Poster: The Unknown Unknown: Cybersecurity Threats of Shadow IT in Higher Education

Jan-Philip van Acken
Utrecht University
Utrecht, Netherlands
j.vanacken@uu.nl

Joost F. Gadellaa
Coöperatie SURF U.A.
Utrecht, Netherlands
joost.gadellaa@surf.nl

Slinger Jansen
Utrecht University
Utrecht, Netherlands
slinger.jansen@uu.nl

Katsiaryna Labunets
Utrecht University
Utrecht, Netherlands
k.labunets@uu.nl

ABSTRACT

The growing number of employee-introduced IT solutions creates new attack vectors and challenges for cybersecurity management and IT administrators. These unauthorised hardware, software, or services are called *shadow IT*. In higher education, the diversity of the shadow IT landscape is even more prominent due to the flexible needs of researchers, educators, and students.

We studied shadow IT and related cyber threats in higher education via interviews with 11 IT and security experts. Our results provide a comprehensive overview of observed shadow IT types and related cyber threats. The findings revealed prevalent cloud and self-acquired software use as common shadow IT, with cybersecurity risks resulting from outdated software and visibility gaps. Our findings led to advice for practitioners: manage shadow IT responsibly with cybersecurity best practices, consider stakeholder needs, support educators and researchers, and offer usable IT solutions.

CCS CONCEPTS

• **Security and privacy** → **Usability in security and privacy**; **Social aspects of security and privacy**.

KEYWORDS

Shadow IT, Higher Education Institutes, Cyber Threats, Cyber Risk Management, Qualitative Study

ACM Reference Format:
Jan-Philip van Acken, Joost F. Gadellaa, Slinger Jansen, and Katsiaryna Labunets. 2023. Poster: The Unknown Unknown: Cybersecurity Threats of Shadow IT in Higher Education. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23), November 26–30, 2023, Copenhagen, Denmark.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3576915.3624395

1 INTRODUCTION

Higher Educational Institutions (HEIs) manage an ever-increasing demand for high-quality and extensive IT services that combine a regular "corporate" IT environment with collaboration software for students and researchers, innovative research equipment, and blended learning tools. Some researchers call HEIs *"the least secure place in the universe"* [2]. HEIs deal with large amounts of (personal) data, intellectual property, and computational power, making them attractive targets for malicious actors.

Another trend also in HEIs is growing shadow IT, i.e., *"hardware, software, or services built, introduced, and/or used for the job without explicit approval or even knowledge of the organization"* [6]. A recent survey[1], revealed that 46% of IT directors agreed that *"direct purchasing of SaaS and other non-sanctioned software by individuals and business units makes it impossible to protect all their data, systems and applications"* and 21% of participants reported shadow IT-related cyber incidents. Shadow IT can be deployed by departments, end-users (incl. students), or research groups to get work done with no malicious intent (e.g., research groups in Whatsapp, freemium tools for integrating digital whiteboards, in-lecture voting apps, virtual cloud machines for research, and other). Security risks of shadow IT have been highlighted in the literature [7]. On top of it, shadow IT introduces issues related to compliance, lack of integration with other IT, and loss of synergy. With these risks, it would be obvious to ban all shadow IT in HEIs, if it were not for the value that shadow IT brings. Identifying and managing all shadow IT instances would not be possible or pragmatic, and there is evidence that sanctioning is not a practical solution to shadow IT occurrence [6]. Moreover, shadow IT systems can be innovative and a helpful response to a lack of organizational agility.

Many institutions take a risk-based approach to shadow IT and cybersecurity, aiming to get a grip on the risks of shadow IT instead of banning the phenomenon. Thus, our study aims to gain empirical insights into shadow IT presence and the related cyber threats within higher education organizations as the first step towards comprehensive cybersecurity risk management on this topic.

This paper presents the preliminary results of the first qualitative study of cyber threats of shadow IT with 11 experts from Dutch HEIs. We extended the existing taxonomy of shadow IT by Mallmann et al. [11] with an inventory of shadow IT types in Dutch HEIs. We explore the cyber threat landscape of HEIs through the lens of CORAS threat modelling notation and provide recommendations on enabling the responsible use of shadow IT and improving cyber risk management in this light.

2 RESEARCH APPROACH

Following the study aim, we formulated our research questions:

RQ1: *What types of shadow IT are observed in Dutch HEIs?*
RQ2: *What cyber threats do experts associate with shadow IT?*
RQ3: *How cyber threats are connected to specific shadow IT types?*

CCS '23, November 26–30, 2023, Copenhagen, Denmark
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0050-7/23/11.
https://doi.org/10.1145/3576915.3624395

[1] Forbes Insights and IBM 2019 Survey Report *"Perception gaps in cyber resilience: Where are your blind spots? The hidden costs of shadow IT, cloud, and cyber insurance."* Report available at: https://edu.nl/eadeq. Slides: https://edu.nl/evxjj.

Variable	Scale	Mean/SD	Distribution
Institution type			63% Higher Professional Education (HBO); 36% Scientific Education (WO)
Role			36% CISO; 45% Security and/or Privacy Officer; 18% Other
Years at institution	Years	14.93 ±10.35	27% worked 5 years or less; 9% worked 6-10 years; 27% worked 11-20 years; 27% worked > 20 years; 9% did not report;
Work experience	Years	25.67 ±8.35	27% had 10-20 years; 27% had 21-30 years; 27% had > 30 years; 18% did not report
Security education			63% had it; 18% did not have; 18% did not report

Participants were directly asked if they have security and privacy-related educations or certificates (e.g., CISSP or DPO).

Table 1: Demographics of the participants

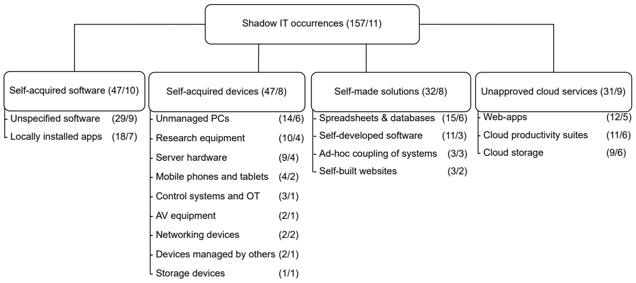
Shadow IT occurrences (157/11)
Self-acquired software (47/10)
Unspecified software (29/9)
Locally installed apps (18/7)
Self-acquired devices (47/8)
Unmanaged PCs (14/6)
Research equipment (10/4)
Server hardware (9/4)
Mobile phones and tablets (4/2)
Control systems and OT (3/1)
AV equipment (2/1)
Networking devices (2/2)
Devices managed by others (2/1)
Storage devices (1/1)
Self-made solutions (32/8)
Spreadsheets & databases (15/6)
Self-developed software (11/3)
Ad-hoc coupling of systems (3/3)
Self-built websites (3/2)
Unapproved cloud services (31/9)
Web-apps (12/5)
Cloud productivity suites (11/6)
Cloud storage (9/6)

Figure 1: Shadow IT types with groundedness and coverage

To address these questions, we conducted semi-structured interviews with 11 experts working at various Dutch HEIs. To recruit the participants, we adopted a mixed method approach combining maximum variation purposive sampling and convenience sampling [12]. We interviewed security experts from the policy side of cybersecurity practice (such as CISOs) and the more practical side (such as SOC analysts and CERT members) with almost 15 years on average working in the institution. Table 1 shows the demographics of our participants and the organisations they represent.

The semi-structured interviews were designed to answer RQs through the lens of the cyber threat-related concepts in the CORAS notation [10]. This way, we aimed to reveal emerging patterns among shadow IT types and associated cyber threats.

Ethics: The study was approved by Utrecht University's Science-Geo Ethics Review Board (ref. Beta S-22770). Participant consent was obtained after informing them of the study details, risks, and our use of collected information. We anonymised the interview transcripts and analysis project to protect our participants' privacy and mitigate possible risks for them and their organisations. The participants were asked to review the anonymised transcripts and provide final consent for their publication.

Data analysis: We employed a coding process inspired by grounded theory [3], a widely used approach [1, 8]. Two authors coded shadow IT types based on Mallmann et al.'s taxonomy [11]. The CORAS notation was used to facilitate discussion and analysis of cyber threat components. Through threat modelling, we linked the observed shadow IT types with perceived cyber threats and shed light on the threat landscape for HEIs and the impact of shadow IT.

Code Saturation: To assess thematic code saturation, we followed Guest et al.'s approach [5] by comparing new codes for each run consisting of three interviews to a base set of four interviews. After two runs (totalling ten interviews), we achieved saturation with 0% new codes for shadow IT types and 2.7% for threats, indicating a comprehensive understanding of emerging themes in the topic.

3 RESULTS

The eleven interviews resulted in a codebook of 63 codes: 18 for types of shadow IT and 45 for the different components of threat modelling, which were applied 448 times. We present the most prominent results for each RQ and the related quotes from our experts (E01-E11).

RQ1 - Types of Shadow IT: Figure 1 shows the identified types of shadow IT with the groundedness and coverage metrics per code. *Groundedness* is the frequency of the code across all interviews, while *coverage* shows how many interviewees mentioned this code at least once. All experts agreed to a common shadow IT definition beforehand, yet not all of them identified types from every category.

Self-acquired devices is the most varied category of shadow IT. The top three frequently mentioned examples are 'unmanaged PCs', 'research equipment', and 'server hardware'.

> *"We also have a [unit], and you also have all kinds of equipment with computers built in that is also not purchased centrally. Then an [research device] comes in, and there's a [legacy] device in there. "* [E07]

Self-acquired software is the next most frequent shadow IT type among our experts that mentioned 'unspecific software' and 'locally installed apps' as related types.

> *"[. . .] those teachers who think they should have the full version of [software] and the [department] thinks they shouldn't, so the money isn't there for it, and who then just [. . .] download a cracked version, because 'I have to have it for work'. [. . .] that too is not stopped."* [E01]

Self-made solutions include mainly 'spreadsheets and databases' duplicating data from official systems and 'self-developed software' created by researchers for their own use:

> *"There are the researchers who are developing something anyway, put it into production, and we don't actually know what they are doing. Besides, they haven't applied privacy and security by design either."* [E04]

Unapproved cloud services included three types of cloud-based services deemed shadow IT instances: 'web-apps', 'cloud storage', and 'cloud productivity suites'.

> *"We have our friends here from [a specific department] who [...] have everything running at [a shadow cloud provider]. They have their own mail domain on [provider] they have their own [cloud productivity tool]."* [E01]

RQ2 - Cyber Threats of Shadow IT: Due to space limit, we focus on the identified **vulnerabilities** (see the full list in Table 2). Our experts frequently mentioned 'outdated software' containing known security flaws:

> *"Well, the server was also not managed very professionally, so security patches were not installed. So yes, it had been an easy target to get in for a hacker [. . .] from the outside [. . .] to attack our network."* [E05]

Table 2: Shadow IT Types and Vulnerabilities Co-Occurrence

The items on axes are sorted by the number of times mentioned in total. The table provides insight into the density, the number of links from a code to other codes, of the different vulnerabilities and shadow IT types.

Shadow IT type/Vulnearbility		Outdated software	Lack of control of data	Lack of access control	Lack of authorization policy	No contract with the supplier	No visibility of vulnerabilities	Not actively managed	Re-used password	Remotely accessed	Not secure by design	Users have installation rights	Lack of data encryption	Firewall or antivirus lacking	Human error	Lack of backups	Lack of logging	Unsuitable hardware
Self-aquired devices	Unmanaged PCs	×	×	×	×		×	×	×			×	×	×	×			
	Research equipment	×				×		×	×					×			×	
	Server hardware	×		×	×					×	×			×				×
	Mobile phones and tablets	×	×	×									×					
	Control systems and OT	×		×				×										
	AV equipment	×						×		×								
	Devices managed by others	×																
	Networking devices	×																×
	Storage devices	×	×															
Self-aquired software	Unspecified software	×	×				×			×								
	Locally installed apps	×					×											
Unapproved cloud services	Web-apps		×	×		×	×		×							×	×	
	Cloud storage		×	×	×	×		×	×							×		
	Cloud productivity suite		×	×		×			×									
Self-made solutions	Spreadsheets and databases		×	×	×						×							
	Self-developed software	×					×				×							
	Self-built websites	×									×		×					
	Ad-hoc coupling of systems	×	×	×			×								×			

A 'lack of control of data' cannot be directly exploited by a threat, but it was used by our experts to describe what enabled further problems with shadow IT. A 'lack of access control' was mentioned when there is insufficient access control:

> *"Who would know that if you go to [application].domain.nl/employees.csv that you would simply get all the employees with all dates and addresses"* [E01]

RQ3 - Relations of Shadow IT Types and Cyber Threats: In our study, we see an occurrence of shadow IT as a tangible thing. Thus, in CORAS, we can relate it to a *vulnerability* that may be present in a specific shadow IT type and potentially be exploited by a threat actor to attack organisational assets. To explore the role of shadow IT in the cybersecurity landscape, we build a co-occurrence matrix for identified shadow IT types and vulnerabilities (see Table 2). We marked a co-occurrence when an expert mentioned a *vulnerability* strictly related to an occurrence of shadow IT.

'Outdated software' is related to all types of shadow IT, except for spreadsheets and the different cloud services. By now, this is clearly a top-of-mind concern for our experts across organisations and applicable to most of the IT landscape [4, 9].

Among vulnerabilities, *'re-used passwords'* stands out due to limited semantic overlap with other categories. This uniquely challenges certain shadow IT types: user-chosen passwords for cloud services and devices, and an application or device is not linked to the single sign-on solution of the institution. Mobile devices, however, use passcodes instead. If a weak passcode is used, it becomes a *lack of access control* issue, allowing unauthorised access.

We can also group vulnerabilities by specific shadow IT types. For example, for cloud services, *'re-used passwords'* and *'lack of contracts with the supplier'* are distinctively visible. *'Lack of contract'* for a cloud service is similar to *'not actively managed'* for self-acquired devices: nobody looks after the security and data of the system in your interest. Self-acquired devices have some specific issues that are not perceived for other shadow IT types, such as the *unsuitability of the hardware.*

Recommendations: We also identified a number of security measures mentioned by our experts. To minimise shadow IT, our expert suggested learning about the users' preferences, monitoring the network, and providing employees with managed devices. For devices that were initially unmanaged, it was suggested to take those under the managed umbrella as soon as they became known, *and* to provide managed alternatives: *"At one point we adopted the policy of: yes, we can also manage [network attached storage systems]. [. . .] If you want, we also manage your [(unmanaged) storage] [. . .]. We have deployed large-scale central storage [. . .] with redundancy, backup, ransomware protection, [etc]. "* [E07]

4 CONCLUSION

This paper presents the first qualitative study of cyber threats from shadow IT in HEIs, identifying prevalent types and associated vulnerabilities that can lead to cyber threats. One of our interviewees called shadow IT an *'unknown unknown'*. Our findings show that shadow IT is ingrained in the IT environments of HEIs. Interviews with 11 experts highlighted that almost every IT form has a shadow facet. A taxonomy of types is established, expanding on Mallman et al.'s framework and exploring links to vulnerabilities and cyber threats. Finally, we report recommendations for practitioners on enabling the responsible use of shadow IT and improving cyber risk management in this light.

REFERENCES

[1] Veroniek Binkhorst, Tobias Fiebig, Katharina Krombholz, Wolter Pieters, and Katsiaryna Labunets. 2022. Security at the End of the Tunnel: The Anatomy of VPN Mental Models Among Experts and Non-Experts in a Corporate Context. In *Proc. of USENIX Security* (31 ed.). USENIX Association, Boston, MA, 3433–3450.

[2] Ivano Bongiovanni. 2019. The Least Secure Places in the Universe? A Systematic Literature Review on Information Security Management in Higher Education. *COMPUT SECUR* 86 (2019), 350–357.

[3] Juliet Corbin and Anselm Strauss. 2008. *Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory.* Sage.

[4] Stephanie de Smale, Rik van Dijk, Xander Bouwman, Jeroen van der Ham, and Michel van Eeten. 2023. No One Drinks From the Firehose: How Organizations Filter and Prioritize Vulnerability Information. In *Proc. of S&P.* IEEE Computer Society, Los Alamitos, CA, USA, 203–219.

[5] Greg Guest, Emily Namey, and Mario Chen. 2020. A Simple Method to Assess and Report Thematic Saturation in Qualitative Research. *PLoS One* 15, 5 (2020).

[6] Steffi Haag and Andreas Eckhardt. 2017. Shadow IT. *BUS INF SYST ENG* 59, 6 (2017), 469–473.

[7] Stefan Klotz, Andreas Kopper, Markus Westner, and Susanne Strahringer. 2019. Causing Factors, Outcomes, and Governance of Shadow IT and Business-managed IT: A Systematic Literature Review. *International Journal of Information Systems and Project Management* 7, 3 (2019).

[8] Faris Bugra Kokulu, Ananta Soneji, Tiffany Bao, Yan Shoshitaishvili, Ziming Zhao, Adam Doupé, and Gail-Joon Ahn. 2019. Matched and mismatched SOCs: A qualitative study on security operations center issues. In *Proc. of ACM CCS.* 1955–1970.

[9] Platon Kotzias, Leyla Bilge, Pierre-Antoine Vervier, and Juan Caballero. 2019. Mind Your Own Business: A Longitudinal Study of Threats and Vulnerabilities in Enterprises.. In *Proc. of NDSS.*

[10] Mass Soldal Lund, Bjornar Solhaug, and Ketil Stølen. 2010. *Model-driven Risk Analysis* (2011 ed.). Springer, Berlin, Germany.

[11] Gabriela Labres Mallmann, Aline de Vargas Pinto, and Antônio Carlos Gastaud Maçada. 2019. Shedding Light on Shadow IT: Definition, Related Concepts, and Consequences. In *Information Systems for Industry 4.0.* Springer International Publishing, Cham, 63–79.

[12] Nel Verhoeven. 2019. *Doing Research: the Hows and Whys of Applied Research* (5 ed.). Boom, Amsterdam.