Editors: Melissa Dark, dark@purdue.edu | Jelena Mirkovic, mirkovic@isi.edu | Bill Newhouse, newhouse@nist.gov

Scoping the Cyber Security Body of Knowledge

Awais Rashid | University of Bristol
George Danezis | University College London
Howard Chivers | University of York
Emil Lupu | Imperial College London
Andrew Martin | University of Oxford
Makayla Lewis | Brunel University
Claudia Peersman | University of Bristol

Cybersecurity is becoming an important element in curricula at all education levels. However, the foundational knowledge on which the field of cybersecurity is being developed is fragmented, and as a result, it can be difficult for both students and educators to map coherent paths of progression through the subject. By comparison, mature scientific disciplines like mathematics, physics, chemistry, and biology have established foundational knowledge and clear learning pathways. Within software engineering, the IEEE Software Engineering Body of Knowledge (SWEBOK; www.computer.org/web/swebok) codifies key foundational knowledge on which a range of educational programs may be built. There are a number of previous and current efforts on establishing skills frameworks, key topic areas, and curricular guidelines for cybersecurity (see sidebar). However, a consensus has not been reached on what the diverse community of researchers, educators, and practitioners sees as established foundational knowledge in cybersecurity.

The Cyber Security Body of Knowledge (CyBOK) project (www.cybok.org) aims to codify the foundational and generally recognized knowledge on cybersecurity. In the same fashion as SWEBOK, CyBOK is meant to be a guide to the body of knowledge; the knowledge that it codifies already exists in literature such as textbooks, academic research articles, technical reports, white papers, and standards. Our focus is, therefore, on mapping established knowledge and not fully replicating everything that has ever been written on the subject. Educational programs ranging from secondary and undergraduate education to postgraduate and continuing professional development programs can then be developed on the basis of CyBOK.

Starting in 1 February 2017, we undertook a range of community consultations (see Tables 1 and 2), both within the UK and internationally, through a series of different activities designed to gain as much input as possible and from as wide an audience as possible. In addition, analysis of a number of relevant texts (44 in total), such as tables of contents of textbooks, calls for papers for conferences and symposia, standards, and existing certification programs, was undertaken to complement the insights gained from the community consultations. The insights from these activities were synthesized to develop a scope for CyBOK and 19 top-level knowledge areas (KAs).

Scoping Research

We describe the scoping research before discussing the KAs that emerged.

Consultation Workshops

One hundred and six attendees from UK industry and academia met—in a collaborative and creative environment—to discuss the KAs that ought to be included in CyBOK. Some workshops were dedicated to consultation with academia and others to consultation with practitioners. A subset also included representatives from both academia and practitioner communities.

The workshops were based on a supermarket metaphor (Figure 1) whereby participants were encouraged to think about what they considered to be the key KAs to be included in CyBOK. Participants discussed and identified a range of KAs collectively and put each KA into one of the four supermarket areas:

- In the trolley—KAs to be included;
- On the shopper's heart—KAs that are of interest to participants but not necessarily to be included;
- On the shelf—KAs to be discussed further; and
- In the bin—KAs deemed out of scope.

Copublished by the IEEE Computer and Reliability Societies

Table 1. Scoping research activities and number of participants/responses.

Input	Participants and no. of responses
Online survey	44 responses received
Analysis of relevant texts	44 separate texts analyzed
In-depth interviews with key experts	10 interviews undertaken
Community workshops across the UK	11 workshops
	106 attendees
Call for positions statements	13 statements received
Panel at Advances in Security Education Workshop at USENIX Security Symposium, Vancouver, Canada, October 2017	Paper-based exercise with 28 attendees

Table 2. Distribution of input from academia and practitioners.

Input	Academic (%)	Practitioner (%)
Online survey	51	49
In-depth interviews with key experts	50	50
Community workshops across the UK	55	45
Call for positions statements	62	38

This sorting exercise was followed by a 15-items-or-less task during which participants were asked to sort the "in the trolley" KAs into groups of top-level and sub-level KAs.

This workshop design allowed for small group discussions on where KAs would be best placed and why. It also led to subtopics within knowledge areas to be identified.

In addition to these workshops, consultations were also held at the Higher Education Academy Conference in Liverpool in April 2017 and the Cyber Security Professionals Conference in York, UK, in May 2017.

A panel discussion was also organized at the Advances in Security Education Workshop at the USENIX Security Symposium in Vancouver in August 2017, and views on the relative importance of particular topics emerging from the above workshops were sought via a paper-based exercise.

Complementing the Consultation Workshops

The workshop consultations were complemented by an online survey involving a series of open- and closed-ended questions on KAs that may form part of CyBOK. The survey sought participants' views on topics such as the KAs that had been the most important background knowledge in their career, key KAs that ought to be covered in CyBOK and those that should be out of scope, and topics that would be of most importance over the next five years.

Semistructured interviews were conducted with 10 leading international experts in cybersecurity. The interviews included both technical experts in computer security and those studying topics such as human factors, governance, regulation, risk, and law.

A small amount of input was also received through an open call for position papers.

Analysis of Various Texts Listing Key Topics

We complemented the data arising from the above community

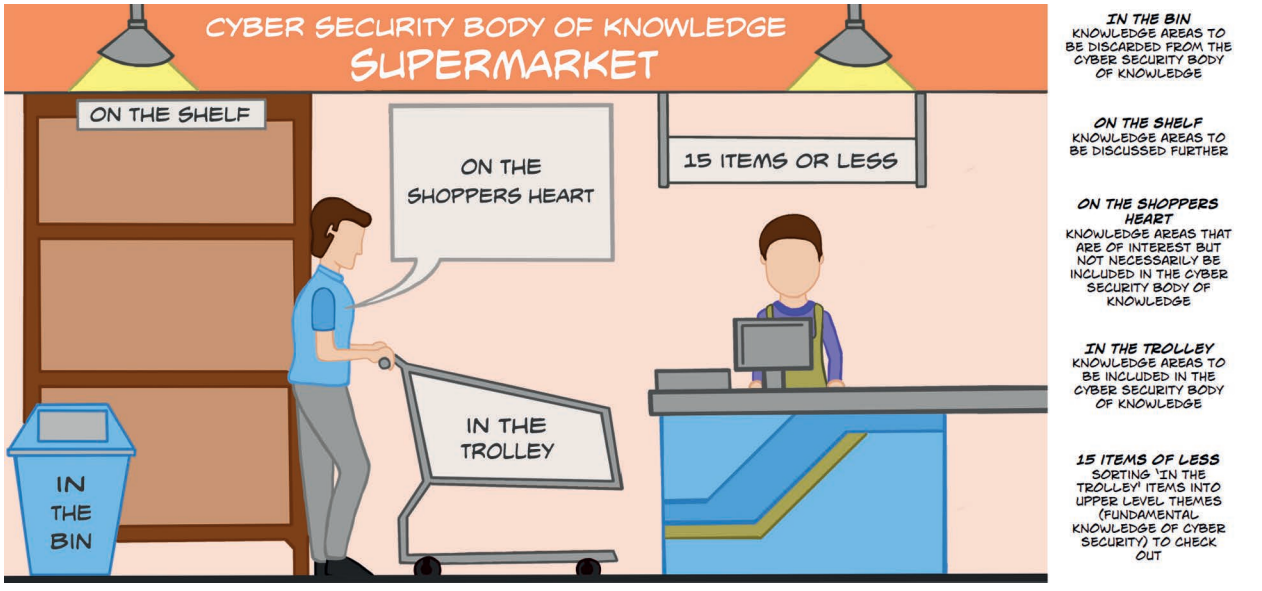

Figure 1. The supermarket metaphor used in participatory workshops to determine KAs for CyBOK.

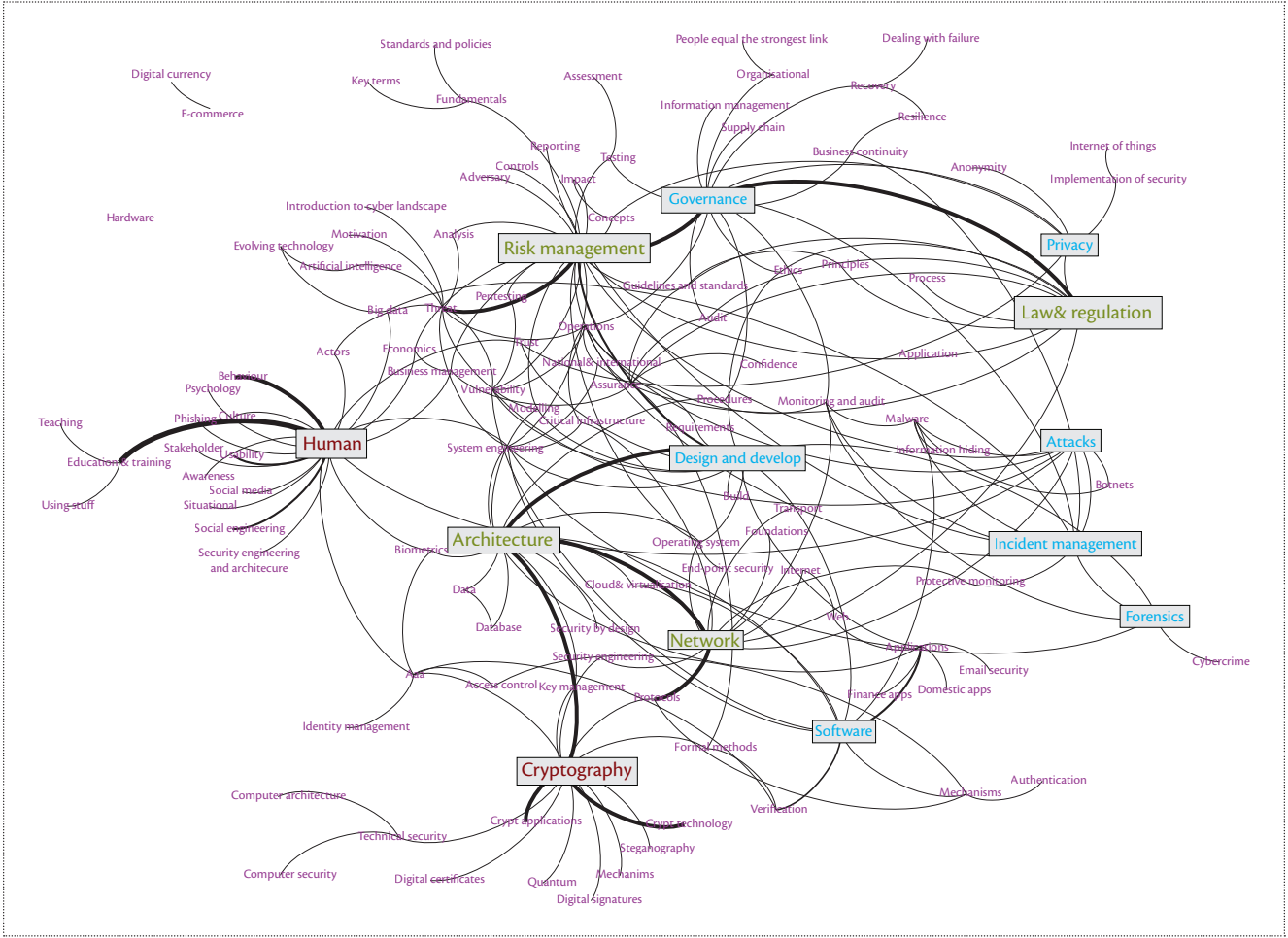

Figure 2. An example graph of distilled KAs derived from the analysis of workshop data.

consultations with analysis of a number of documents that typically list key topics relevant to security. Example documents included:

- Categorizations, such as the ACM Computing Classification System (CCS) taxonomy;
- Certifications, such as Certified Information Systems Security Professional (CISSP) and the Institute of Information Security Professionals (IISP) Skills Framework;
- Calls for papers such as IEEE Symposium on Security and Privacy and USENIX Symposium on Usable Privacy and Security;
- Existing curricula, such as the ACM computer science curriculum and the work of the Joint Task Force on Cybersecurity Education;
- Standards, such as BS ISO-IEC 27032 2021 and NIST IR 7298; and
- Tables of contents of various textbooks.

We used a variety of text-mining techniques, such as natural language processing and automatic text clustering to group relevant topics and identify relationships between topics. Techniques included semantic word cloud visualizations, word vectors, ward clustering, K-means clustering, and Latent Dirichlet Allocation.

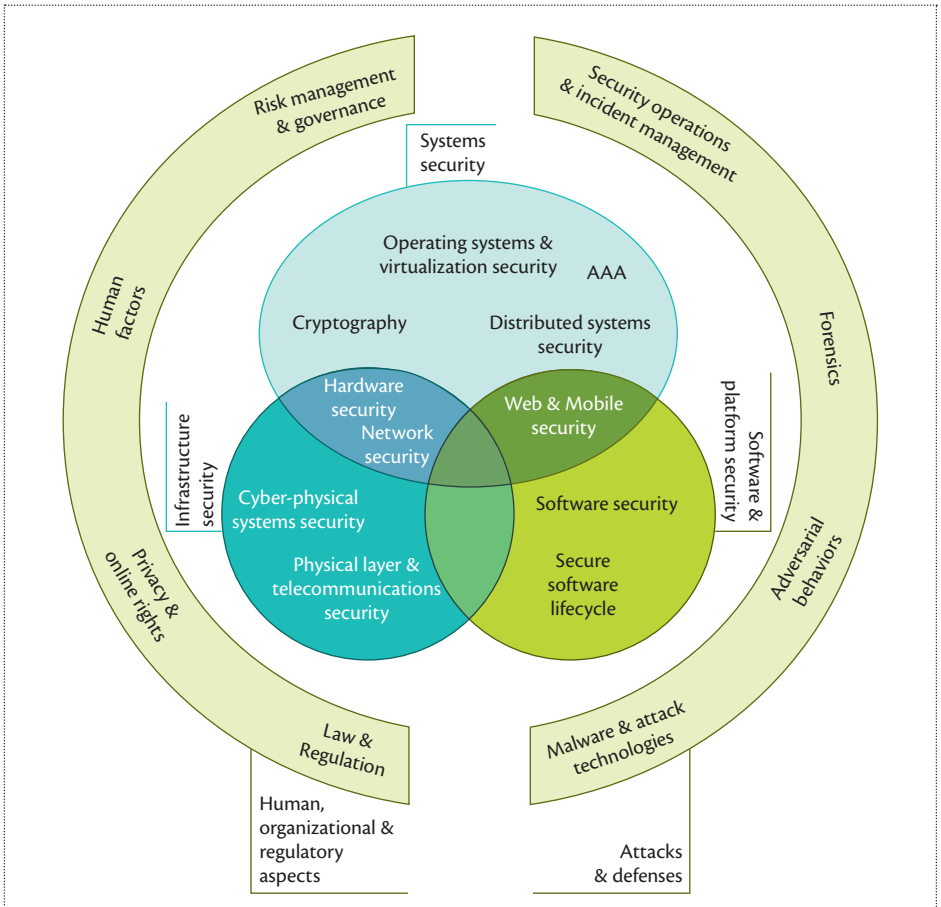

Figure 3. The 19 knowledge areas and their categorization within CyBOK.

Distilling the Knowledge Areas

Workshop participants identified key topic areas together with subsidiary topics that they believed should be included in each area. This provided the opportunity to visualize the workshop data as a graph in which nodes were highlighted according to the strength of recommendation as a key area, and edges weighted to show the strength of relationship between topics. Inevitably, the workshops resulted in a large number of unique terms for topics: a total of 906 unique terms, with 660 occurring only once in the record. Some data cleaning was therefore necessary, and was carried out via an alias list that could be inspected and reviewed. This resulted in 483 unique terms, in which a core of 144 topics occurred more than once. The final graph was filtered by edge weight to allow review of the data at different levels of granularity. An example graph is shown in Figure 2. Note that the colors represent visual distinction into quartiles and that the size of each topic name is proportional to its frequency as a nominated key area.

The graph in Figure 2 is one example of the types of graphs that served as the starting point for our distillation of KAs. The thematic clusters emerging from such graphs were cross-referenced against the data from the survey, the interviews, and the position statements. The subtopic lists in these thematic clusters were further compared to the clusters identified through the text-mining analysis of documents listing key topics in cybersecurity. During this synthesis, we particularly attended to topics—for instance, hardware security and cyber-physical systems security—that appeared disconnected or did not formulate large clusters in the graphs but were highlighted by our survey or interview participants as key emerging topics of importance over the next five years. This analysis and synthesis resulted in 19 KAs, grouped into five broad categories (see Figure 3 and Table 3). Figure 3 shows that

Table 3. Overview of the 19 knowledge areas.

Human, Organizational, and Regulatory Aspects	
Risk Management and Governance	Security management systems and organizational security controls, including standards, best practices, and approaches to risk assessment and mitigation.
Law and Regulation	International and national statutory and regulatory requirements, compliance obligations, and security ethics, including data protection and developing doctrines on cyber warfare.
Human Factors	Usable security, social and behavioral factors impacting security, security culture and awareness as well as the impact of security controls on user behaviors.
Privacy and Online Rights	Techniques for protecting personal information, including communications, applications, and inferences from databases and data processing. It also includes other systems supporting online rights touching on censorship and circumvention, covertness, electronic elections, and privacy in payment and identity systems.
Attacks and Defenses	
Malware and Attack Technologies	Technical details of exploits and distributed malicious systems, together with associated discovery and analysis approaches.
Adversarial Behaviors	The motivations, behaviors, and methods used by attackers, including malware supply chains, attack vectors, and money transfers.
Security Operations and Incident Management	The configuration, operation, and maintenance of secure systems including the detection of and response to security incidents and the collection and use of threat intelligence.
Forensics	The collection, analysis, and reporting of digital evidence in support of incidents or criminal events.
Systems Security	
Cryptography	Core primitives of cryptography as presently practiced and emerging algorithms, techniques for analysis of these, and the protocols that use them.
Operating Systems and Virtualization Security	Operating systems protection mechanisms, implementing secure abstraction of hardware, and sharing of resources, including isolation in multiuser systems, secure virtualization, and security in database systems.
Distributed Systems Security	Security mechanisms relating to larger-scale coordinated distributed systems, including aspects of secure consensus, time, event systems, peer-to-peer systems, clouds, multitenant data centers, and distributed ledgers.
Authentication, Authorization, and Accountability	All aspects of identity management and authentication technologies, and architectures and tools to support authorization and accountability in both isolated and distributed systems.
Software and Platform Security	
Software Security	Known categories of programming errors resulting in security bugs, and techniques for avoiding these errors—both through coding practice and improved language design—and tools, techniques, and methods for detection of such errors in existing systems.
Web and Mobile Security	Issues related to web applications and services distributed across devices and frameworks, including the diverse programming paradigms and protection models.
Secure Software Lifecycle	The application of security software engineering techniques in the whole systems development lifecycle resulting in software that is secure by default.
Infrastructure Security	
Network Security	Security aspects of networking and telecommunication protocols, including the security of routing, network security elements, and specific cryptographic protocols used for network security.

Table 3. Overview of the 19 knowledge areas (cont.).

Hardware Security	Security in the design, implementation, and deployment of general-purpose and specialist hardware, including trusted computing technologies and sources of randomness.
Cyber-Physical Systems Security	Security challenges in cyber-physical systems, such as the Internet of Things and industrial control systems, attacker models, safe-secure designs, and security of large-scale infrastructures.
Physical Layer and Telecommunications Security	Security concerns and limitations of the physical layer including aspects of radio frequency encodings and transmission techniques, unintended radiation, and interference.

Related Work on Identifying Core Concepts in Cybersecurity

The ACM, IEEE, Association for Information Systems Special Interest Group on Security (AIS SIGSEC), and International Federation for Information Processing Technical Committee on Information Security Education (IFIP WG 11.8) Joint Task Force (JTF) on Cybersecurity Education has developed guidelines for undergraduate curricula in cybersecurity (http://cybered.acm.org). Eight principal knowledge areas are considered, based on the entities to be protected: data security, software security, component security, connection security, system security, human security, organizational security, and societal security. These are complemented by crosscutting concepts such as confidentiality, integrity, and availability. Undergraduate cybersecurity curricula can then be designed for particular disciplines, for instance, computer science, software engineering, and so forth, and/or linked to particular application areas. In contrast, the Cyber Security Body of Knowledge (CyBOK) project (www.cybok.org) aims to codify foundational knowledge that can inform the design of cybersecurity education and training programs at a range of levels: from secondary and undergraduate to postgraduate and continuing professional development. It complements the work of the JTF by providing in-depth coverage of knowledge areas (KAs) and key resources that curriculum designers can utilize.

The National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework[1] has established a set of specialty areas and mapped them to roles in the cybersecurity workforce. The focus is on skills and the tasks a person in a particular role ought to be able to perform. CyBOK can form the basis of charting the learning pathways that such skilled roles may need to take across the 19 KAs (or a subset thereof) in order to be able to proficiently perform the required tasks.

The Cybersecurity Assessment Tools (CATS) project has undertaken a Delphi study identifying the importance, difficulty, and timelessness of particular cybersecurity topics.[2] Such understanding is essential to the design of cybersecurity education programs. It would be interesting to explore where the topics of most difficulty and importance appear in the 19 CyBOK KAs and—combined with charting of learning pathways for the NICE framework—how this may inform pedagogical approaches to cybersecurity.

The security counterpart to the IEEE Software Engineering Body of Knowledge (SWEBOK; www.computer.org/web/swebok) is "Software Assurance: A Guide to the Common Body of Knowledge to Produce, Acquire and Sustain Secure Software."[3] This has similar style and chapter headings to SWEBOK and provides a summary of knowledge relating to software and the software lifecycle. Software development is within the scope of CyBOK, which in contrast has the wider scope of fundamental and applied knowledge in all aspects of cybersecurity.

Knowledge required for the Certified Information Systems Security (CISSP) examination has also been codified in a body of knowledge.[4] The CISSP CBK documents the knowledge required for a specific examination in a summary textbook; this is in contrast to other bodies of knowledge and CyBOK, the contents of which guide readers to knowledge contained in authoritative references.

system, infrastructure, software, and platform security is shaped by human and organizational factors and vice versa. At the same time, cybersecurity of technologies, people, and organizations requires a deep understanding of attacker behaviors and attack technologies as well as effective responses for analysis of attacks and incident management and response. We note that other possible categorizations of these KAs may be equally valid. Also the categories are not necessarily orthogonal.

Next Steps

The initial CyBOK scope and KAs identified above were made publicly available for community comments in September 2017. Although none of the 19 KAs needed to be removed or new ones added on the basis of the feedback, the topics to be covered under each KA have been refined. As a next step, authors will be invited to write detailed descriptions of KAs, which will be reviewed by a small panel of peer reviewers before being made available for public consultation. As each KA description is finalized, it will be made available on the CyBOK website. We aim to complete all KA descriptions by the end of July 2019. In addition, learning pathways through CyBOK and exemplar curricula at different education levels will be developed. We will undertake a series of consultations through workshops and interviews with stakeholders not only involved in university education but also from primary and secondary education as well as industrial training programs. Combined with desk research on curricula, such consultations will form the basis to develop a set of exemplar learning pathways as a set of case studies for utilizing CyBOK in educational programs.

Cybersecurity is a rapidly changing and evolving field. As such, CyBOK will never be "finished" per se. Future iterations will need to be undertaken to ensure that the coverage remains up to date and the KAs reflect both the current state of knowledge in cybersecurity and emerging needs. The inclusion of KAs such as hardware security and cyber-physical systems security in the current scope reflects such emerging needs. Any future maintenance of CyBOK will need to ensure that, while not ignoring the needs of contemporary and legacy systems, the CyBOK scope also reflects key challenges arising from the increasing integration of technology—and hence cybersecurity—into the very fabric of our society. ■

References

1. W. Newhouse et al., "National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework," NIST Special Publication 800-181, NICE Framework, Aug. 2017; https://dx.doi.org/10.6028/NIST.SP.800-181.
2. G. Parekh et al., "Identifying Core Concepts of Cybersecurity: Results of Two Delphi Processes," *IEEE Transactions on Education*, 2017.
3. S.T. Redwine, ed., "Software Assurance: A Guide to the Common Body of Knowledge to Produce, Acquire and Sustain Secure Software," Version 1.1, US Department of Homeland Security, Washington, 2006.
4. A. Gordon, ed., *Official (isc) 2 Guide to the Cissp CBK*, CRC Press, 2015.

Acknowledgments

The CyBOK project is sponsored by the National Cyber Security Programme in the UK. The authors also thank Yvonne Rigby, project manager for CyBOK, for her excellent work on coordinating the various strands of research across the project.

Awais Rashid is professor of cyber security at the University of Bristol, UK. Contact at awais.rashid@bristol.ac.uk.

George Danezis is professor of security and privacy engineering at University College London. Contact at g.danezis@ucl.ac.uk.

Howard Chivers is honorary fellow at University of York, UK. Contact at howard.chivers@york.ac.uk.

Emil Lupu is professor of computer systems at Imperial College London, UK. Contact at e.c.lupu@imperial.ac.uk.

Andrew Martin is professor of systems security at University of Oxford, UK. Contact at andrew.martin@cs.ox.ac.uk.

Makayla Lewis is a research fellow at Brunel University London, UK. Contact at Makayla.Lewis@brunel.ac.uk.

Claudia Peersman is senior research associate at University of Bristol, UK. Contact at claudia.peersman@bristol.ac.uk.