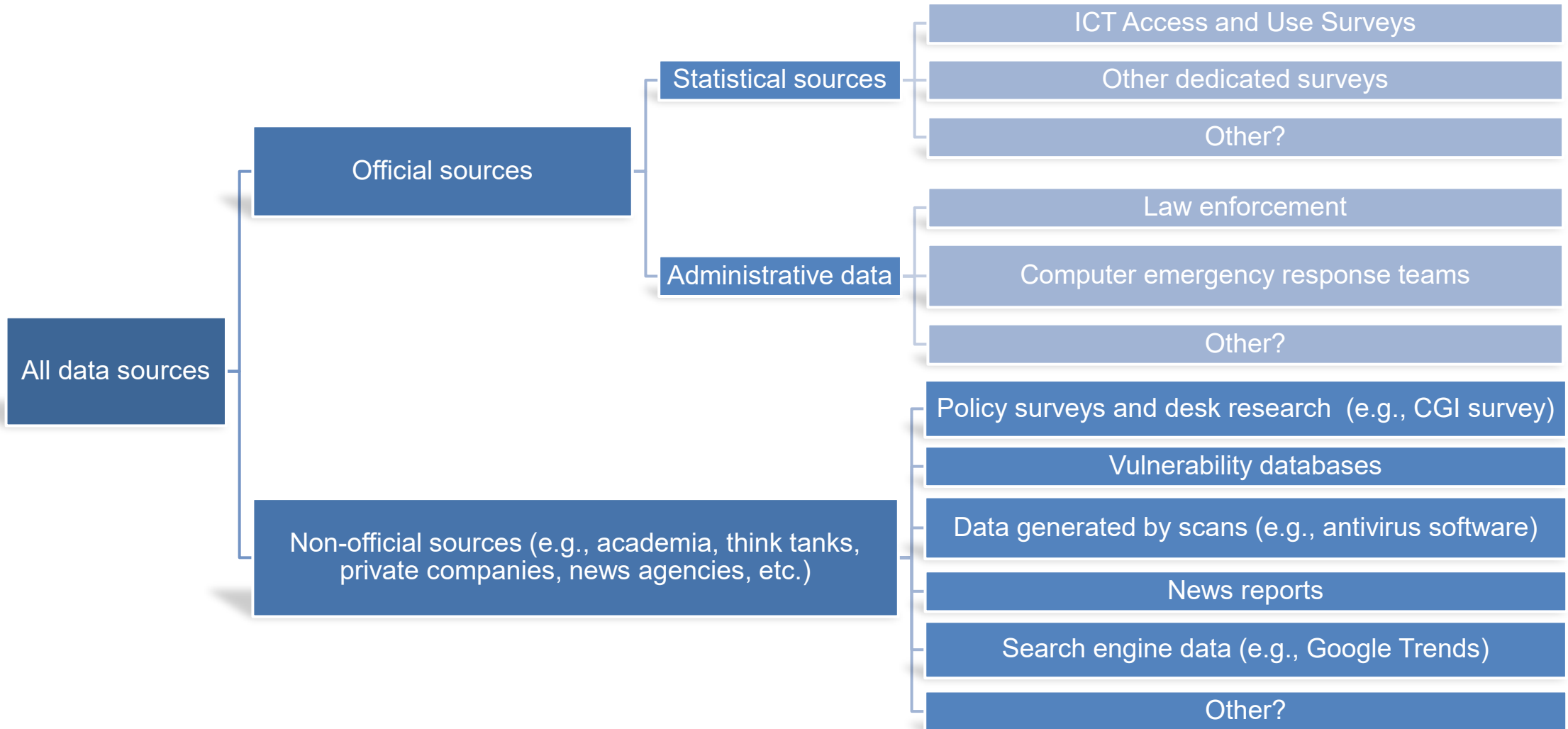
Source: Authors' elaboration



# AN OVERVIEW OF DIFFERENT CYBERSECURITY DATA SOURCES



# Mapping cybersecurity data sources





# Official data sources: Spotlight on sample surveys

*Indicators from sample surveys on outcomes should be used only very carefully, while data on security measures are likely informative*

**Annex Table 1. Current availability of relevant indicators from the OECD's ICT Access and Use databases**

Enterprises (10 employees and more, business sector excl. financial services) and individuals (aged 16-74), 2015-2021

Indicator	Indicator caption	2015	2016	2017	2018	2019	2020	2021
<b>Enterprises (40 countries, 38 OECD member countries, Croatia, and Romania)</b>								
E3	Businesses experiencing ICT Incidents (security breaches) (%)	5.0%	12.5%	7.5%	<b>80.0%</b>	10.0%	12.5%	<b>70.0%</b>
E7	Businesses with formal policy to manage ICT privacy risks (%)	5.0%	5.0%	7.5%	5.0%	10.0%	7.5%	2.5%
<b>Individuals (42 countries, 38 OECD member countries, Brazil, Bulgaria, Croatia, and Romania)</b>								
F4G	Individuals did not submit forms to public authorities: personal data protection and security concerns - last 12 m (%)	64.3%	64.3%	<b>69.0%</b>	<b>66.7%</b>	<b>66.7%</b>	<b>66.7%</b>	<b>66.7%</b>
H1J	Individuals having caught a virus or other computer infection with impacts - last 3 m (%)	2.4%	4.8%	7.1%	2.4%	0.0%	0.0%	0.0%
I3	Individuals having experienced a financial loss from fraudulent payment - last 3 m (%)	<b>71.4%</b>	7.1%	9.5%	9.5%	<b>73.8%</b>	9.5%	4.8%





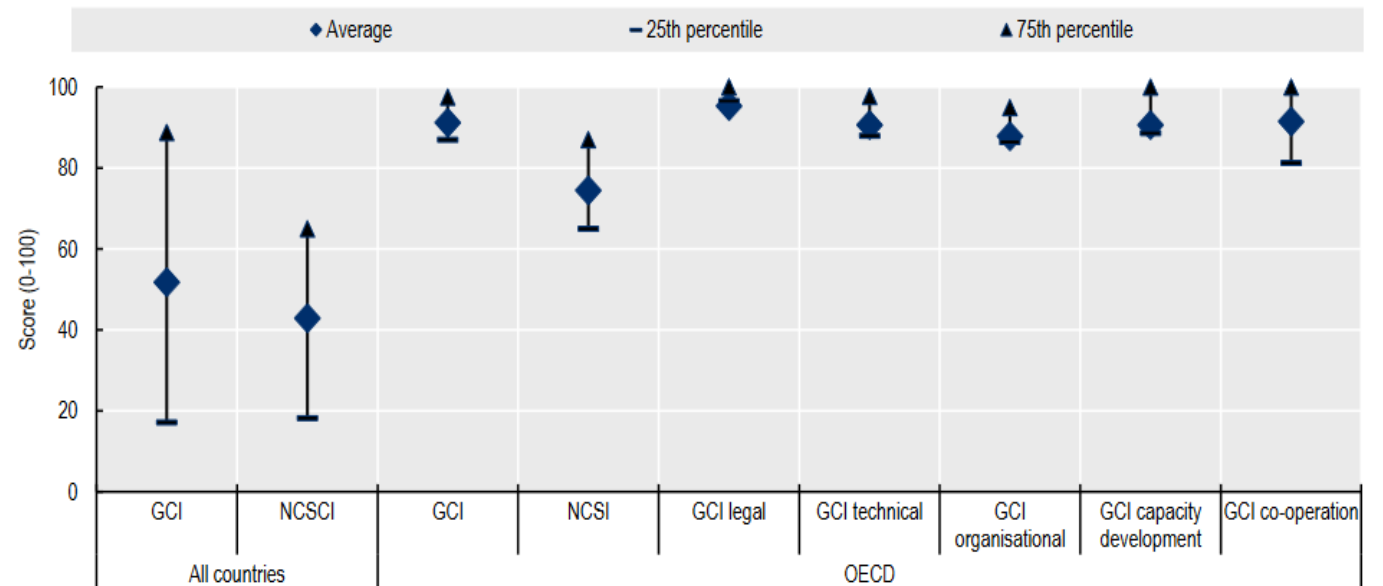
# Non-official data sources: Spotlight on policy surveys

Policy surveys are administered to relevant government authorities that aim to collect information about the institutional framework and potentially governments' actions:

- The United Nations Institute for Disarmament Research's *Cyber Policy Portal*
- The International Telecommunications Union's (ITU) *Global Cybersecurity Index (CGI)*
- The Potomac Institute's *Cyber Readiness Index 2.0*
- The e-Governance Academy's (eGA) *National Cybersecurity Index (NCSI)*

Figure 6. There is limited variation in the ITU's *Global Cybersecurity Index* across OECD countries

*Global Cybersecurity Index, 2020, and National Cybersecurity Index*





# TWO INNOVATIVE APPROACHES TO MEASURING CYBERSECURITY UNCERTAINTY



# A cyber uncertainty index based on news reports

- News reports about major incidents reflect the **state of cybersecurity**, and shape **public opinion** → might alter **online behaviour**.
- **Economic Policy Index** (Baker, Bloom and Davis, 2016)\*.
- Index construction:
  1. Identify one **leading newspaper** by country,
  2. Define **search criteria** (see Table 4),
  3. Compute the **relative frequency of reports talking about cybersecurity**,
  4. **Normalise** the index.
- Index values = **proportion of cybersecurity-related articles** in each newspaper → **relative prominence of cybersecurity in the media landscape**.
- **High frequency** of articles on cyber-incidents → potentially **high uncertainty** → **potentially affect** investment and uptake of online services.

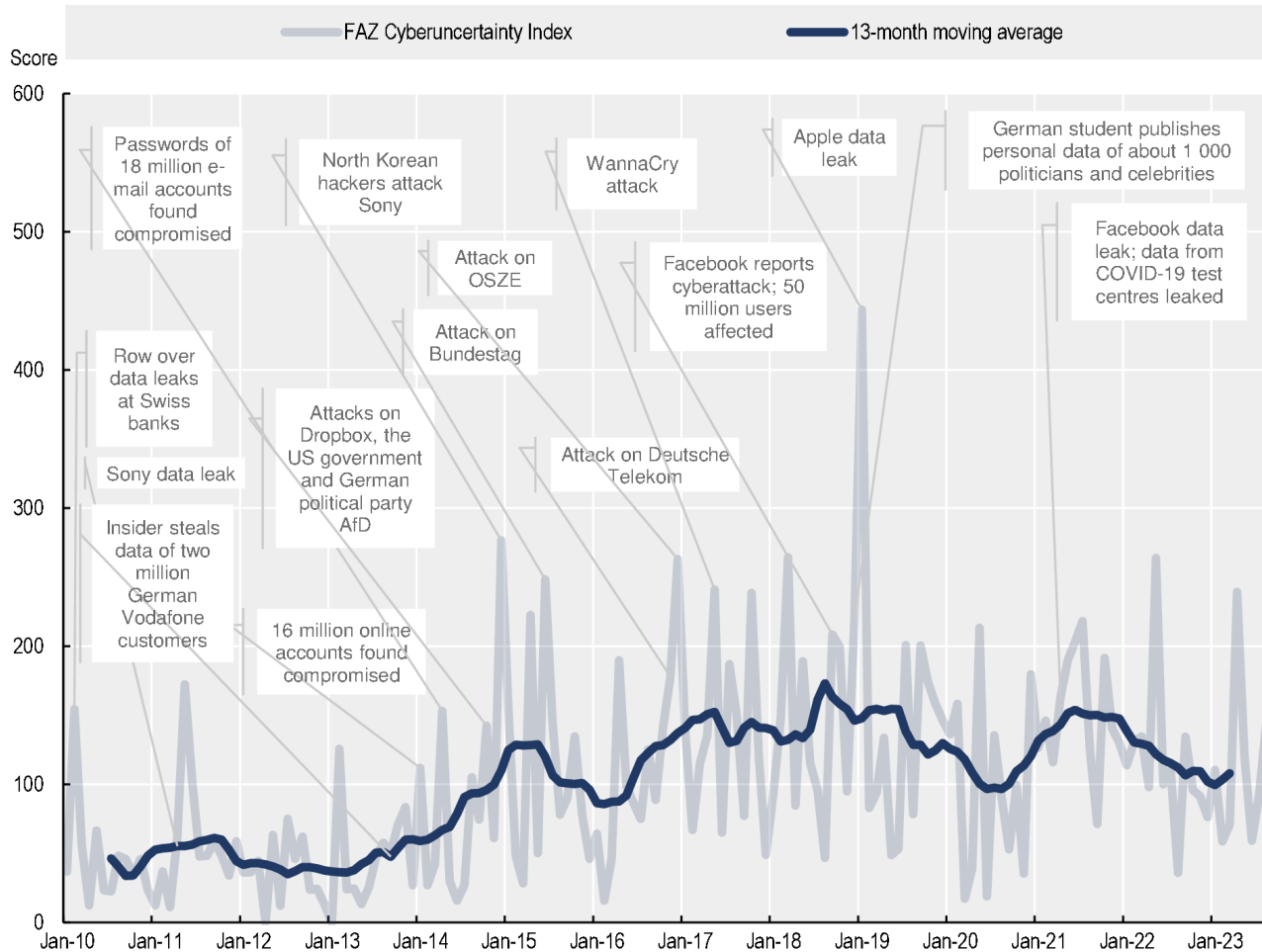
Table 4. Alternative search criteria for *Le Figaro*, the *FAZ* and *The New York Times*

	Search terms	English translation
Le Figaro	"cybersécurité"	"cybersecurity"
Frankfurter Allgemeine Zeitung I	"IT-Sicherheit" OR "Cybersicherheit"	"IT security" OR "cybersecurity"
Frankfurter Allgemeine Zeitung II	"IT-Sicherheit" OR "Cybersicherheit" OR "Datendiebstahl" OR "Datenleck" OR "Hackerangriff" OR "Datenpanne,"	"IT security" OR "cybersecurity" OR "data theft" OR "data breach" OR "hacker attack" OR "data breach"
Frankfurter Allgemeine Zeitung III	((("IT-Sicherheit" OR "Cybersicherheit") AND ("Angriff" OR "Vorfall" OR "Panne")) OR ("Datendiebstahl" OR "Datenleck" OR "Hackerangriff" OR "Datenpanne"))	((("IT security" OR "cybersecurity") AND ("attack" OR "incident" OR "breach")) OR ("data theft" OR "data breach" OR "hacker attack" OR "data breach"))
The New York Times	"cybersecurity"	

\* Baker, S. R., Bloom, N., & Davis, S. J. (2016), "Measuring economic policy uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593-1636.



# The FAZ Index for Germany (2010-23)

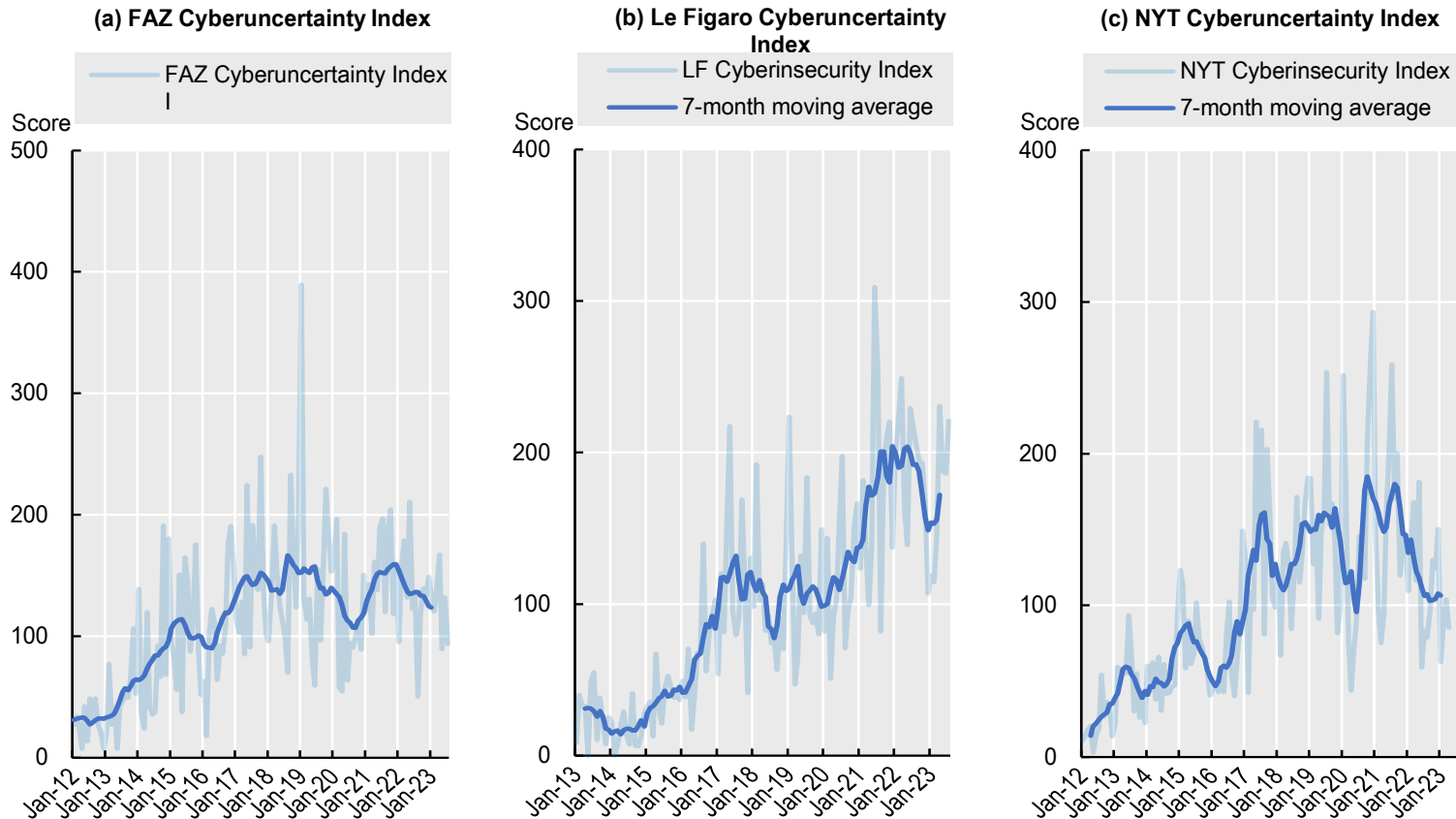


Note: The dark blue line indicates a 13-month moving average centered on the date for which it is reported. Source: Authors' elaboration based on data from the *Frankfurter Allgemeine Zeitung*.

- Leading paper for Germany: the *Frankfurter Allgemeine Zeitung* (FAZ).
- Search **criteria n°III** (see Table 4); articles selected contain:
  - Either “IT security” or “cybersecurity”, and,
  - At least one term indicating a threat/cyber incident.
- **Spikes** reflect major incidents; both with primarily **national** (e.g., German parliament hack, 06/2015) and **global repercussions** (e.g., WannaCry ransomware attack, 05/2017).
- Observed **secular trend**.



# Comparing newspaper indices for France, Germany and the United States (2012-23)



Note: The dark blue line indicates a 7-month moving average centered on the date for which it is reported.

Source: Authors' elaboration based on data from the *Frankfurter Allgemeine Zeitung*, *Le Figaro* and *The New York Times*.

- **Similar trends** observed using search **criterion n°1** (e.g., “cybersecurity” only).
- Selected leading newspapers:
  - *FAZ* for Germany,
  - *Le Figaro* for France, and,
  - *The New York Times* (*NYT*) for the United States.
- **Strong correlations across the three indices:**
  - $\text{corr}(\text{Le Figaro-FAZ}) = 0.46$ ,
  - $\text{corr}(\text{Le Figaro-NYT}) = 0.56$ ,
  - $\text{corr}(\text{FAZ-NYT}) = 0.46$ .



# The Google Trends Cyber Uncertainty (GTCU) Index

- 3 motivations for using online searches: **relevance, timeliness & coverage.**
- **Google Trends Uncertainty Index** (Castelnuovo and Tran, 2017)\*.
- Index construction:
  1. Select 9 **topics related to cyber risks** (see Table 6),
  2. Elect a **benchmark** topic and **combine the data** from multiple searches to get **relative frequencies**,
  3. **Sum up topic frequencies**,
  4. **Normalize** the index ( $\mu = 100$ ;  $\sigma = 30$ ).
- Index values  $\rightarrow$  **varying search interest** or popularity regarding topics linked to **cyber threats**, aligning with different **levels of uncertainty** related to cyber **incidents**.

Figure 11. The relationship between Google Trends keywords, topics, and categories

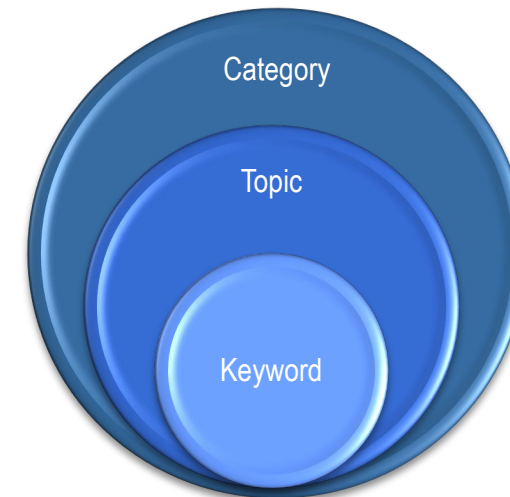


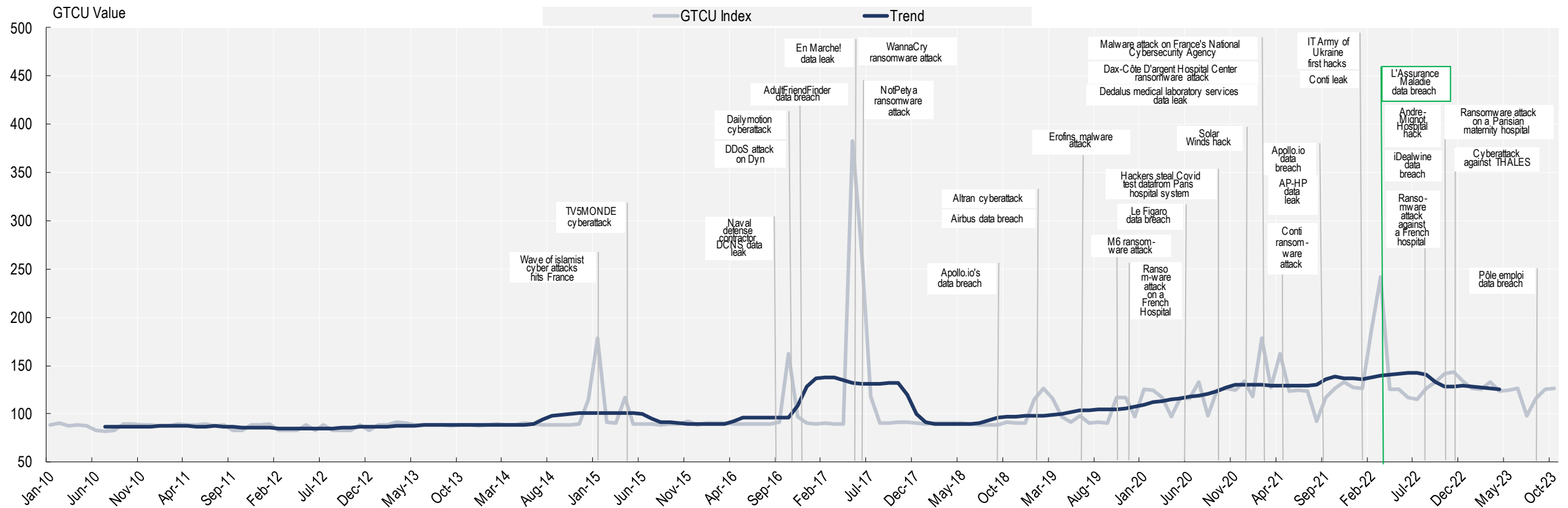
Table 6. Matching nine topic labels with corresponding IDs in Google Trends

Google Trends topic label	Google Trends topic ID
Cyberattack	/m/0p78w_d
Data breach	/m/03c18t5
Spyware	/m/075fh
Malware	/m/0582c
Ransomware	/m/0657nv
Computer virus	/m/01qgr
Phishing	/m/027b9k
Identity theft	/m/018npy
Cybercrime	/m/01y4q_

\* Castelnuovo, E., & Tran, T. D. (2017), "Google it up! a Google Trends-based uncertainty index for the United States and Australia," *Economics Letters*, 161, 149-153.



# The GTCU Index for France (2010-2023)

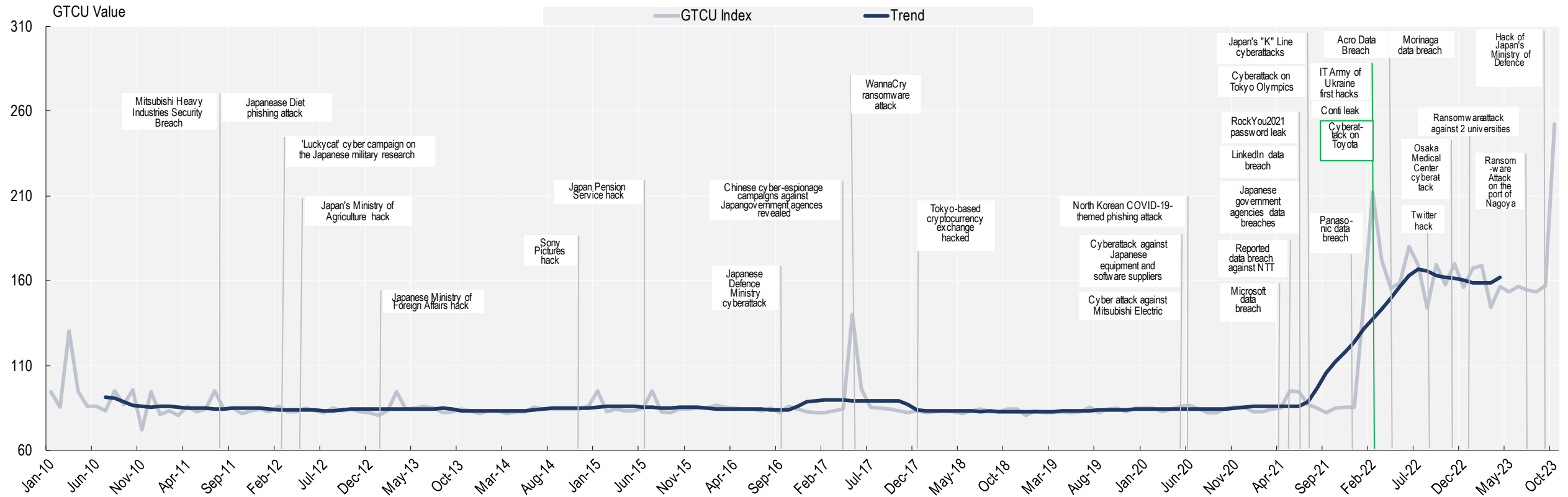


Note: The dark blue line illustrates the trend component obtained through seasonal decomposition, capturing the underlying long-term behavior or trend in the dataset.

Source: Authors' calculations using Google Trends data.



# The GTCU Index for Japan (2010-2023)



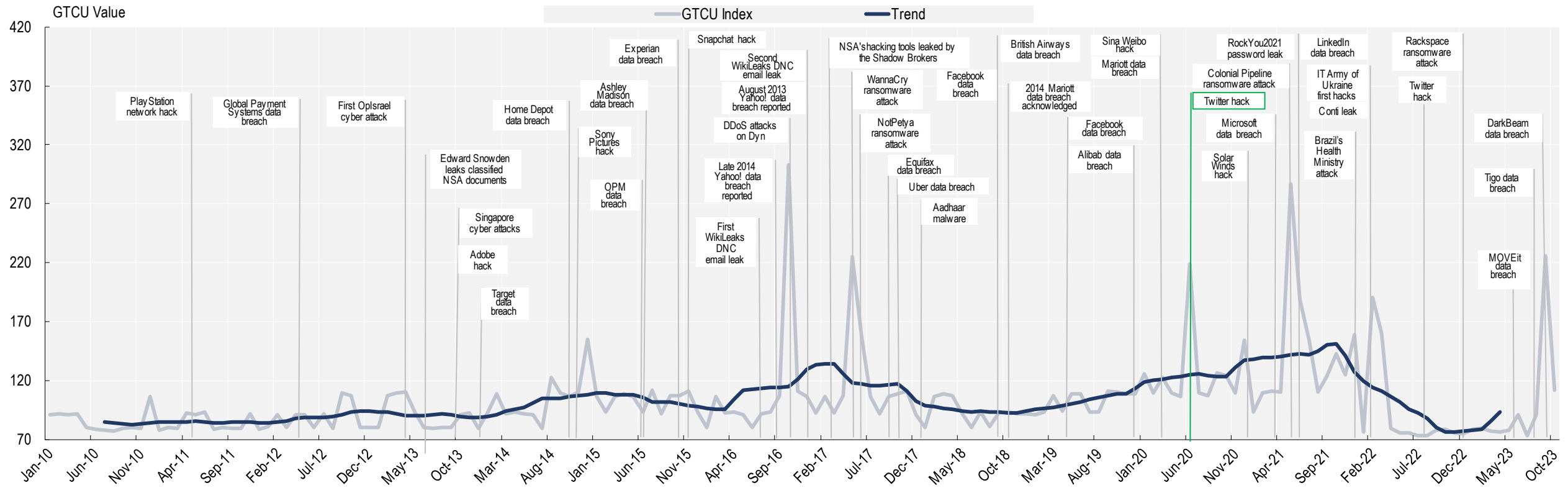
Note: The dark blue line illustrates the trend component obtained through seasonal decomposition, capturing the underlying long-term behavior or trend in the dataset.

Source: Authors' calculations using Google Trends data.





# The GTCU Index for the United States (2010-2023)



Note: The dark blue line illustrates the trend component obtained through seasonal decomposition, capturing the underlying long-term behavior or trend in the dataset.

Source: Authors' calculations using Google Trends data.



# GTCU index series: Main insights, advanced features and future steps

---

- Index series are informative:
  - Several peaks match major cyber incidents (with primarily national & global repercussions),
  - Suggest that cyber incidents aimed at governments generate greater levels of uncertainty,
  - Public perceptions differ across countries,
- Beyond index series: exploring two advanced features of the index:
  - Top related topics → inform cybersecurity policies (e.g., awareness campaigns),
  - Rising related topics → detect unusual cybersecurity activity online real-time.
- Future steps would involve:
  - Refining the methodology → address biases (e.g., sampling variance),
  - Expanding to more countries
  - Assessing its predictive capacity (e.g., ICT investment or cybersecurity job postings).



# News reports and Google Trends cyber-uncertainty indices: Strengths and weaknesses

	News Reports	Google Trends
Strengths	<ul style="list-style-type: none"><li>• <b>Well-established methodology</b> (U.S. Federal Reserve Bank).</li><li>• <b>Applicability</b> to several OECD economies.</li><li>• <b>Monthly</b> frequency.</li><li>• <b>Transparent results</b> that are <b>easy to explain</b>.</li></ul>	<ul style="list-style-type: none"><li>• <b>Very high frequency</b> and <b>comprehensive data</b> (e.g., weekly/daily, sub nationally).</li><li>• Easily applicable to all OECD countries.</li><li>• <b>Cross country comparability</b>.</li><li>• Inform <b>cybersecurity policies</b>.</li><li>• Detect <b>emerging security threats/concerns</b>.</li></ul>
Weaknesses	<ul style="list-style-type: none"><li>• Limited to <b>reported incidents</b> by journalists.</li><li>• <b>Limited benchmarking value</b> → different <b>reporting styles/priorities</b>.</li><li>• Scalability implies relying on <b>black boxes</b>.</li><li>• <b>No single-date cross-country comparisons</b>.</li></ul>	<ul style="list-style-type: none"><li>• <b>Limited transparency</b> (user confidentiality, prevent data manipulation).</li></ul>



Next steps

---

OECD Digital Economy Paper  
to be published in **June 2024**

<https://doi.org/10.1787/20716826>



# Thank you for your attention

For any questions, please contact:

[molly.lesher@oecd.org](mailto:molly.lesher@oecd.org)