This report is a summary of proceedings of a meeting organised by The Royal Academy of Engineering in December 2011. The meeting was attended by Fellows, representatives of industry, government and other relevant organisations. The report reflects the discussions that took place at the meeting but it should be noted that, while the conclusions and recommendations reflect the majority opinion, they do not necessarily represent the policies of the Academy or the other organisations involved.
The future of computing: indispensable or unsustainable?
Contents

1 Introduction from the chair 5
2 The future of ICT 6
3 The future of ICT in the work environment 9
4 The future of ICT in government 12
5 The future of data centres 15
6 The future of critical metals for ICT 19
7 The future of economic development using computing 23
8 Discussion 26
Introduction from the chair

For almost a decade, my interest has been in the crossover between computing and sustainability. Over that time, issues around sustainable computing have gained prominence within the ICT industry, academia and wider society. However, a show of interest does not mean anything has really changed. The motivation for this meeting was to try to get a feel for what the current situation is from various different sets of perspectives and different sectors. The summary of our meeting, as set out in this publication, will form the basis of further policy discussion both within and outside of the Academy.

“... I have no doubt that some of our predictions today will be totally wrong, but we have to be optimistic when we see what has happened already. This meeting has illustrated this very well.”

Sir Anthony Cleaver, meeting participant

Professor Andy Hopper FRS FREng, Head of the Computer Laboratory, University of Cambridge
The future of ICT

Richard Lanyon-Hogg, IBM, Chief Technology Officer for Energy Efficiency

Six years ago, IBM started working with the Department for Environment, Food and Rural Affairs (Defra) on issues around ICT and sustainability. We set about answering two examination questions:

• How to determine the energy and carbon footprint of an ICT system?
• How to determine whether one end-user system is more energy efficient than another?

That work has shaped my thinking as to how I see ICT unfolding over the next 10 to 20 years. Three areas of importance will be devices, behaviour of people and ICT infrastructure.

2.1 Devices

As of 31 October 2011, the world’s population, hit 7 billion. As the population increases so too do the number of devices we have, both permanently and intermittently connected to the internet. The future of these devices will be shaped by several factors: heat, materials and development cost versus revenue.

• **Heat:** As processors have become more complex and performed more work, they have generated more heat. Back in the 1960s, water cooling was introduced to tackle this issue; in the mid-1980s, bipolar transistors were introduced initially cooling the system but ultimately leading to as much heat as a domestic steam iron being emitted (5 watts/cm²). Bipolar gave way to CMOS transistors, which required less power and so did not become as hot; but as processors developed, heat problems soon returned, with device surface temperatures climbing beyond 10 watts/cm². As we look to the future, clearly advances beyond

“When I started out, I worked for a company called Walter Somers Forgemasters. I walked into their data centre suite and there was an ICL 2903 mainframe. Bliss – they came in orange in those days.”

Richard Lanyon-Hogg
‘lower power’ processors are required. We need much more innovation, particularly with new transistor designs such as FinFETs and 3D.

• **Materials:** The increases in processor power have been achieved thanks to developments in the numbers of chemical elements that could be used alongside silicon-based technologies. Before the 1990s, six major chemical elements were employed in silicon technology; through the 1990s and up to 2005 a further eight were deployed; recently this number has increased to a total of 53. Many of the chemical elements are no longer easily available, or accessible. When you reflect upon the acceleration of the global population and the projected growth in devices, the question of the sustainability of the current technology and business models becomes ever sharper.

• **Development cost versus revenue:** As far back as the mid-1950s, the annual growth rate of R&D expenditure was nearly twice that of the revenue it generated; slowly but surely, these two values have begun to merge. This has led to an acceleration of knowledge-sharing and manufacturing. Over the next 10 to 20 years, my feeling is there has to be a further sharing of technology and a consolidation in the industry if companies want to continue to remain profitable. The current model is unsustainable.

### 2.2 The behaviour of people

Over the next 10 to 20 years, the ICT industry as a whole will have to become more inclusive in its thinking. At the end of the decade, one third of the world’s population will be over the age of 50. We mustn’t forget that, currently, two thirds of the population in developed countries struggle to use existing technology such as mobile phones and the internet. The industry has no other choice but to reflect these challenges in the new products and services it wishes (and I believe is required) to offer.

*“Thank heavens for man’s innovation and ingenuity! Could you imagine a mobile phone made of discrete transistors or valves? You would need a shopping trolley to wheel it around in.”*  
Richard Lanyon-Hogg
2.3 ICT infrastructure

Besides resource and demographic challenges, a large and looming threat to today’s current infrastructure is data. The daily volumes of data are now being calculated at 15 petabytes. To put that into context, the British Library has many millions of items and it is calculated as holding about 3 petabytes of data. So, daily, we are now generating five times what is in the British Library. The generation of such vast amounts of data heralds the era of big data; we have no other choice but to use analytics to mine data to yield information enabling individuals, communities, people in the public and private sector to make more informed and meaningful decisions.

2.4 Conclusion

If we were standing here in 2031, what might it be like? It is quite conceivable that some of the world of ICT could actually be in us. We have lived through paper tape, cards, keyboards and touch. Potentially, however, we could be talking about biological interfaces into systems in 20 years’ time. In terms of programming paradigms engineers are now talking about non-Von Neumann-based architectures, systems that will learn from the environments in which they are placed. Perhaps if we come back in 2031, by which time I’ll be even greyer and older, ICT will slowly start to become invisible, and it could even begin to become a little cognitive.
“The chaos monkey is an application that runs on Netflix’s production infrastructure and it goes round randomly shutting things down – in a live environment. They have built to defend against the chaos monkey so remain running when there are power outages.”

Joe Baguley

3 The future of ICT in the work environment

Joe Baguley, Chief Cloud Technologist, VMware

In the early 1990s when I started my career in ICT, everything that I had that was cool in technology was being given to me by my employers. They had control over my devices and my access. However, that has now been completely turned on its head, and it is not devices that are the consumerised element, it is users that have become consumerised. Users now have consumer attitudes and that is what ICT departments and the industry have to adjust to.

It is no longer the case that users have one log-on to their company; they have multiple log-ons: Gmail; Dropbox; Facebook, where people lie to their friends; Twitter, where people are honest with strangers; banks, retailers, the list goes on. My company is just something that provides me with an email system in addition to the email systems I already have. We are now in the post-PC era; people are no longer limited to a desktop accessing data from a server. They want applications that are running on the cloud that can be accessed by multiple devices with different operating systems (OS). Because of all this, the way that enterprises are developing and deploying applications is changing dramatically.
3.1 Applications will become more device-, OS- and browser-neutral

If we go from 1996 to 2020 (Figure 1), the mid-blue area below shows the percentage of Windows-specific applications being developed. The light blue area in the middle is browser-specific applications, usually Internet Explorer 6. The dark blue area above shows people developing applications that run on any device. This is happening because developers have realised that the consumption model is changing.

3.2 The end of the ‘Corporate LAN’

Historically, companies built infrastructure because they had to. This has left most organisations with huge, insecure and very expensive infrastructure. Applications were developed in a flawed way because security was never considered; it was assumed that everything would sit...
behind a firewall. If we, as human beings, had evolved in the same way, we would all be walking around in plastic bubbles with ‘firewall’ written on them, with no immune system whatsoever. As soon as someone popped that bubble and put a virus in there, you would die. That is how corporate ICT is done.

However, you can think differently and shrink all of your delivery down to a central collection of services and then deliver those to the internet. To do this, you should assume that every device that connects with them is compromised from a security perspective. That kind of security thinking is what is happening now in companies and that is the evolution we are seeing: organisations are starting to give up control.

3.3 Platforms to build services

Cloud computing is now giving you the choice to get the infrastructure from somewhere else as a service. Virtualisation is turning the server into a commodity such that, if a physical piece of hardware disappears at any one time, there is no problem and the service continues. Resilience has been built into the systems that deliver the service.

3.4 Virtual machines

The platform is now a service; static machines plugged directly into the system are no longer needed. Our electronic devices, smartphones, tablets, have become ‘virtual machines' and because they can plug straight into systems they are now disposable, consumable items. Software built as a service leads towards consumption. So cloud, simply, is us moving up the stack of abstraction, moving up the stack of resilience and redundancy and up the stack of commoditisation and componentisation.
What is green ICT? It is no longer just concerned with power efficiency and ‘switch off’ campaigns, it is now moving to grapple with the environmental impact of the ICT life-cycle, from cradle to grave, including: how materials are mined, produced and presented; the conditions and energy for manufacture and development; the delivery of ICT to consumers; energy and consumables in operation; the re-use of kit and its eventual decommissioning and disposal at the end of life.

Why do we need green ICT? The most important issue to understand is that we are facing an acceleration in the use and demand for electronic products. In the UK there are already 10 million office PCs and nearly 50% of the adult population use PCs at work; this is set to rise to 70% by 2020. Currently ICT consumes 15% of office power and this is predicted to rise to 30% by 2020. Domestically, it is predicted that 45% of power will be consumed by electronic products by 2020.  

4.1 UK government ICT strategy

The UK public sector spends £14 billion on ICT with significant energy and carbon costs; how are we trying to respond to this? A cross-departmental group has now published the UK Greening Government ICT strategy, its vision being A cost effective and energy efficient ICT estate, which is fully exploited to enable new, more sustainable and efficient ways of working for staff organisations and customers (that’s you and me!)

As part of the Green Strategy, a roadmap has been produced, highlighting 14 key areas for improvement. These include the adoption of Government Buying Standards, which cover principles that seek to minimise
the environmental impacts of new hardware across its lifecycle; better print management; virtualising and sharing applications; energy-efficient operation of data centres; and, reporting on ICT disposals to be sure old equipment does not end up in landfill. Another direction being taken with the Green Strategy is to tackle the supply chain so that we can better understand the footprint of all ICT services (in-house or out-sourced) used in government work.

4.2 Using ICT to reduce environmental impacts in other areas of government

The government has proposed to cut 25% of greenhouse gas emissions across our estate and our operations by 2015. In tackling these challenges ICT has a key role to play. In return for the 10%+ of an organisation’s emissions, there is the opportunity to use ICT to tackle the remaining 90% of an organisation’s footprint by providing us with the means to be more efficient in delivering services.

4.3 Developing the green economy

We have developed six key principles for the optimal path in the transition to the green economy:

• providing policy certainty and simpler regulation
• choosing the right mix of policy instruments
• making it ‘easier to be green’
• incentivising R&D and innovation
• promoting business and economic resilience
• working through EU and international channels
4.4 Technology change?

Finally, I have some questions. Are we really running to stand still? Will the actual energy and resource demand we are creating in our pursuit of technology innovation and exploitation exceed our ability to produce it and our planet’s ability to sustain it? Will we ultimately reach a level of miniaturisation and energy efficiency beyond which we cannot go? Then there is the oft-quoted Jevons’ law (see Box 1) that consumption of a resource expands to fill the capacity released.

**Box 1: Jevons’ law and demand elasticity**

In 1865, William Stanley Jevons recognised that, despite more efficient steam engines being invented, our consumption of coal went up. This was because steam engine use became more cost-effective, leading to its increased use. With this he had identified one of the underlying problems in modern economics, demand elasticity: if we reduce the marginal cost of a unit of production, we are very likely to drive up demand so fast that we drive up our overall consumption of the resource.
5 The future of data centres

Liam Newcombe, Chief Technology Officer, Romonet

Data centres are a growing opportunity market. Since 2007, there has 56% growth in energy consumption. Data centres in Europe use the same amount of electricity as Portugal – a country. In the US and the Pacific Rim, they use more. A single, managed server may be rented for about £2,000 a year, which might lead you to believe that it is low impact. But putting that server into a typical corporate data centre and running it for a year will lead to more CO$_2$ emissions than if I drove a modern four-by-four around the Equator$^1$.

5.1 Data centre demand

Within the ICT industry, demand for data centres is now increasing at a spectacular rate for multiple reasons:

- rise in demand for traditional data centre services by corporations
- displacement of physical process onto ICT systems
- new retail services
- social networking, gaming and entertainment
- increase in the use of cloud services as more people start to use ‘thin client’ or ‘virtual desktop infrastructure’

Data centres are always on, so we are displacing activity from a pattern of variable energy consumption into potentially higher fixed energy consumption. For example, my laptop is sitting in my bag, using no energy. However, it is connected to a virtual desktop machine sitting in my office. That machine plus the mobile network is probably still consuming 80% of the power I would be using while actually interacting with it.

$^1$ Speaker’s own calculation
“We have made data centres more efficient… Nobody has said, ‘Great, I’ve got half the power bill!’ What they say is, ‘Great – I can deploy more servers.’”

Liam Newcombe

5.2 Data centre efficiency

Power bills, as well as threats from various regulators, woke up the ICT industry to their inefficiency, and a league table in the Carbon Reduction Commitment, which is the UK’s mandatory climate change and energy saving scheme. We now have Energy Star ratings for servers, the SPECPower benchmarks and the EU Code of Conduct for Data Centre Energy Efficiency. The industry understands and competes on efficiency, but it does not really care about carbon yet. It is not yet on the radar because it doesn’t cost (enough) money.

Previously, from the power station to the processor, 0.3% of source energy was actually being used to do processing work in a data centre; the rest was lost (Figure 2). New data centres use about 2.5% of our source energy (Figure 3). This is a huge increase from the previous value of 0.3%, but there are still opportunities for improvement.

![Diagram showing where energy is lost between source and processor in a 2007 vintage data centre (image courtesy of Liam Newcombe)](image-url)
5.3 Things need to change

- **Fixed-energy consumption**: Many data centres do nothing for much of the time, but use 95% of the energy that they use when they are running at full capacity; this is a vast fixed-energy cost and the major opportunity for improvement.

- **Transparency**: If companies were transparent, internally and externally, about how much energy they have consumed to deliver a service, market forces could realistically drive competition on efficiency.

- **Embedded energy**: The industry needs to start calculating the embedded energy that has gone into creating their products. Many people are now replacing their servers on an 18-month cycle. Although the new servers may be more efficient, there is a capital cost from an environmental perspective (the embedded energy), which must also be weighed against the operational cost of the energy consumption in use.
• **Demand elasticity:** Data centres have become more efficient, yet their power consumption has continued to rise as Jevons’ law predicts. If the policy goal is to constrain energy consumption, then demand growth management is far more likely to succeed. On the other side of the argument, the use of ICT can reduce energy consumption in other areas, frequently with a large net saving. In this case, we should be encouraging demand growth, but only for those areas where a real chance of reduction exists.
The future of critical metals for ICT

Andrew Bloodworth, Head of Science, Minerals & Waste, British Geological Survey

We live in a world where, as we say in the business, if you cannot grow it, you have to mine it. We depend on minerals for all sorts of things. Just about everything we touch and do on a daily basis has a minerals component to it, including what goes into digital technology and communications. Major increases in demand for particular metals have been forecast: should we be worried?

6.1 Physical depletion

How do we manage our resources, as a species? In two ways: we can manage supply and we can manage demand. Supply is either from the Earth or it is from what we are starting to call the ‘urban mine’, which is the metals in buildings and devices in the anthropogenic environment. Demand is rising rapidly but can be managed in part by substituting a scarce resource or by learning how to do more with less.

Going right back to the days of Malthus, people have worried that we are going to run out of things, but what Malthus and Jevons did not consider was the power of science and technology to continue to deliver new ways of doing things. Advances in exploration and extraction technologies, underpinned by scientific research, mean that physical exhaustion of metals from the Earth’s crust is unlikely any time soon. However, access to these Earth resources may be restricted by a range of human and environmental factors.

As demand increases, the ‘urban mine’ will also become critical in supplying metals, although the availability and flow of these metals are notoriously difficult to measure. Access to materials in our urban mine is governed by how long products last and how easy it is to extract metals.

“Of all the platinum group metals ever mined as a species, we have mined 75% of them since the 1980s”
Andrew Bloodworth
from them. Currently, whether a metal is recycled is heavily influenced by its price. However, recycling can only achieve so much. If consumption rates continue to rise as predicted, we will always need to extract more.

We can try to decrease the demand of scarce metals by replacing them with less critical material. The scope of that substitution is limited by the fundamental properties of the material you are trying to substitute. For example, finding an acceptable substitute for the very high field strengths imparted to magnets by some rare earth elements is very difficult. The other option is to dematerialise – doing more with less. Again, metal prices are very powerful drivers for innovation in these areas.

### 6.2 Human factors

Human factors can threaten supply:

- **Resource nationalism**: Governments of resource-rich countries may try to boost their revenues by gaining control of those resources, which can include nationalising them.

- **Geopolitics**: Economies of scale mean the production of metals is concentrated into a few large mines and/or countries. That concentration of supply tends to increase the tension between those who have, like China, and those who do not, like Europe or Japan.

- **Conflict**: Large revenues from mining flowing into resource-rich, economically-poor countries can lead to corruption and conflict.

- **Skills shortage**: Looking into the future, there will be a big demand for new mines and new extraction operations, but there will not be the workforce with the technical knowledge to build and run them.

- **Business models**: Many metals have not been used much before and, although they are critically important, they are not used in massive volumes. This

---

“As geologists, when we have found dispersed materials in the crust, we concentrate them together and sell them – they are then used in electronics and are dispersed again. That is a very open loop.”

Andrew Bloodworth
means that they tend not to be a priority for large mining companies because limited production of these materials does not fit into their business model. Also, getting these metals from discovery to mine to market can take a long time – up to a decade.

6.3 Environmental limits

Extraction is a very energy-intensive process. As lower grade materials are recovered, more energy is required to separate them out and processing requires even more. Can we afford the carbon cost of recovering lower and lower grades from both primary materials and also from recycled material? Decarbonising our resource use presents a major challenge to the industry.

6.4 Scarcity or scare story?

Both the Earth resources and the urban mine together represent an abundant resource: there is material out there, but accessing that material is a big challenge. Many of the metals that ICT devices now use are at risk of supply disruption, mainly in the short- and medium-term, from human factors such as resource nationalism, geopolitics or skills shortages.

What do we need to do?

• Invest in the science of finding and separating materials.
• Develop better resource diplomacy.
• Consider ethically sourcing materials.
• Diversify our supply.
• Develop good dialogue down the supply chain to ensure the correct materials are being used for different components.
Reduce dispersion. This is a really important message for the ICT industry: recycling rates have to be increased. Products should be designed for recycling and reuse. There are also social barriers to overcome in terms of getting people to recycle items rather than keeping them as spares.

In the long term, the biggest challenge for the mining industry is, how do we decarbonise all these processes? That is a major challenge which we are nowhere near solving.
The future of economic development using computing

Dr Andrew Rice, Lecturer in the University of Cambridge Computer Laboratory

During this meeting, the point has been made several times that it is very difficult to calculate the environmental cost of computer processes. But what are the environmental benefits? We have looked at general ways in which we might make computing as applicable as possible; if we can do more things with computing, we can amortise the cost further.

7.1 The UN development goals

The UN development goals are internationally agreed, and include access to drinking water; gender equality in education and dealing with HIV and AIDS. The UN Development Report published in 2011 discussed the importance of dealing with these human development issues in tandem with how we deal with sustainability issues. I will speculate about how computing can perhaps contribute to some of these questions.

7.2 E-lancing

E-lancing is a term used to describe people who are contracted to do work electronically rather than physically. There are a number of start-ups in Silicon Valley acting as brokers for this kind of work. You can post your job online to a selection of people who will do the work and then they will come along and take on those contracts for you. Amazon’s Mechanical Turk lets you outsource very small units of work. Individuals browse through the jobs, choose the ones they want to do and do them.

E-lancing sounds like it could be a digital sweatshop, with people working at minimum possible cost to try to get a job done. In fact, the demographics of Amazon’s Mechanical Turk show a large percentage of workers being

“… I do a Google search and it is instrumental in me developing a new source of green energy. Rather than thinking about the cost and the benefit of particular things like that… instead look at what computing gets for us.”

Dr Andrew Rice
parents in the United States who look after their children and are also trying to make a little extra money at the same time. Previously subscribers to Amazon Mechanical Turk were paid in US dollars or in Amazon vouchers. A few years ago, Amazon addressed this and also started allowing payment in rupees. So now, if you look at the stats, the breakdown of workers on Mechanical Turk is about 50:50 from the US and India. Allowing payment in rupees allowed many more workers to enter.

One of the things that e-lancing can do is to deliver wealth straight to people – depending, of course, on the middle man and how much of a cut they take. This is one of the key tenets of the UN development goals because often, when you push aid into a country, it is very hard to get that aid to the people on the ground. In the next 10 years, as smartphones start to become available in areas such as Africa, how might the available workforce expand further?
7.3 Participatory sensing

Participatory sensing, is an idea that allows a community of observers to collect sensor information or observations and send them into a central point. This kind of information could be used to support the development goals by asking questions such as, ‘What is the status of your building? How far did you walk for water today?’ A successful participatory sensing platform would need to be able to manage data integrity, while a preformed set of people are contributing the values to it. A payment option could also be used to provide further incentives. A system like this could also be put in place to contextualise the readings of sensors.

7.4 Informing future designs

How do we try to maximise the utility of platforms now, so that in the future people can e-lance and take part in participatory sensing schemes, earn money, and have that money delivered straight to them? First, if you are trying to shape a mobile phone platform, you need to understand that platform and how people use that platform. Currently, the uncertainty about how people use their smartphones is at least as large as the uncertainty about how much energy is used in making them. Our project ‘Device Analyzer’ has been designed to analyse how users use their smartphones and is available free to download on the Android marketplace. The information collected may inform the design of this next generation of handsets.
8 Discussion

Below, is a summary of challenges for the future that arose from the discussions that followed each presentation.

8.1 Challenges for the future

• **Make simple changes:** there are basic things which companies could be doing to save energy but they are not (BCS publication, Greening your workspace, www.bcs.org/category/10552).

• **Enable users:** people want to use ICT efficiently; they just don’t know how to.

• **Incentivise change:** give someone a cash sum of 10% of any energy-saving idea they had.

• **Calculate embedded energy:** without that information, consumers cannot compare the efficiency of different service providers.

• **Change the philosophy of ICT device designers and engineers:** everything doesn’t need to be on all of the time.

• **Design for the future:** devices should be designed to be reused and recycled. We don’t want to repeat the mistakes of the past with lots of hardware ending up in landfill.

• **Filter and classify data for storage:** one of the biggest challenges that we are facing with big data is how to access it all efficiently.

• **Update regulation:** what are the data protection issues and what are the guardianship issues we have to address as more and more people start using cloud facilities.

“It is interesting that the hero of the day is obviously Jevons, who has been mentioned by at least three of our speakers.”

Sir Anthony Cleaver, meeting participant
“Will there be a point when the system is mature enough to start to ration how much energy or bandwidth it uses?”
meeting participant

“I like to think that we will just carry on riding Jevons’ law and providing ever more capacity until we start to reach a point where there will be no need to add more.”
Dr Andrew Rice

“… what humans do is they find new ways to use old stuff and new ways to use new stuff, which is the challenge … perhaps there will come a point when we say that we have pretty much found out the best way we can consume all of this. But I am not sure about that.”
Joe Baguley
The Royal Academy of Engineering

As the UK’s national academy for engineering, we bring together the most successful and talented engineers from across the engineering sectors for a shared purpose: to advance and promote excellence in engineering. We provide analysis and policy support to promote the UK’s role as a great place from which to do business. We take a lead on engineering education and we invest in the UK’s world class research base to underpin innovation. We work to improve public awareness and understanding of engineering. We are a national academy with a global outlook and use our international partnerships to ensure that the UK benefits from international networks, expertise and investment.

The Academy’s work programmes are driven by four strategic challenges, each of which provides a key contribution to a strong and vibrant engineering sector and to the health and wealth of society.

Drive faster and more balanced economic growth
The strategic challenge is to improve the capacity of UK entrepreneurs and enterprises to create innovative products and services, increase wealth and employment and rebalance the economy in favour of productive industry.

Lead the profession
The strategic challenge is to harness the collective expertise, energy and capacity of the engineering profession to enhance the UK’s economic and social development.

Foster better education and skills
The strategic challenge is to create a system of engineering education and training that satisfies the aspirations of young people while delivering the high calibre engineers and technicians that businesses need.

Promote engineering at the heart of society
The strategic challenge is to improve public understanding of engineering, increase awareness of how engineering impacts on lives and increase public recognition for our most talented engineers.