Re:AKT

December 2002 www.aktors.org/reakt

From the Director

Prof. Nigel Shadbolt

WELCOME to *Re:AKT*. I hope you find this and future newsletters interesting and useful. We want to communicate our excitement about the AKT project – to describe its goals and ambitions, results and achievements, the challenges and opportunities that lie ahead.

AKT represents a new way of funding research in the UK. It is the result of a national competition – five projects were funded out of a field of some 120 bids. Selected projects had to demonstrate that they were world-class, innovative, important and timely. They had to be multi-site, interdisciplinary and collaborative. In return AKT received six years of substantial, unfettered funding.

Why is AKT important? We all recognise that we are drowning in a deluge of data. We are suffering from a new kind of technological pollution – infosmog – too much information to make timely or effective decisions.

Knowledge

AKT is all about getting the right content to the right place, at the right time and in the right form. Content managed in this way is practical knowledge.

AKT is about making this vision of knowledge management a reality. It is about understanding why this is hard. It is about integrating where we can, and inventing where we must, the methods and techniques to realise the vision.

AKT is using the World Wide Web as its primary source of content. Tim Berners-Lee, the web's inventor, outlined his vision for a Semantic Web close to AKT's original aspirations. This newsletter articulates the assumptions behind the Semantic Web. AKT is tackling one of the great challenges of the 21st Century – realising the potential of the extraordinary information repository that humankind is building.

■ **To find out how you or your organisation might become AKTively involved, please contact us at info@aktors.org.**

Helping the fight against cancer

IN an average year in the United Kingdom, 118 women out of every 100,000 will be diagnosed with breast cancer.

The Government's breast screening programme has been effective, but it is massive – in 1999–2000, almost 1.5 million women were screened.

This produces a huge quantity of information; the difficulty is how to process it quickly and accurately.

Expertise

AKT has joined with another British IRC (Medical Imaging and Signals) to help tackle this problem, setting up project MIAKT (Medical Imaging and Knowledge Technologies).

Currently, breast screening pools the expertise of clinical, mammographic and histopathological experts and MIAKT will support and record this complex process.

Researchers from the two groups met at Guy's Hospital on May Day 2002, with teams from Sheffield, Southampton, the Open University (OU), Oxford University and King's College informally exchanging ideas.

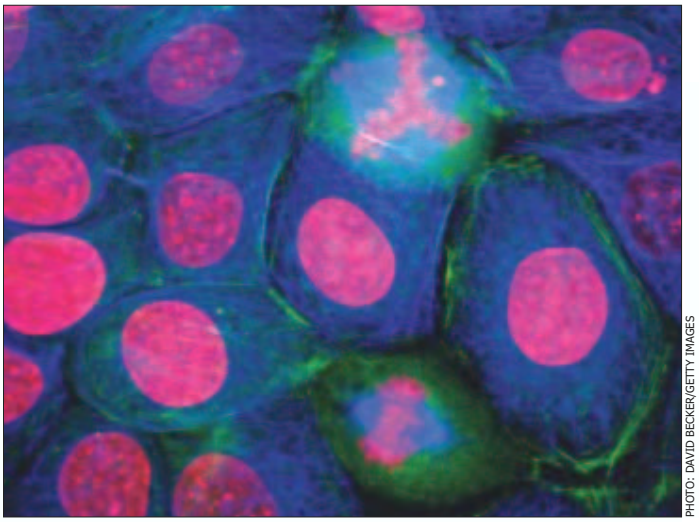


PHOTO: DAVID BECKER/GETTY IMAGES

Danger signs: Cancer cells from human breast.

Challenges MIAKT hopes to meet include developing ontologies and process models that can show the complex diagnostic processes, and adapting knowledge technologies for what is primarily a visual task involving processing images.

Empowering meeting attenders near and far

FACE-TO-FACE meetings offer many challenges to both attenders and organisers: everyone knows how hard it is keep meetings in focus and on time.

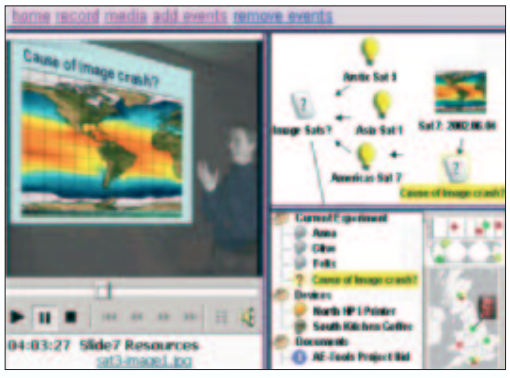
The challenges escalate considerably when remote participation is involved, particularly with scientists of international stature who may have detailed technical points to convey, debate and possibly reanalyse with multiple participants. Facilitating such meetings is the goal of CoAKTinG (Collaborative Advanced Knowledge Technologies in the Grid), a sub-project of AKT funded by the EPSRC as part of the UK's e-Science Initiative on Grid computing.

Collaboration

CoAKTinG will provide tools to assist scientific collaboration by integrating 'intelligent meeting spaces' (which help understand when key events happen at a meeting) with techniques for taking live video and audio streams from internet-hosted meetings.

Southampton's HyStream tools can annotate the video/audio with semantic tags to assist retrieval and analysis later on. The Open

■ *Continued on back page*



CoAKTinG scenario: This is how it might look to replay the key segment from a missed meeting, following a BuddySpace alert and schematic Compendium network overview of the ebb and flow of the discussion.

Next-generation

Agents on the web

ANYONE who's ever arrived in an unfamiliar city and wished for a personal guide to help them make the most of their visit will appreciate the vision behind the Agentcities project.

Funded by the EU 5th Framework, Agentcities is aiming to deploy intelligent network services capable of advising mobile users about an environment in which they find themselves.

One Agentcities scenario envisages planning an evening's entertainment for a traveller, selecting dining, entertainment and accommodation according to their preferences.

Interface

Agentcities is funding small-scale projects to deploy existing services on the emerging Agentcities network. In the second half of 2002, the Aberdeen AKT team has been working to interface selected AKTbus components to the Agentcities infrastructure.

This work will allow AKT knowledge services to operate as software agents and serve Agentcities users. In conjunction with the Aberdeen software agents group, a demonstrator is being developed that offers entertainment options to people visiting Scotland's Granite City.

THE World Wide Web is like a large library. But searching through it is unintelligent and often painful. As it gets larger, the dangers of infosmog will increase.

The Next-generation Semantic Web is intended to get around these problems by developing ways of sifting intelligently through information. By providing different layers of representation for content, the Semantic Web allows increasingly sophisticated and more powerful inferences to be made about the content, as the details are abstracted from the raw material.

For instance, in our figure (*opposite*), we see an actual news story as it appears in the Open University's *KMi Planet* e-journal, with a headline 'Outstanding contribution award for Daisy Mwanza'. Here, the content has been 'rendered' – generated on the fly – for display on the web. The layer above shows the HTML encoding, which gives the 'look' of the page behind the scenes. For instance, we can see that what makes the headline the headline is the pair of tags, **<H1>** and **</H1>**, which show where it begins and ends.

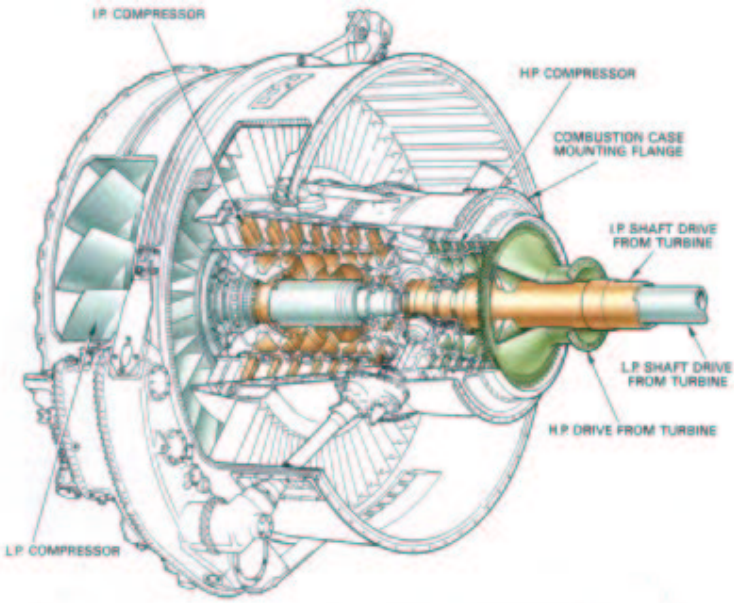
Abstract

The next layer above is the extended mark-up language XML, which allows a more abstract – and hence more powerful – representation. Now the headline is delimited by the tags **<Headline>** and **</Headline>**. These tags are arbitrary and their only meaning is by mutual agreement by specific communities of users. So the people behind KMi Planet have decided that **<Headline>** should be interpreted as **<H1>**. But mobile phone users, for example, may interpret the **<Headline>** tag differently to allow it to appear in a very small font. Hence XML allows what is called repurposing of content.

So far so good. But we want intelligent look-up. If we move up another layer, we can now see some inferential capability, via the **Resource Description Framework** (RDF). RDF represents relations as triples, two objects and a relation between them. Both objects and the relation are actually locations on the web.

This is a simple but amazingly powerful idea, allowing much more complex searches than are possible on the World Wide Web, enabling principled, large scale indexing and retrieval by machines.

But it is not yet intelligent. The top layer of representation shows a language called **OWL** (Web Ontology Language). This, and languages like it, provide the capability to reason about the relations stored in RDF triples, thus going



Better by design

AKT is strongly concerned with industrial relevance. As one example of this, we are working closely with Rolls-Royce to provide support for its aero engine designers. In collaboration with the Transmissions and Structures team at Derby, work is under way on developing a Designers' Workbench software system.

The current prototype allows designers to interact with a familiar engineering drawing interface. As their design progresses, the Workbench checks that the new parameters are in accordance with a variety of pre-existing constraints, using constraint-satisfaction technology developed by the Aberdeen AKT group.

Workbench users can also access the rationales for these constraints and should they wish to override them, they can provide their own rationale for doing so. In this way, the Workbench builds up a repository of design knowledge and experience.

The prototype Designers' Workbench will be trialled by Rolls-Royce early in 2003.

Going to town on technology

ALTHOUGH AKT has a remit from the EPSRC to think as broadly and as innovatively as possible, this does not mean we spend all our time in the 'blue sky'. Relevance to the real, information-related problems in industry and within organisations that follow from changes in technology is a priority.

As part of our commitment to industrial relevance, AKT has instituted a series of 'town meetings', in which our technologies can be seen in action, via presentations, demonstrations and informal chats – those 'coffee break' moments where the serious questions can be asked.

The first of these meetings was held on September 18, 2001, at the Institute of Materials in London. The AKT Manifesto was launched (**http:www.aktors.org/publications/Manifesto.doc**) and technology was demonstrated, including AKT-0 (an early integration of AKT technologies), tools and methods for enriching web-based knowledge content and language-processing technology.

An exciting day was enjoyed by all. The audience included representatives of a number of enterprises ranging from massive bluechips to the nimblest technology startups.

Opportunity

A second meeting is in order, to demonstrate the continuing success of AKT and to showcase the achievements of the last 12 months. It will provide an opportunity for individual organisations to discuss specific knowledge management challenges and how these relate to AKT technologies.

Places are limited for this one-day event, and are by invitation only. If you have an interest in attending, please email Susan Davies (AKT administrative co-ordinator) at **sdd@ecs.soton.ac.uk,** or telephone **023 8059 3523.**

http://www.aktors.org: it's no ordinary website

A WEBSITE is essential nowadays for scientific, technological or business communication and, in common with most contemporary collaborative projects, AKT has a website which can be found at the address above.

But aktors.org is no ordinary website – we have applied our own principles and technologies for the creation of the site.

The site contains information about AKT and its aims, associated events and the AKTors Club. There are links to the AKT Manifesto and to the individual websites of the consortium partners.

But site editor Hugh Glaser is most excited about the use of our own technology in the site. Much of our internal knowledge management is done via the project Intranet, but even on the public site we can glimpse the future – a website incorporating ontologies, authoring tools and instantly generated discussion environments.

The welcome page includes the latest story from *AKT Planet*, an intelligent online news server and our ontology-mediated e-print archive gives immediate access to the technical papers produced within the project.

There is also access to public discussion spaces of the papers, using the Digital Document Discourse Environment (D3E).

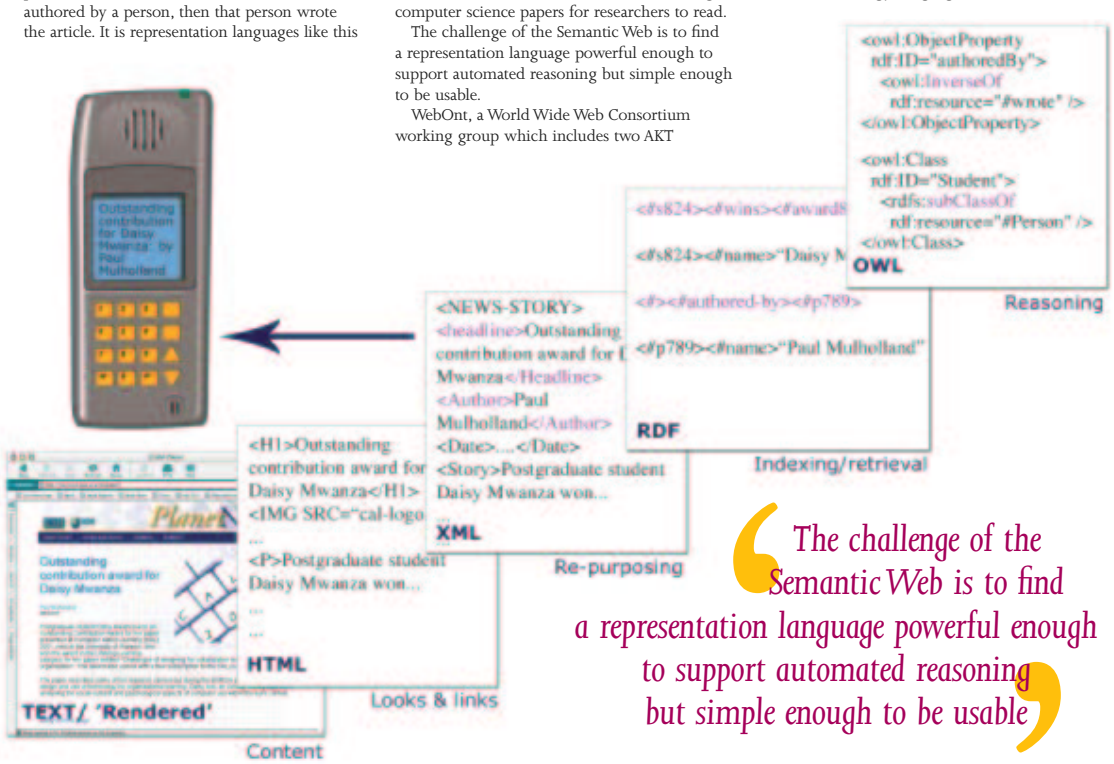
We are committed to providing the best technology for information dissemination. The beauty of it is that we can use it to tell the world about ourselves!

Semantic Web

beyond search and retrieval, giving the machine some capacity to understand the implications of particular relations – for instance if an article is authored by a person, then that person wrote the article. It is representation languages like this that will support the applications described in *Re:AKT*, such as planning an evening's entertainment in Aberdeen, or recommending computer science papers for researchers to read.

The challenge of the Semantic Web is to find a representation language powerful enough to support automated reasoning but simple enough to be usable.

WebOnt, a World Wide Web Consortium working group which includes two AKT members, is wrestling with this problem. Only one puzzle remains. If it is so intelligent, why is the Web Ontology Language not called WOL?



> *The challenge of the Semantic Web is to find a representation language powerful enough to support automated reasoning but simple enough to be usable*

Knowledge services

THERE is a huge quantity of information on the web – and our society is getting more and more information-hungry. As the Semantic Web allows increasingly intelligent search and retrieval of information, AKT can begin to provide the technology to satisfy information hunger from the immense resources on the web.

One of the most ambitious AKT projects is to set up an AKTive portal to provide knowledge services for British computer science. There is a huge amount of data about computer science sitting on the various departments' websites, waiting to be Hoovered up – about projects, staff, research interests, courses, prototypes, etc. This data can be used to produce a snapshot of the state of British computer science.

The AKTive portal is an information space organised by an ontology of the important concepts in computing research. Knowledge services which the portal may provide include: recommending relevant people or documents; supporting discussions of published or unpublished papers; issue tracking – the generation and maintenance of a 'to do' list; narrative or template generation for reports or presentations; collecting research intelligence – which projects are on the go in such-and-such a research area.

Once the research is proven, it can be used in fields other than computing.

The information to support all these services is already on the web. AKT's aim is to provide the technology to use this previously unexploited resource.

Amilcare: the key to automating annotation

THE Semantic Web is dependent on the annotation of documents with metadata – all the stuff about the document itself that is so useful for reasoning about it but so hard to collect.

Amilcare, developed by Dr. Fabio Ciravegna at Sheffield, provides the leading solution for partially automating the annotation process. It is an adaptive Information Extraction engine with the unique ability to handle the whole range of web pages from plain text to structured data.

Into these texts, it can insert annotations indicating the location of significant information, making it available for subsequent processing.

The software is designed with both novices and experienced professionals in mind, and permits a sophisticated Information Extraction system to be trained and deployed with minimal input from the user. Amilcare has already found wide acceptance as proved by its integration with a number of Semantic Web tools including CREAM, an annotating and authoring environment from the University of Karlsruhe.

Improving business workflow

MODERN organisations are virtual entities. People are distributed about the place, with different capabilities and responsibilities, all needing to interact and collaborate. Not only that, some of the agents in a business workflow are artificial!

These different agents must get together to achieve complex tasks requiring specialised – read *scarce and expensive!* – expertise and sophisticated technology.

Co-ordinating resources for optimal efficiency is a hard problem and demands much organisational effort.

Expertise

The AKT consortium has special expertise in this field. The Edinburgh team has a great deal of experience in Workflow and Business Process Modelling and is focusing on the problem of providing an open-architecture virtual workflow system that is distributed, knowledge-based and agent-based, whose ability can be enhanced by arbitrary intelligent agents where appropriate.

This approach integrates wonderfully with Aberdeen's work on problem-solving agents, and the KRAFT system, a constraint-based problem-solving agent, has been used to integrate with the distributed workflow system.

Other tools across the AKT consortium that might be integrated in this way include, for example, the OU and Southampton's hyperlink tool, process modellers in general, the AKTbus, Netmeeting, Edinburgh's I-X, and Aberdeen's KRAFT system and its inbuilt Constraint-Solver

The ultimate goal of the effort is to provide an environment and means where multiple tools, methodologies and disciplines can be brought to bear to improve organisational performance.

AKTors dominate EKAW conference

THE start of the university academic year saw AKTors dominate EKAW, the premier European conference on knowledge management and the Semantic Web, held in Sigüenza, Spain.

In the most visible sign of this, AKT's Director, Prof. Nigel Shadbolt, was a keynote speaker – together with Prof. Tom Mitchell, a pioneer of machine learning, and Prof. James Hendler, chair of the W3C working group (WebOnt) on the Semantic Web.

Tim Berners-Lee, the inventor of the web, insists that the Semantic Web – incorporating meaning into web pages (*see pages 2/3*) – will have a larger impact than the original web.

During his talk, Prof. Shadbolt presented examples of how AKT technologies were leading the way. For instance, Ontocopi automatically identifies emergent communities of practice within organisations and uses 3D visualisations to highlight significant clusters of collaborators.

Another indicator of AKT's impact at the conference was the number of AKT-related papers accepted. Seven papers were co-authored by AKTors and four of the five AKT sites were represented at the workshop. Also, two of the three top paper authors were AKTors – Dr Fabio Ciravegna and Dr John Domingue, each of whom co-wrote three papers. Indeed, the AKT consortium had more papers at the conference than any single nation.

Variety

The seven AKT papers covered a number of key AKT technologies in a wide variety of contexts. For example, Dr Ciravegna described how the Melita system would enable information extraction – automatically extracting items of interest from natural language text – to be carried out by non-experts. Melita achieves a usability improvement by semi-automating document markup – standard information extraction requires users to mark-up about 100 texts.

Dr Enrico Motta presented MnM, a web-based tool, which also helps annotate documents both automatically and semi-automatically.

Finally, Dr Srinandan Dasmahapatra presented work which automatically identifies when terms in web documents refer to the same entity.

Seven things everyone should know about the World Wide Web . . .

1 **Infinite content:** As far as the casual web user is concerned, it is in effect an infinitely large dynamic multimedia library.

2 **Community of authors:** Unlike a library, however, the proportion of readers who are also authors is very high.

3 **HTML = content layout + links:** The main web language is HTML (HyperText Markup Language), which 'tags' text to be displayed in different styles/fonts and enables you to click on 'links' to jump from one page to any other.

4 **Web pages on the fly:** For big websites, the pages you see are created 'on the fly'. Content is stored in a generic database, and actual web pages are displayed ('rendered') according to design templates, customised to the person reading them.

5 **XML = context-sensitive layout:** XML (eXtended Markup Language) goes beyond HTML by tagging abstract functions such as 'headline' or 'author'. These abstract tags allow the display to vary according to the viewing device (eg PC vs mobile phone).

6 **Unwieldy and not future-proof:** The web could collapse under its own weight. And we must remember that as old formats fall into disuse, older content will become harder to retrieve.

7 **Becoming proactive:** Literal-minded look-up is all very well, but the web should be more like a smart and helpful librarian, intelligently helping users by understanding what they mean. This is the challenge of the Next-generation Semantic Web, the challenge that AKT has taken up.

Empowering meeting attenders

■ *From page 1*

University will add some of its pioneering approaches – such as BuddySpace, a suite of enhanced presence management and instant messaging tools, and Compendium, which helps capture the information generated in a discussion in the form of a visual 'group memory' available to all stakeholders.

Key issues for discussion and action will be flagged to all participants using the University of Edinburgh's I-X Process Panels, which help monitor progress and delegate tasks across complex task structures.

Even those who missed a key discussion at a particular meeting will be able to relive it afterwards: a BuddySpace alert will take them straight to the relevant Compendium diagram, provide schematic highlights of the pros and cons of the debate, and let them jump directly to an 'instant replay' of the video and audio streams recorded and annotated at that point in the discussion!

AKTors Club

INDUSTRIAL relevance is a key factor in AKT. The problems of infosmog are first and foremost real problems within real organisations. Hence we see it as a key test of AKT that our technologies address these problems – and this is a self-imposed test, not one that our funding body, the EPSRC, has dictated.

Our original bid for funding was supported by a number of industrial firms and organisations that were interested in our approach and knew of our partners' reputations. These initial supporters included bluechip firms such as Unilever, BA, Rolls-Royce, Boeing and BP Amoco, and technology firms such as Parametric Technology, Epistemics, Teknowledge and Multicosm (now Active Navigation).

This nucleus of support has been expanded into the AKTors Club, a grouping of organisations that support our research. Club members can be 'players', in which case they display an active interest in AKT, and have a named representative who can track the development of AKT's technologies, and a monitored web page on the AKT website. A second level of membership is that of the 'audience', which contains those companies that have expressed an interest in AKT.

■ **The AKTors Club can be accessed via the website at http://www.aktors.org/club/. To join, please send an email request to club-request@aktors.org.**

Meet an AKTor

WHAT sort of people work for AKT? We went to Edinburgh to meet Jessica Chen-Burger, who has worked as a research fellow on AKT since January 2001.

Jessica took her first degree in Business Studies at SooChow University, Taiwan. She has two masters degrees – from Missouri and Edinburgh – and a Ph.D. in Artificial Intelligence from Edinburgh. She specialises in modelling business processes and before AKT worked on the US DARPA-funded initiative Air Operations Enterprise Modelling.

■ **What are you working on at the moment?**

My current work has two strands: one is knowledge management using business process modelling and workflow techniques; the other is part of an initiative – the AKT Map – for providing an intelligent knowledge portal for AKT capturing the research activities of AKT fellows, the people they collaborate with, the important literature and software systems that influence their work and the results of their research activities.

■ **And who are you collaborating with?**

In AKT, I collaborate mainly with Aberdeen, but through the AKT spin-off project, CoAKTinG, I work with the OU and Southampton. I've also really enjoyed getting to know researchers who are part of the wider AKT family.

■ **What are your future AKT plans?**

I hope to bring formality and automation to knowledge management where possible. I have already done some of this in the past and hope to pursue this line of work in the future.

■ **And what do you think AKT's legacy will be?**

Given the internet, obtaining almost any kind of information is no longer an issue. Instead, the most important problem today is to determine whether a piece of information is relevant and trustworthy.

With the diversified talents and commitments of AKT members, I am confident that AKT can shed new light on these important and urgent issues.

AKT in numbers

- **6 years**
- **5 universities**
- **£7.3 million**
- **55 researchers**