CASE FOR SUPPORT

# AI-Trust: Computer Architecture for Confidential and Private Computing for AI in HPC Environment

**Dr. Ahmad Atamli**
UNIVERSITY OF SOUTHAMPTON

## 1 Vision

The vision driving this project is to **establish** new research discipline of confidential and privacy preserving computing for Artificial Intelligence (AI) in High Performance Computing (HPC) environment. A technology that is designed for intensive computation on large datasets.

In this project, inspired by real-world use-case requirements from modern computation, I intend to devise, investigate and optimise a new solution that will result in a *change of paradigm* in computer architecture and allow **confidential and privacy preserving execution** for AI applications. I envision the outcome of this work to be universally applicable beyond AI, and pave the way towards exploration of similar approaches in future computational platforms including quantum computers.

This project and all its activities fall squarely within UKRI vision and mission of creating a secure and scalable systems technologies to benefit and enrich lives nationally and internationally across multiple disciplines and sectors.

## 2 Research and Innovation Excellence

The developed world makes ever-increasing demands for networked computation that is AI-based in support of society's needs and the everyday lives of citizens.

Specifically, one technique called "deep learning" has been widely used in diverse areas of research including image captioning, speech recognition, robotics, genomics data analysis, biomedical science, diagnostics, finance, and self driving cars.

The potential of AI in modern applications has led to explosive demand for computing and storage to operate on large datasets. This resulted in hardware manufacturers racing to provide specialised hardware (e.g. TPU, GPU) that was specifically designed for raw compute, parallelism, and high memory bandwidth . Many organisations tried for years to maintain and manage their infrastructure(s). However, the rapid advancement in the semiconductor world left them lagging behind due to lack of budget to maintain and renew old equipment. As a result, the world turned to cloud computing to provide the computing and storage necessary for AI applications.

**Cloud computing** provides resources to users on demand, free of the overhead of system management and maintenance [7] associated with physical resources. It delivers on the promise to reduce the costs of power, maintenance, and operations. Cloud environments are based on sharing of hardware resources among many users, hundreds or more. This allows the cloud model to scale as the number of users increases, while keeping it affordable for the user and profitable for the cloud provider. But security in this setting remains a very serious concern.

When cloud computing was introduced, the industry rushed to implement it using the conventional processor architecture already in place for desktops and laptops, without thorough evaluation of the risks associated with this new model [6]. The security of cloud infrastructures differs from classical computing: it's complex and challenging and introduces new regulations and requirements. For example, medical records of the UK citizens are not to be processed and stored by computing platforms outside the UK. Enterprises and organisations need guarantees of confidentiality and integrity: guarantees that a malicious actor cannot leak or tamper with users' data. This includes both a malicious actor in the cloud organisation and other enterprises/users sharing the same resources.

The crux of the problem is that the monolithic structure of a cloud application makes it vulnerable to malicious attack [4]. The problem is serious, as demonstrated by recent breaches suffered by airlines and health providers[1].

The cloud environment has a unique infrastructure where appropriate security solutions are lacking. Current solutions for security problems using differential privacy, homomorphic encryption, and Trusted Execution Environment (TEE) [3, 1, 2] are not easily transferable to HPC due to various contextual factors, including lack of appropriate technologies, suitability to AI, performance degradation, and limited computation

[1] WannaCry - https://bit.ly/34KkKE6

functionality.

In a previous project (ASSURED)[2], I have proposed a new platform architecture as part of a European-led effort to tackle the challenge of integrity assurance in systems-of-systems. In ASSURED, the new proposed platform architecture establishes trust between all components in the system prior to any execution. Thus, allowing only *good* hardware devices to be connected.

But the question remains on how we can maintain security of AI applications when the data is processed outside the TEEs, e.g. by specialised hardware accelerators such as GPUs and FPGAs.

To overcome this challenge I propose that trusted execution should not reside in a single place or component (e.g. a TEE) and that modern computers must provide a suitable root to secure data outside a TEE restricted to the main CPU. This proposal aims to identify and investigate capabilities that can enhance security in complex platforms, while keeping them purposed for modern application of AI and data analytics, affordable, readily deployable and energy efficient. My approach focuses on capabilities rooted in hardware, based on the core concept of trust. Operations and elements that are rooted-in-hardware cannot be modified and thus cannot be compromised by malicious actor.

### 2.1 Novelty and Timeliness

In this project I have set-out on an unprecedented path to address the security challenges faced in *modern* HPC environment: complex assemblies of CPUs, I/O and communications infrastructure, and specialised hardware devices. Old computing architecture relied on the main CPU to execute the work load and increased the number of computing cores on demand for more computation. This approach of horizontal scaling of resources has several limitations of security, energy consumption, increased manageability, increased waste, and maintenance due to the increase of chip die-size and number of platforms to deliver on performance requirements. HPC platforms improve efficiency by using specialised hardware to offload functionality (e.g. deep learning). My novel and timely programme of research will be the first to treat the problem of confidentiality and privacy in HPC environment, introducing a new discipline that has never been explored before. As the digital society is faced by increasing computation requirements on private data across many sectors it is of paramount urgency to address those challenges in HPC.

The novelty of the proposed project lies in exploring a new research area of confidential and privacy preserving computing for HPC, repurposing an existing technology concept referred to as **confidential computing** and propose a novel computer architecture appealing to modern world applications of AI and data analytics.

This application is a step-change in the HPC world and will result in moving the global industry forward to adopting a novel computer architecture, with all the technical security, privacy, and economic benefits. I will work in the second half of this fellowship (years 5-7) to accelerate the delivery of this architecture with my industrial partners to provide tools for other sectors to advance in employing the capabilities for confidential and privacy preserving AI.

## 3 Research Hypothesis and Objectives

My **research hypothesis** is that the confidentiality and privacy guarantees implemented in CPU hardware can be extended to computing platforms that include specialised hardware devices outside the CPU (e.g. GPU, TPU). The **ultimate aim** of this work is to make a step-change in the security of HPC, allowing privacy preserving and confidential execution. My **immediate aims** are (1) to explore the fundamental technological basis for extending the architecture across the platform and (2) to validate this with a significant case study: off-CPU AI module training.

Pursuant to these aims, the main **objectives** of this research are to:

1. Perform an extensive and comprehensive analysis of requirements for AI-based applications in the health, finance and government sectors.

2. Devise and demonstrate the hardware architecture and software abstractions needed to extend **trusted execution** from user code running on the CPU to specialised devices with access to host memory and between specialised hardware devices (e.g. PCI, InfiniBand).

3. Devise, prototype, and evaluate the hardware design **trusted execution** from 2 in the context of performance and latency.

4. Devise and demonstrate mechanisms for off-CPU, specialised devices to gain and

[2] https://www.project-assured.eu

exercise **trusted execution** with respect to applications running on the CPU.

5. Demonstrate and evaluate the experimental technology developed under objectives (2) and (4) with a significant **case study** of an AI application.

6. Evaluate the scalability, overall performance, economics, efficiency, waste, and energy consumption of the proposed technology compared to state-of-the-art for the **case study** .

These research objectives represent a path-finding endeavour to develop a computer architecture for security in complex HPC platforms that heavily rely on using acceleration hardware devices. This project looks to answer two concrete questions. First, can there be computer architecture that allows extending trusted execution environment to specialised hardware that provide security, performance, and efficiency of execution of modern world AI applications? Second, how can the proposed hardware architecture improve the scalability, overall performance, economics, efficiency, waste, and energy consumption?

My research programme is advocating for a change of paradigm where we no longer rely on existing and wasteful solutions but rather develop novel architectures which will empower multitude of novel software architectures to meet the requirements of today's modern technological challenges. The research aims are squarely aligned with the overall vision and objectives of UKRI and the research vision to make the UK a leading power in secure communications. My project will conduct aspiring fundamental research that is novel and spans technologies beyond CPU-focused industries (e.g. Intel and AMD) and include high speed protocols and AI applications. This proposal lays the foundations for new opportunities that across different technologies: CPU, protocols, network, and acceleration. Each of these individual components represents an exclusive niche dominated by specific industry with no single party adopting a holistic view of the whole system. My view of delivering a holistic investigation will only be possible through the close collaboration with my academic and industrial partners which will accelerate the delivery of this interdisciplinary work to maximise its impact on society and facilitate the adoption of this innovative technology by the industry and its dissemination to other sectors.

## 4 Programme and Methodology

This research programme as shown in Figure 1 has five main work-packages (WP): WP1 focused on studying the frameworks of AI used by use-cases applications , WP2-WP4 aimed at developing the fundamental mechanisms and architecture , and WP5 will demonstrate use-cases applications, and evaluate the proposed solution.

This methodology benefits from my previous track record (see **CV**) and experience in using specialised hardware devices to extend secure execution outside the main CPU [5], as well as long experience in designing chips and Systems on Chip (SoC) hardware, and the accompanying firmware and software in industrial setting. I will also acquire new technical skills and venture to establish new discipline through partnerships and training as explained for each WP.

### 4.1 WP1 - AI framework and Data

This WP will curate the requirements to shape the software and hardware architecture development and prototyping. As part of this activity, I will work with the **Alan Turing Institute** and **Microsoft** to understand their AI applications and current Infrastructure and how their applications make use of AI frameworks and specialised hardware devices.

There are several types of data, varying in size and algorithms used by AI frameworks. Many AI frameworks exist to facilitate the work with specialised hardware. The aim of this WP is to perform comprehensive investigation into use-cases applications and AI frameworks (e.g. PyTorch) commonly used by researchers and industries. Many AI application uses PyTorch, an open source machine learning framework that accelerates the path from research prototyping to production development. PyTorch works with hardware acceleration and is used by the partners providing the use-cases. This research will illustrate the partition between CPU and hardware acceleration, user space code and privileged code execution, shard buffers size, and volumes.

**Outcome.** Specifications of new CPU architecture and communication protocol extending trusted execution to specialised hardware device.

### 4.2 WP2 - Trusted Execution Enablement for AI applications

Current TEE architectures by Intel, AMD and IBM allow TEE computation to execute on the main CPU, while blocking access by external software and hardware devices. As AI applications rely

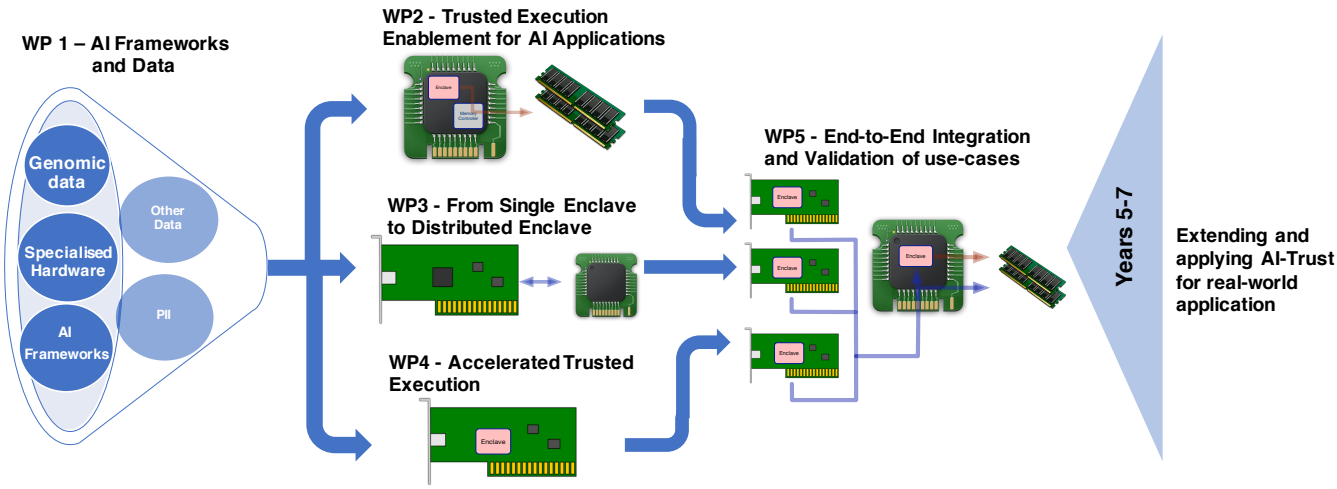


Figure 1: AI-Trust Work-Packages

on the computation capabilities of external hardware, it is imperative to enable access to an TEE for secure operations. The aim of this work-package is to investigate the hardware primitives in the access control module of the main CPU and specialised hardware. The new model will enable presence of the TEE in multiple places (e.g. CPU and GPU) and allow efficient sharing of memory space and exchanging data while blocking other software and hardware. This will be developed either as a new CPU instruction that is used by the application to grant access to an external device, or firmware flow handled by an internal co-processor. This will mandate a change within the Memory Management Unit (MMU) that will extend the memory access module of a TEE memory regions on a host machine to an external specialised hardware.

#### WP2.1 - Hardware Enablement

A new CPU architecture will be developed and prototyped to provide a TEE with the ability to extend the trust model from its internal execution to an external hardware. This will manifest in changes in the CPU architecture allowing peripheral access to the TEE memory, caches, register and internal execution.

#### WP2.2 - Software Enablement

A software abstraction will be developed to be used by AI frameworks when using specialised hardware, this library will leverage the new hardware architecture to configure and authorise the right policy and access for a hardware device.

This work-package will benefit from the partnership with Queens University Belfast and decades of experience in designing SoC by Professor Sakir Sezer. Prof. Sezer will advise on the new micro-architecture and software-hardware interface that will allow an application to enable access of specialised hardware device to TEE memory space.

**Outcome.** New CPU architecture and design that supports TEE memory access from specialised hardware device.

### 4.3 WP3 - From Single TEE to Distributed TEE

Specialised hardware accesses the host machine through Interconnect buses (e.g. the PCIe Bus). Current computer architecture allows communication with the main CPU and peer-to-peer communication over the PCIe bus. But this is under the complete control of the resident OS/hypervisor/cloud provider and the protocol lacks security such as preventing an interposer device or malicious software from observing and modifying the controls of the interconnect, hence, does not provide confidentiality and integrity of the data transported. The aim of this work-package is to design security in the hardware PCIe protocol between components on the platform. This will securely connect TEEs running on the host machine and associated computation running on the specialised hardware.

Specifically, the objective of this work-package is to create an Interconnect mechanism for extending trust from software TEEs in the CPU to specialised hardware. To perform the design and evaluation, an FPGA board will be used as a prototype to develop the new hardware and protocol proposed. why are you not using GPU for prototype

#### 4.3.1 WP3.1 - Interconnect Protocol Design

A secure-by-design protocol will be developed to establish the channel between TEE components in the system (e.g. between TEE instance in the main CPU and TEE's instance in a specialised hardware device and specialised hardware to specialised hardware). As part of this activity, the new approach's correctness, cryptographic strength, performance, bus utilisation, and latency will be evaluated. We will investigate the two most relevant protocols for this work used in HPC platforms, the PCIe protocol and InfiniBand (IB) .This work-pacakge will benefit from the mentorship of Peter Peneah at Nvidia. Peter has 20+ years in designing PCIe protocols and will advise on the new micro-architecture and protocol design of security-by-design PCIe protocol that will form a trusted channel between TEE and on-board specialised hardware device.

#### 4.3.2 WP3.2 - Interconnect Protocol Validation

In addition to designing the protocol, I will work with in partnership with Professor Vladi Sassone and co-supervised PhD student on verifying the correctness and confidentiality and privacy assumptions of the designed protocol. This work-package will benefit from the mentorship of Professor Vladi Sassone on formal. Professor Sassone has 20+ years in formal methods and verifying the correctness and assumption of hardware design. Professor Sassone will recruit and co-supervise with myself a PhD student that will work on verification of the correctness and assumptions developed in this work.

**Outcome.** A novel interconnect protocol for establishing trusted communication between TEE's instances in different components in the same platform. The outcome of this WP will be communicated to the PCI Express working group, of which I am an active member, making it available to all industries.

### 4.4 WP4 - Trusted Execution on Specialised Hardware

This work-package will explore how to achieve trusted execution by a specialised hardware (e.g. FPGA). The architecture of most specialised hardware differs from the main CPU where the users run their application. The assumptions regarding resource allocation are significantly different. For example, an FPGA might be serving 1-2 users with 1-2 TEEs at a time, unlike host machines that can serve hundreds of users.

#### 4.4.1 WP4.1 - Confidential and Privacy preserving FPGA Architecture

I will explore a new approach to trusted execution between different TEEs within the specialised hardware device. As part of this work-package we will prototype the abstraction to provide trusted execution between multiple TEEs on the same host machine. This work-package will benefit from the partnership with Dr. Suhaib Fahmy from King Abdullah University for his experience working on FPGAs. Dr. Fahmy will recruit and fund a PhD student who will be dedicated to prototyping the FPGA implementation which will be developed as a proof-of-concept for this work and the <name project at Suhaib end>. The PhD student will be co-supervised by Dr. Fahmy and myself.

#### 4.4.2 WP4.2 - Energy Efficent Embedded System Design

An evaluation of the energy consumption as a result of the new changes that rely on cryptographic operation. The evaluation will guide the implementation and options supported by the proposed protocol. This work-package will benefit from the results of the PRiME project and the mentorship of Prof. Bashir Al-Hashimi at King's College London, who has track record researching the areas of energy consumption. Professor. Al-Hashimi will advise on the techniques to evaluate the energy consumption and design techniques to reduce power consumption.

**Outcome.** Proof-of-concept trusted execution architecture and design on an FPGA.

### 4.5 WP5 - End-to-End Integration and Validation of Use-cases

This work-package will evaluate a number of applications and case studies of our core technology. We aim to address the following questions. What opportunities are provided by the proposed hardware in securing trusted execution outside the main CPU? How does this approach perform compared to software running on host machine? Can the on-board accelerators such as DMA engines, GPUs, dedicated hardware, and FPGAs allow for a more efficient, economic, and low latency security solution compared to current approaches (e.g. differential privacy, homomorphic encryption, TEEs)?

**WP5.1: AI Applications Integration**

This work-package will focus on porting the AI applications to the proposed architecture, while considering the different software layers: the kernel, device driver, operating system, and the

hypervisor. This is an end-to-end integration and evaluation of the AI applications we curated in WP1. We will aim to evaluate applications against native performance and latency and split between the CPU and the specialised hardware accelerator. This work-package will benefit from collaboration with the Alan Turing Institute, Microsoft, and University of Oxford which will provide the use-case AI application, algorithms, and datasets.

**WP5.2 - End-to-End Evaluation against state-of-the-art**

This work-package builds on the integrated system flow from WP5.1 and brings together the application and data we will use to evaluate our work and falsify or support my main research hypothesis. To that end, I will evaluate the developed architecture against current state-of-the-art computing platforms through several case studies. Specifically, our metrics will include performance, latency, scalability, economical model, efficiency, and power consumption, and waste generated. By the end of this WP we would have clear view of the capabilities of the proposed architecture, and identify specific aspects that require further refinement to meet and exceed current standards. As this would be the ultimate proof-of-concept to our methodology this WP will benefit from the partnership with the Alan Turing Institute and Microsoft to evaluate the results against current state-of-the-art.

**Outcome.** 1) An end-to-end AI application integration with trusted execution between the specialised hardware and main CPU. 2)Realise the potential of proposed computer architecture for AI applications by evaluating the superiority of the approach given all the relevant constraints.

add a small paragraph on plans for years 5-7

### 4.6 Project Management

The success of the management structure of this large-scale project is through an authoritative **advisory network** with the different partners and technical mentors. All are industry and academic leaders with deep and relevant technical knowledge, as well as profound understanding of the business context for the impact of my work. My view is that a project of this kind will benefit more from an agile and flexible approach, rather than (say) yearly formal advisory board meetings. I will throughout the project consult the different partners on specific and focused topics, as well as maintain their awareness of our primary results. The specific responsibilities and involvement of each staff member and partner is highlighted in the workplan attached to this proposal.

## 5 Applicant and Development

I have led a cross-disciplinary career at the interface of industry and academia for more than a decade. Alongside my academic training (See **CV**)I have actively engaged in partnerships and advised industrial companies including Nvidia, EZchip, Mellanox, Voltair, Custdio, and Microsoft on strategy, vision, and security in next generation System on Chip (SoC) and datacenters. I have spearheaded the research behind two proposals that received €1M industrial funds from Horizon2020 (ASSURED, TrustedFog) to develop trust in modern computing platforms using trusted computing and led the development of an innovative proposal that secured $780,000 award for Custdio [3]. I have a track record of delivering cross-disciplinary research **(outputs list)** where I continuously communicate with researchers, product engineers and customers alike. My communication skills across the different levels transpires from the numerous invitations to present my research, share knowledge, and give training and webinars in both academic and industrial spheres throughout the years with 26 talks in merely two years (10 in 2019 and 16 in 2020).

I constantly endeavour to engage and develop new relationships across disciplines, with the aim of solving real world challenges and making an impact on society. I am actively engaged with PCI Express group and Confidential Computing Consortium where I have contributed to novel definition of specifications adopted by semiconductor industry. As part of my research vision, I am consolidating my current collaborations and expanding them to include pioneers from the health sector (Dr. Shivan Sivakumar, Pancreatic Oncologist, Oxford), artificial intelligence (Prof. Jon Crowcroft, Alan Turing Institute) as well as sustainable technologies (Prof. Bashir Al-Hashimi, KCL). The outlined vision in this FLF is purposefully ambitious and involves a network of support from different academic institutions and industrial organisations. This core feature will give me the mentorship, technical support and confidence to deliver the outcomes of this project. I have identified world leaders in different disciplines and with specific expertise instrumental for each work-package.

More importantly I have devised a detailed

[3] https://securitybrief.asia/story/singapore-cybersecurity-awards-celebrate-top-cyber-talent

personal development plan. I have already taken time, research and product management courses to acquire a set of skills that will enable me to effectively and smoothly run my first formally independent group. As part of the FLF I intend to undertake training in Public Speaking and engagement (Media coaching), Coaching for Professional Leadership (Sir John Whitmore), science communication (UWE Bristol) and Entrepreneurship Development Program (MIT). Furthermore, as a recipient of a UKRI fellowship I will be eligible to join the Innovation and Business of Science training offered by the Royal Society. This programme will enhance my technical skills as well as leadership and outreach attributes. I will also support the PDRAs and PhD students working with me to attend personal development at the University of Southampton (UoS) and public engagement training (NCCPE, Famelab international) and undertake internships with industrial partners to gain new skills.

I have also identified mentors to support me throughout the fellowship including Prof. Shiyan Hu at UoS who will be instrumental to guide me through my early career and provide the advice to succeed as an academic. I will also get mentorship from two industrial partners Peter Peneah at Nvidia and Boris Pinzur at Microsoft who will guide me in identifying the best impact vectors for the industry.

My vision is to create a safe cyber society across different sectors which is amenable to modern technological advancement and is environmentally conscious. In order to achieve this, I will share the knowledge and expertise I gain with my team of two PDRAs and four students with other centres, start-up and industrial partners across the UK and internationally. I have also identified opportunities through the Royal Society to pair with UK parliamentarians and will engage with those to identify the best routes to make an impact on policymakers. I am also leading the curriculum innovation in the Academic Centre of Excellence in Cyber Security Education (ACE-CSE) at UoS which will allow me to engage with students and train them with emerging novel concepts from this work.

**Career Intentions** : Receiving a FLF at this stage of my career will have a significant impact on my future progression. On the long-term, I aspire to lead an internationally recognised research and innovation team with global impact to shape creative and holistic security solutions adaptable to developing and emerging technologies. My professional mission statement is that research can no longer be done in isolation, and we need to work closely with other sectors in order to deliver innovative solutions of real-world challenges to improve human life through partnership with the industry and policymakers.

Over the short-term the FLF will enable me the flexibility and freedom to develop and coordinate a large-scale project to explore a novel territory. This will allow me to focus on consolidating my skills and leadership, continuously engage with partners, disseminate the outcome and deliver on this project. By the end of the fellowship, I will have an international reputation, introduce and lead new research discipline of confidentiality and privacy preserving technologies in HPC. I will also have gained new personal, communication and technical skills, which will support my trajectory to become a renowned professor.

My mid-term vision includes spin-out to a new business that focuses on building new hardware enabling security in HPC environment, consultation to businesses and sectors implementing this technology, and conducting workshop and academic courses in the field.

In the long-term, this award will support me in devising appropriate growth strategies to create a cyber-resiliant society. The close mentorship by multiple academic and industrial partners as well as the national and international network of peers and collaborators across multiple sectors will be valuable assets to join my future work.

## 6 Impact and Strategic Relevance

In recent years, the UK and the western world in general, have suffered from a surge in cybercrime instances on government, public and private institutions, resulting in financial loss of billions of pounds and privacy exposure to thousands of users. AI has become an important tool to solve today's problems, but its adoption in cloud HPC environemnt poses security and privacy problems. Users' and companies are increasingly suspicious about sharing their data even with the health sector or academics who strive to employ AI and data analytics for diagnostics and data interpretation for the public's benefit.

The outcome of this research will ensure safe and trusted cyber society by producing new technology for people and businesses that overcome the performance and limitations of differntial privacy, homomorphic encryption, and current TEE technology. It will pioneer a novel infrastructure for **confidential and privacy preserving computation** on AI large volumes of data while con-

sidering efficency, scalability, reliability, and environmental impact. This will maintain the health in other sectors such as the medical sector. It will allow medical practitioners to analyse users' data without requiring patients to waive their privacy. It will also allow researchers to recruit unlimited resources of cloud computing saving on huge sums of money on infrastructure that will soon become outdated and require constant updates. The proposed solution will provide governments and orgnisation the level of confidentiality and privacy requierments needed. The UK Research Councils have identified the importance of creating a secure and scalable systems as part of their Information and Communication Technologies delivery theme since 2013. It comprises 7% of the ESPRC funding portfolio and spans many sectors including healthcare, energy and finance. Cyber-security tops the 6 global uncertainties that pose risk to modern society and need to be addressed in order to help governments and societies better predict, detect, prevent and mitigate threats. Out of the top 7 future priorities of this challenge this specific proposal addresses Cyber-crime, secure management and use of data, making systems more resilient and understanding and monitoring systems and networks. With more than 90% of large corporates reporting security breaches this is a timely proposal that cannot be realised without the generous support and flexibility offered by a Future Leaders Fellowships which fosters both technical and academic skills development but also identifies the importance of cross-disciplinary approaches and allows for me and my team to gain experience through training and interacting with other sectors such as the healthcare, finance and stake holders.

## 7 Research and Innovation Environment

I recently joined the Cyber Security Group (CSG), School of Electronic and Computer Science (ECS) at UoS to establish my research group where I am leading a new strategic area of HPC Infrastructure Security. I have already started establishing the foundation which is instrumental for realising this innovative project. I have secured initial support from the ECS to establish the first Cyber Security Lab in UoS equipped with new servers and hardware accelerated devices worth £12,000. In addition, Mellanox Technologies has provided hardware accelerated devices BlueField SmartNIC. This equipment will be used in WP2-WP4. CSG members have complementary skills to mine and will contribute to the realisation of this innovative project. Prof. Sassone and Prof. Buttler have recently been awarded ESPRC funding for "Security-by-Design". In addition to their expertise in machine learning, blockchain and formal verification. Prof. Hu who is a leader in efficient scheduling, economic scheduling will act as my personal mentor to enhance my career progression. UoS has multiple research council funded Doctoral Training centres and I will be supported by the school's internal funds to recruit two PhD students in year 1 and 3 of the project. I am part of the UoS Academic Centre of Excellence in Cyber Security Research, which brings together cyber-security related research across a broad academic spectrum, as well as the UoS Cyber Security Academy. I am also leading the vision of ACE-CSE to recruit and train innovators and graduate students working at the interface of academia and industry and gain relevant skills for innovative hardware security. UoS is a member of the SPRITE+ Hub on challenges in security, privacy, identity and trust in the digital economy. A particular strength relevant to my vision is the university's strong enterprise agenda and excellent relationships with business and industry which will be instrumental for the dissemination of the research outcome and its realisation in a timely manner during the second phase of this fellowship (years 5-7).

## 8 Intellectual property

In the case of an IP generated from this project, all the related licensing will be handled by the Research and Innovation Services at UoS and Licensing office with dedicated contract and IP mangers. UoS has a dedicated business unit to realise the full benefits of any funded research and to support staff to accelerate the validation of research outcome. They have a network of companies and industry specialists and investors, provide additional prime funding to drive any IP to a viable development stage. The appropriate contracts and agreements with any commercial/ business partner involved will be drafted and put in place by the relevant contracts' teams at UoS to ensure the research outcome is shared with the community.

## References


[1] Amd secure encrypted virtualization. "https://developer.amd.com/sev/".

[2] Ibm secure execution. https://www.ibm.com.

[3] S. Arnautov, B. Trach, F. Gregor, T. Knauth, A. Martin, C. Priebe, J. Lind, D. Muthukumaran, D. O'keeffe, M. L. Stillwell, et al. SCONE: Secure linux containers with Intel SGX. In *12th USENIX Symposium on Operating*

*Systems Design and Implementation (OSDI 16)*, pages 689–703, 2016.

[4] A. Atamli. *Partitioning the Trusted Computing Baseof Applications on Commodity Systems*. PhD thesis, University of Oxford, 2017.

[5] A. Atamli, G. Petracca, and J. Crowcroft. IO-Trust: An out-of-band trusted memory acquisition for intrusion detection and forensics investigations in cloud IOMMU based systems. In *Proceedings of the 14th International Conference on Availability, Reliability and Security*, page 45. ACM, 2019.

[6] J. Brodkin. Gartner: Seven cloud-computing security risks. *Infoworld*, 2008:1–3, 2008.

[7] P. Mell, T. Grance, et al. The NIST definition of cloud computing. 2011.