AICSCSHistory_Interview001_David_Willshaw

SPEAKERS
David Willshaw, Vassilis Galanos, Xiao Yang

Please note that this transcript has been edited in places relative to the original audio, in order to improve clarity and readability. The original audio is not made publicly available. The timings shown correspond to the original audio.

Vassilis Galanos
So this is Vassilis Galanos. This is the 2nd April 2024 and we have the pleasure to interview Professor David Willshaw for our project on the history of artificial intelligence and computer science at Edinburgh University. I'm accompanied by Xiao Yang, who's my colleague. So without any further ado, I will just start the interview. So before we delve into the specifics of your notable career and the evolution of AI and computer science at Edinburgh, could you please share an introduction about yourself, highlighting your major contributions to the field and your personal journey within AI and or Computer Science?

David Willshaw
Okay, well, thank you. I hope I'll be helpful. So let me tell you my journey, as it were, which gives a factual account of how I got here and why I'm here. So I'm primarily, in fact exclusively, interested in the brain. So I'm one of these people in Informatics who are not AI people or Computer Science people. I've been here on two occasions, early on in the late 1960s to early 1970s, and then, since 1988. So I was here at the start of the Artificial Intelligence activity here. I came up to Edinburgh to be a PhD student of Richard Gregory, Richard being one of the three founders of the Department of Machine Intelligence and Perception in 1967, who was a perceptual psychologist, a visual psychologist. And I worked for him for a short time and then I transferred to be a PhD student of Christopher Longuet-Higgins. Christopher was a notable theoretical chemist, who at the age of 40 decided he wanted to study the mind. And he came to Edinburgh with Richard to join Donald Michie. Together they had a large grant, which enabled the Department of Machine Intelligence and Perception to be set up. I was here as a PhD student, then a postdoc. Then I went to Germany to study neuroscience, to study how the brain develops. I then went to London to continue my work, and I became an employee of the Medical Research Council in the UK. And after a few years, my boss, the eminent neurobiologist Mike Gaze, who had been in Edinburgh, decided he didn't want to be in London any more. And he persuaded the powers-that-be to transfer himself and his group to Edinburgh.. So I came back to Edinburgh to the Zoology Department at King's Buildings, in 1984. And in 1988, I became a scientific staff member in my own right with my own group, and I moved to Buccleuch Place to be part of the Centre for Cognitive Science. I had a group interested in neural networks, from the biological side, and also from the algorithmic side. And when the School of Informatics was formed, my research group became the nucleus of the Institute for Adaptive and Neural Computation (ANC) and I became honorary director. I couldn't be the director because I was already employed by the Medical Research Council. ANC grew from a very small number, including Chris Williams who was one of the first two academics - we only had two to

start with. We were the only Institute which wasn't formed from an existing department. I was succeeded as director, by Chris and then Nigel Goddard, Matthias Hennig and now Douglas Armstrong. And in about 2010, I transferred to the University, and retired in 2013. So that's my career.

Vassilis Galanos
That's fascinating. Would you like to elaborate a bit more on your own contributions to the field? How they relate to computer science or AI? Or how they do not relate?

David Willshaw
Yeah, right. Well, I know something about computers but I'm not a computer scientist. So I should imagine my main contribution to the field is in the AI sphere. This is a subject which I have worked on peripherally, although my primary focus is on brains. The difference is that, of course, in studying brains you have a structure you have to understand. In Artificial Intelligence, you make a structure, and then you try to use it. So it's two sides of the coin, you might say. My supervisor, Christopher Longuet-Higgins, emphasized that, basically, we're thinking of these two things, understanding human processes, which can be thought of in terms of how they actually work, and how we would design something that would be like a human. Mike Fourman, who initiated the School of Informatics, had a slogan, which was "Informatics is the study of natural and artificial computation", which is really nice, actually. So my contribution, I suppose, is really to do with the work that I did as a PhD student. I worked on a structure called the associative net, which you might say now is a simple neural network. And another person who came a few years after me, who was interested in similar topics, although his PhD was slightly different, was Geoff Hinton, who, obviously, you know, is one of the great archpriests of neural networks I always think of him as that. I later on designed an algorithm, which is quite interesting, it's not mainstream AI but it's in the way of designing algorithms to do clever things. I don't know whether you know this classic computer science problem, The Travelling Salesman Problem. The Travelling Salesman Problem is a classic problem in computer science, which you can think of in terms of a salesman who has a number of stops to make in the country, say, Britain, and he or she wants to do this in the minimum time possible, and so what he/she wants to do is to minimize the route to take. And so obviously, he/she should visit in a circle. And the question is, okay, if the towns are spread out quite uniformly, it's quite easy. But when there's lots of towns in different places, then it's complicated. So this is one of those problems which are very difficult to solve. What's the best solution? Can you make an algorithm that does this as efficiently as possible, minimize the length of the path, a circular path. It's related to all sorts of other problems in computer science. It's so hard, you can't really solve it cheaply in general. But if you could solve it, there's lots of other problems you could also solve. Another related problem is the knapsack problem, which is how do you stuff a whole lot of objects into a knapsack in an efficient fashion, so to you get as many of them in as possible. I devised an algorithm, which is actually based on how I thought the brain made connections, which gives a good algorithm to solve this particular problem, which is called the Elastic Net.
So, that's a small thing I did, but I mean, my main contribution, I would say, is probably being here. Like in many other parts of Informatics, there are activities going on that feed into what we now think of as AI. In my case, you know, one of our interests in ANC and my research group was neural network algorithms. At that time, there was a rivalry between people who tried to solve problems by neural networks and people who wrote logic programs to carry out logical deductions - the symbolic approach against

the subsymbolic approach. And in fact, there was a great rivalry between the two. And the curious thing is that after years of rivalry the mainstream of AI is actually driven by neural network algorithms run on great big computers.. So I suppose having an environment where this could flourish is essential and I was part of this environment.

Vassilis Galanos
This is extremely interesting. And we'll return to several of these subjects in the course of this interview.

Xiao Yang
Can I cut in with a question? So I feel, like a lot of people in your generation, you have a perception to study brains and biology and to understand how the brains work very well. How do your conception or perception change during this way? Because now it's actually in a department doing more like machine or machine learning, machine intelligence? So how do you think during the whole career, how do you feel this perception changes and are you happy with it, do you find the thing you want to find out first?

David Willshaw
Well, the people haven't changed. People who study or are interested in brains like Matthias Hennig, still study brains. I mean, the perception hasn't changed. I don't quite see what you're getting at in terms of perception changing.

Vassilis Galanos
I guess I think I see what Xiao also sees here. But so you mentioned earlier that, you know, that motto that Informatics was the study of both biological and mechanical processes so we just have the feeling that the biological might be a bit lost in the Informatics department. And you mentioned rivalries, you know, the course of history like machine learning taking over. So, I'm wondering whether there's something even material, maybe the infrastructures, the buildings have changed, I don't know any, powers, factors.

David Willshaw
The business about biology being lost is that most of the biological work in Informatics has been about working on brains from the mathematical modelling point of view. To do real biology would require a different type of building. In fact a lot of biology has been done through collaborations with other departments. One of the initiatives that was very successful was our Doctoral Training Programme in Neuroinformatics from 2002. It was one of the first big PhD training programmes funded by UK research councils. There are now quite a few in the School of Informatics, in different subjects. The first two in the UK were one in Oxford on vision, and one in Edinburgh on neuroinformatics, which is really computational approaches to neuroscience. We had lots of people doing neurobiology on that programme, doing experimental work in the biological departments of the University. So, on that perspective, I didn't have a feeling the biology was, you know, put aside. It's rather that all sorts of different collaborative activities have been carried out. That's fine, because it gives a wonderful environment. And, of course, we have a very big successful activity in Cognitive Science here in Informatics. The cognitive scientists, in terms of, say, language processing, have great big datasets. And obviously, neural network type algorithms have been ideally suited to analyse these. So that's another advantage of having a large community of people doing neural network, machine learning type things.

Vassilis Galanos
Yeah. Yeah, I think that came up naturally. Because you mentioned that sort of model before. So that's, that's extremely interesting. And you spoke about the environment here. And so I want you to want to push you a bit like what factors do you think made Edinburgh Edinburgh in terms of AI, Cogsci, Computer Science? What made Edinburgh so unique?

David Willshaw
Well, I think this comes down to the activities in the early days. I have to here mention my supervisor Christopher Longuet-Higgins. This is going back to when the Department of Machine Intelligence and Perception was set up. In order to set up something new you have to have lots of money to do that. And as far as I understand it, the idea was to set up an organisation that studies not only AI, but also the mind. Longuet-Higgins and Richard Gregory were interested in the mind. Donald Michie was interested in machine intelligence. He was a geneticist by training. but he had ideas, very early on about say, the benefits of computerised medicine, for example, which is a long time before anybody else. Also to mention is that there is a classic type of algorithm called reinforcement learning, which is important in neural network theory (and neuroscience), and the people who developed that in the USA always credit Donald as the first person who'd had the idea in a very simple way. What Longuet-Higgins did was, firstly he had a group in neural networks with Geoff Hinton as a student. He then developed an interest in the mind and Cognitive Science, a term which he is said to have invented. Concurrently he set up what I always regarded as a discussion forum of eminent people across the disciplines, in the so-called School of Epistemics - you know about this?

Vassilis Galanos
Yes, yes, Xiao has told me a lot about this.

David Willshaw
He involved an array of people from outside the Department because he was a very eminent guy with very broad interests. He was, by training, a theoretical chemist, and some people reckon that he should have got a Nobel Prize. As an undergraduate student, he made major contributions to chemistry that enabled his supervisors to become Fellows of the Royal Society. He became a Fellow himself at the age of 28. So he was a superstar. He knew people from the arts, from zoology, from biology, from mathematics, chemistry, and they would come and give talks. I don't know whether you've looked at the archives about the School of Epistemics. It is interesting seeing the correspondence.

Vassilis Galanos
So no one recorded the actual events in the sessions?

David Willshaw
There are minutes of the Epistemics Steering Group. visitors would come and give three or four invited lectures, there'd be no record. Or the only record would be their overhead transparencies

Vassilis Galanos
Yes. I heard there were papers sometimes or manuscripts.

David Willshaw
I looked at the archives as well. This activity is what eventually developed into the Centre for Cognitive Science. So Christopher sort of broadened the perspective. And as far as I know, he was interested in AI, but he didn't practise AI. It wasn't his thing. He talks about understanding the mind as an execution of a computer program. I've got his book here, which is a set of collected papers, and there is a small introductory section there.

Vassilis Galanos
The first papers are about AI and sort of defining the field.

David Willshaw
Here's a quote.
When in 1967, Richard Gregory and I packed our bags for Edinburgh, it was on the shared conviction that the workings of the mind could not possibly be as tedious as the psychologists made them out to be, or as peripheral as the physical scientists tended to assume. We were welcomed to our new university by Donald Michie, the British pioneer of machine intelligence, C.H. Waddington, not so much a biologist as a renaissance man, he was a geneticist, and James Thorne an eloquent expositor of linguistic theories of Noam Chomsky.
Christopher put all these things together. He was able to do that. And sadly, you know, he left in 1974 to go to Sussex. He and Donald had disputes about, I don't know exactly, but there were disagreements probably about where the Department was going, because Donald had his big robot project. And Christopher was really happy for him to be doing that. But he didn't particularly want to be involved with that himself, because he had his own agenda. Richard also had gone to Bristol a few years earlier. Originally there were three people each with their own agendas. One of those agendas was AI, and there were other agendas. So there was a fusion, which was good for a time, which developed a multidisciplinary activity. So Edinburgh had this multi-disciplinary approach. Going forward in time, I am reminded, as I was thinking about this, of the Centre for Speech and Technology Recognition, CSTR, who were initially set up by a grant. And their plan, which is a very sort of AI-ish type plan, in the 80s, was to deliver a speech-driven word processor. Henry Thompson, who's still here, he's Emeritus, he's a computer scientist. He was, as far as I remember, one of the people involved, together with linguists, phoneticians, computer scientists, electrical engineers, and other people. All sorts of people coming together. Edinburgh had the environment to be able to support that very multidisciplinary project that enriched the environment. They must have done neural network type stuff, because I used to teach a course on neural networks unofficially. We had colleagues from CSTR teaching on that course. They taught the fundamentals of the classic algorithm in neural networks, the backpropagation algorithm, which is how neural networks became so good at pattern recognition. So, it's an environment, which is multidisciplinary, and I think that's the strength of Edinburgh. From the point of view of the leaders, you have to not just have one core theme, you have to have maybe a core theme and other people doing things are related to that. From this you get inspiration, and the university is there for feeding off other people and getting ideas, you know, getting together.

Vassilis Galanos

Very interesting. For those who are reading or listening to this interview, would you like to expand a bit more on the School of Epistemics, but also give examples from research projects or initiatives that have been quite prominent in that time? You know, what were the sort of research projects that kind of established Edinburgh? You mentioned something at the very beginning, you know, that you need funding to set up something big. So what kind of projects gave the funding in this respect?

David Willshaw
The CSTR example might be an example of that, it takes a few years to get funding for a project together, so. Another later example is the establishment of the Human Communication Research Centre by the Centre for Cognitive Science.

Vassilis Galanos
Would you like to expand a bit more on the School of Epistemics and its value.

David Willshaw
The School of Epistemics had its outward presence and it ran short courses. Speakers from everywhere would come. And you could go and talk to the speakers afterwards. So they were there in residence for a day or two. They would expect that people will go and talk to them about their own research. So it was, I mean, the value of it was that it's another multidisciplinary organization because it involved many of the main professors in the university, The only professors, I understand, who wouldn't take part, or who didn't figure in this were those in Philosophy. So everybody else came - including Genetics, Zoology, Botany, Engineering, Physics, Mathematics, Linguistics, Computer Science. Edinburgh had a very big Linguistics department. John Lyons, who was the professor of Linguistics, was very a powerful Chomsky supporter. So, it was really a meeting point for those people interested in the multidisciplinary account of science and cognition and biology and AI as well, and computer science, all together. So there wasn't any sort of track from lecture to product as it were. Perhaps these days it is different.

Vassilis Galanos
Yeah. When I first heard about it, the first thing that came to my mind was Cybernetics. And the history with the Ratio club. Yeah, so I wonder if that was Edinburgh's version of that.

David Willshaw
It could be. Longuet-Higgins was born in 1926. He could have been part of the Ratio Club who in the 1950's discussed the new subject of Cybernetics. One of the members of the Ratio Club was Donald MacKay who was originally a physicist. He was the first person in Britain with a group, at the University of Keele, on brain modelling and experimental brain science, modelling and brain science together. He was the father of David MacKay. Do you know about David Mackay? Talk to Chris about him. He was one of the big guys in machine learning. Sadly he died of cancer, very young. The Ratio Club was slightly different from the School of Epistemics as it wasn't focused on one University. There is a photo of a meeting of the Ratio Club available on the internet. The impact that the School of Epistemics had was that eventually it became the Centre for Cognitive Science.

Vassilis Galanos
I'm curious, you said earlier, correct me if I'm wrong. philosophers were excluded from the School of Epistemics?

David Willshaw
They weren't excluded. They didn't want to take part, apparently.

Vassilis Galanos
They didn't want to go. Why do you think is that?

David Willshaw
I don't know. I don't know who the professors of philosophy were. Maybe it was just a personal point of view of one or two people who felt that this was not for them. I thought it would be an opportunity to discuss things. There was nobody from theology departments, or from law. Certainly there were many departments involved: genetics, molecular biology, chemistry, biochemistry, zoology, psychology, linguistics, phonetics, physics, mathematics, computer science, machine/artificial intelligence, engineering, everybody was coming to these meetings. And there were lots of other smaller groups who contributed. There was a computer aided design unit, I think it was Science Research Council funded. People from the MRC Speech and Communication Research Unit and The MRC Human Genetics Unit also contributed.

Vassilis Galanos
We see them now. Yeah, yeah. Yeah. Did you encounter the Science Studies Unit?

David Willshaw
I knew about it, but I didn't have much contact.

Vassilis Galanos
That's where I got my PhD from this Department now called Science Technology and Innovation Studies.

David Willshaw
I believe the Director would have come to the Epistemics seminars. I'm not quite sure, I think.

Vassilis Galanos
David Edge was the Director. My supervisor's, I think, most cited paper was coauthored with David Edge. Robin Williams. And we have lots of conversations about C H Waddington, former Professor of Genetics and the connection between genetics and AI. And also the radical science movement representation in the UK. So open questions for historians, I guess.

David Willshaw
There's another area which Christopher Longuet-Higgins also went into, which is theoretical biology. There's a whole series of books called Towards a Theoretical Biology which Waddington edited those. Edinburgh University Press. They're in the special collections in the University Library. These grew out

a series of discussion meetings about the theory of biology in all sorts of ways, you know, from genetics to more behavioural studies.

Vassilis Galanos
Yeah, yeah. I've only gone as far as reading a bit of Maturana and Varela, second order cybernetics kind of thing. I'm not sure if there was any connection between Waddington and these people. I have heard that he was in correspondence with Gregory Bateson, but I haven't verified that.

Xiao Yang
What about Bernard Meltzer? He was the head of the Metamathematics unit, right? He developed this into the Department of Artificial Intelligence or was it totally changed?

David Willshaw
Bernard was originally an electrical engineer and head of this Unit, working on automatic proof methods which contained some very eminent computational logicians, including Pat Hayes and Bob Kowalski, now Emeritus in London at Imperial College London. When the original three Professors of the Department of Machine Intelligence and Perception went their separate ways in the 1970's, the Unit got integrated into the Department of Artificial Intelligence. And Bernard became the Head of that. And then I think in 1977 Barry Richards came on board. Under his leadership the School of Epistemics became the Centre for Cognitive Science. It ran a graduate course and became associated with the Department of AI. The Centre ran a PhD programme and Master's course which all PhD students had to attend. This had quite an unconventional structure. Everybody had to do the same course, and to start with there were no options. So when I said I taught neural networks, I wasn't allowed to teach it formally as it wasn't in the curriculum. There were no formal exams but instead students had to write a take home review paper, which turned out to be extremely demanding. When the Centre got started it had a number of lecturing positions, which were joint positions with other departments. The Departments of AI, Computer Science, Linguistics and Psychology all contributed a staff member position who would jointly be in cognitive science and AI, or computer science etc. So to start with, most of the positions were half positions in two departments. Cognitive Science gradually got their own dedicated positions.

Vassilis Galanos
Extremely interesting, it was a process of absorbing different disciplines. And then out of this process, new disciplines emerged, they became established. Very interesting

David Willshaw
A big change was when Cognitive Science got large research funding from the ESRC, to form the Human Communication Research Centre. Keith Stenning was a lecturer then and became Director, followed by Johanna Moore, who became Head of School.

Vassilis Galanos
What year was that, approximately?

David Willshaw

I'm thinking of 1989. And it involved Glasgow, and also the University of Durham. So in those days there weren't many centres that were multi-university. To start with, it means that you have two or three bureaucracies to agree with each other. It was the ESRC who funded it. It was quite a big organisation. That made Cognitive Science a bigger and more influential organisation.

Vassilis Galanos
Great lessons for us young scholars as well. Yeah, that's extremely interesting. Kind of organically leads me to my next question. So, you have you have mentioned some very notable figures, Longuet-Higgins, Hinton, Michie and so on. Are there less noted individuals that you have particular appreciation for that perhaps history hasn't treated them just as the rest of the people. Somebody that on the occasion of this interview you'd like to highlight?

David Willshaw
Well, somebody who was put in a very difficult position but who made AI survive was Jim Howe. He came to Edinburgh as an associate of Richard Gregory. Jim Howe was a psychologist by training. The main site of their activity initially was at the old Drill Hall at Forrest Hill where originally the University Army Cadet Corps trained. The Drill Hall is a great big barn, where there's a tunnel that Richard and Jim built for doing movement perception. People would be pushed up and down on a railway inside it. They were looking at the problems of visual perception in motion. Richard Gregory had a grant from the United States Air Force to fund this. Jim was one of the leading people on this project. He became a Lecturer in Machine Intelligence and became interested in computer assisted learning. When the various professors started to disagree in 1970, I think he was promoted to director of the department, although he had no power, because he wasn't a professor. He had a very difficult time because he wasn't made Professor until very much later in his career, which was a shame, He had to negotiate between the various factions.

Vassilis Galanos
Yeah, yeah. And I've read a short text written by him on the history of AI, and he's the one who speaks about the rivalries and so on.

David Willshaw
I read that recently and I think it's a very good account. Apart from those rivalries, going back to computer science I think there was a rivalry between computer science and AI as well. I was reminded of that when I was reading something, oh yes it was the Wiki pages of Donald Michie. In the Wiki pages is something about that he started off his own research group by 1963, ‘predating the formation of the computer science unit’. I think there was a rivalry, that one side felt they were doing things in a principled way, and the other side felt they were doing things in a realistic way, where principles are difficult to implement. I didn't really have a lot of contact with the computer scientists. Another person you may want to talk with is Gordon Plotkin. He started with me at exactly the same time. He was supervised for his PhD by Rod Burstall. We both attended the Diploma course run by the Experimental Programming Unit, part of the original Department of Machine Intelligence and Perception..

Vassilis Galanos
Would you like to share any memories or expand a bit on Rod Burstall's influence?

David Willshaw
I don't know much about Rod Burstall. Gordon might be the best person to ask about. I knew him personally, but I didn't really have any contact professionally.

Vassilis Galanos
So that's very interesting. So let's go back to a topic you have touched earlier on, in terms of challenges you can think of in the period between 1960 and 1990. There have been many challenges, right? So it's not just the inner kind of rivalries, but also external kinds of challenges, sometimes our funding challenges, sometimes they are interrelated. What are your recallings from those days?

David Willshaw
The main challenge was the issue over the Lighthill report into Artificial Intelligence in the 1970's. And I think it's simply that this was initiated in response to a big grant proposal put in by Donald Michie to fund his robotic work. And I happened to see some of the proposal. It was clearly an extremely large proposal. And if he had been more modest, he would have probably got the grant. He was asking for a hell of a lot of money. So it's not surprising that somebody will be questioning that, saying is it worth it? This is an interesting lesson for us. It was unrealistic. And the somewhat ironic thing is that, at the time the proposal was written the robot didn't seem to do very much, although a lot of work had been done in the early days of the robot. After the inspection, but before the report was published , the work came along in leaps and bounds. Harry Barrow was very much involved in the project. He was the lead postdoc or maybe a lecturer. He went to the firm Schlumberger in Cambridge and then to the University of Sussex.

Vassilis Galanos
I'm also thinking of, you know, the political climate of the time, whether that had influenced decisions or perhaps, have played any role in the rivalries or alliances between people we see today. You know, some scholars want to, for example, receive money from the military, some don't. I think the 70s were very much post-Vietnam War kind of climate. So I'm just wondering, how did it feel doing that kind of science at that time?

David Willshaw
I think at that time, there probably wasn't much thought about, should we be accepting money from the military say? Richard Gregory himself had a grant of United States Air Force. I don't remember any ethical conversations. Political in the sense that, there's the sort of local political, like what does the government want to do, and whether we agree with what the government wants to do in terms of spending the money. And I suppose to a certain extent, people apply for funding because they realise that there are opportunities available. They don't usually go to the government and say, look here, we've got to do more work on this, give us the money. So it's opportunistic. I suppose it's probably always been opportunistic; I don't know. These are very hard questions.

Vassilis Galanos
Earlier on you mentioned that you have a certain opinion about the Lighthill report. So I'm curious to ask about that as well.

David Willshaw
I think that having an enquiry was justified on the basis of what Lighthill was presented with - an enormous grant proposal. That was the trigger. He was asked to assess the work. He said advanced automation is good. And computerised investigation of the nervous system is good but not the robot work. It's probably that he couldn't answer the question: to what end is the robotic work is being used? What are the benefits given the enormous amount of money being requested? The investigation was coloured by that there was lots of antagonism. Then there was the story about, oh well, British leads, and the lead in Robotics has been cut down, because of this. And there's this dark winter of AI stuff. Maybe that's true, but then I feel if Donald had not been so ambitious, it wouldn't have happened.

Vassilis Galanos
Do you do you see any connection between the so called AI winter and the Lighthill report and the things you mentioned earlier on, you know, the flourishing of different fields, would this sort of idea of the winter kind of inspire people to move on and develop new kinds of approaches in doing AI, or other related disciplines?

David Willshaw
The direct consequences were that people who worked in Edinburgh on robotics didn't have their funding renewed, and so they got jobs elsewhere. Harry Barrow, for example, had been trained as a psychologist. He became an expert in application of AI techniques to engineering and had a successful career in academia and in industry. Donald Michie didn't stop working. He had his own institute in Edinburgh, separate from the Department of AI. And then he went to found in Glasgow the Turing Institute working on AI applications. The Artificial Intelligence Applications Institute was founded in Edinburgh in 1983 and became one of the founding Institute of Informatics. Probably the specific infrastructure underpinning AI was lost. I don't really know enough about this as I was away from Edinburgh between 1974 and 1984. Alan Bundy was appointed in around 1977, when AI was being built up, and so can tell you more about this time. So presumably, the Department of AI itself would have had a negative image with reputational loss, certainly, lots of other things. As far as flourishing of different fields post-Lighthill was concerned, I would see this as that these fields were already there pre-Lighthill and had more time and space to breathe in post-Lighthill in the absence of the big robotics project.

Vassilis Galanos
Yeah. Well, extremely interesting. I'm conscious of time we're over an hour. I am curious to ask you, your message to young scholars now Edinburgh becomes once again a cradle for AI. So many people come here to study AI. So what would your message be to the young scholars? You know why should they be proud for being Edinburgh AI, computer science scholars?

David Willshaw
Obviously, follow your courses, but try and always be thinking more widely. Thinking outside what you're taught, in terms of looking at the opportunities around here. However, now I don't know much about this. And of course, there are many startups that come from Edinburgh. Informatics particularly seems to have a great reputation for that, so many opportunities. The whole thing has changed since I was active, or since I was growing up, the whole environment in terms of what you do and the pipeline

on which you go is changed. Now, for example, people don't talk about having jobs for life - they do one thing and then another thing. So it's so it's a completely different from how I was in schooled, as it were. So there are so many other opportunities that I didn't have. So flexibility is the thing. Adaptability, you must adapt these days

I gave a five minutes talk on history in February or January this year at a symposium tracing 60 years of AI and CS. My message was the importance of diversity which can offer up new challenges. Especially when what you've learned has suddenly changed and there are new opportunities that you could grasp. So that's the message.

Vassilis Galanos
Do you think diversity has been kind of stable throughout Edinburgh's legacy? Has it been challenged?

David Willshaw
I think there's more diversity. But it's probably because there are more types of technology, it's changing so much more rapidly than it was even 20 years ago, 10 years ago. So it's probably that which is the thing. That makes sense.

Vassilis Galanos
So you monitor this sort of global evolution of technology. Do you think what is Edinburgh's role in this global evolution?

David Willshaw
I don't know. I'd say that's a big question.

Vassilis Galanos
I suspect nobody knows. But yeah, from your own point of view from what you collect. So. if Edinburgh has to position itself in a sort of global context of AI, computer science, innovation, what would Edinburgh thrive with?

David Willshaw
I am more interested in addressing specific problems. To know about this, one has to know what the future is going to bring. It is difficult to predict the future but of course there are known challenges. If you take the challenge of climate change, for example, then it is very likely that Edinburgh AI and Computer Science can be brought to bear on the underlying problems.
For this sort of thing to happen, any grouping centred around AI and Computer Science together with a diversity of expertises stretching beyond the core disciplines would be an ideal environment. Secondly, I am a firm believer that to understand a problem and apply one's methods one needs to know technical detail about the problem rather than being brought in as an external consultant with little knowledge of the problem and the technical issues. This was the philosophy we used at our Doctoral Training Centre. We trained students from the physical science to address neurobiological problems and for this they needed to know some neuroscience whether or not they worked directly in the lab. They needed to be multidisciplinary.

Concerning the limits of AI, AI is good at recognising patterns, which is great. Then you step into generative AI, which starts telling you stuff that is dangerous. I'm not an expert to know what to do about that. But the fact is that if you give an AI system the wrong information, then there is a problem. That's what really frightens me about AI. You probably know these simple examples such as: ask a generative AI system for who was the first man to walk over the English Channel and you get a name. Obviously the system in there didn't have the contextual information that we can't walk over water. You see legal cases of people being accused in court of cases that were generated merely because of false correlations that were found when the system was bringing different types of data together. So stuff like that is very worrying to me.

Vassilis Galanos
Well, very interesting to hear that from you. So I think enough with the hard questions. All right. I see you have done some prep work. I'm sure you have many memories to share. So just want to tease you a bit if you have any anecdotes you want to share something that is worth preserving stories you want to share with the posterity.

David Willshaw
I have an anecdote about Geoff Hinton's application to study in Edinburgh, which he may not know. One day I was in Christopher Longuet-Higgins's room when he was looking through PhD applications. On this desk he had a large pile of letters of application – those days people wrote letters – which he was leafing through. Progress was rapid until he came across one written in blue ink, which stood out from the others written in black. He pulled it out and started reading: "This looks interesting" he said. This was Geoff Hinton's application. He'd written down the design of a simple type of associative memory neural network, having had no prior knowledge of a neural network with associative memory. So that was quite interesting to see that he's already had specific ideas in his mind rather than general notions about say solving the brain which many students have. Many students come up, they say, I want to work with you, you know. What do you want to work on? Not quite sure maybe I'll work it out. But Geoff had a plan. He shifted his plan over time, as you do with PhD studies. He also knew how to gain attention - as he has done throughout his career.

Vassilis Galanos
Extremely interesting. Yeah. Do you think there is a story around Christopher Longuet-Higgins reading the Perceptrons book by Minsky and Papert just at the time that Geoff Hinton arrived. Do you think that book played an important role for Edinburgh's history around neural nets and everything?

David Willshaw
That was a beautiful book but I don't think it did. Longuet-Higgins developed an interest in memory networks from two sources. Before starting in Edinburgh he had spent some time at Bell Labs in the USA visiting people working on the hologram which was said to exhibit distributed storage of information – in the sense that if a part of the hologram in which an image is stored is cut off, the image can be reconstructed. This seemed to tie up with findings by the psychologist Karl Lashley which suggested that memories are stored in a distributed fashion rather than in specific locations. Christopher got interested in the idea that the brain stores information not just as in individual slots in the filing cabinet, but similar

to the way a hologram does. If we then examine a portion of the hologram, all the memories are there, because the memories are distributed. In fact my PhD thesis was on distributed associative memories. So that was his inspiration for getting into neural networks. And then the way we got to the neural net associative memory structure which we analysed was by thinking about analogies to the holographic-type memory, rather than from reading Minsky and Papert. Longuet-Higgins didn't like reading books or papers - no, that's not true. he liked to work out everything from first principles. Incidentally, Christopher's first holographic-type memory model would be called a correlational neural network these days. Another anecdote concerns Christopher. A classic text on biophysics was written in the late 19th century by Hermann von Helmholtz, who was a German physicist/psychologist or psychophysicist. In this book he described theories of visual illusions (which was a favourite topic of Richard Gregory). Leon Glass, now an eminent mathematical biologist famed for describing the mathematics of the way the heart works, was carrying out a post-doc with Christopher. Leon was interested in visual illusions (and is now known for an illusion called the 'Glass effect'). He told me he had a supervision with Longuet-Higgins, who said, "I've been thinking about how to understand visual illusions last night, how about this explanation?" And then Leon said "Well, actually, that's very interesting, Christopher, but in fact, if you read Helmholtz it's there." He got a bit annoyed about this - his idea had been thought of already. "So what about this?" "Actually that's there too", and so Christopher went through, one by one, inventing on the spot all the theories of visual illusions that Helmholtz had described his book. So that was quite amazing and showed the sort of person he was. He got a bit upset that his original thoughts weren't original thoughts after all and he wasn't keen on reading the original texts. The sort of neural networks in the Minsky and Papert book is different from the neural networks that we studied. The origin of the Perceptron almost goes back to the time of the Ratio club, in the 1950s.

Vassilis Galanos
From what I understood, Minsky and Rosenblatt (the inventor of the Perceptron) were high school classmates. So it goes back to their conversations from times in high school, most likely. Yeah, but it goes back to McCulloch and Pitts as well, I suspect.

David Willshaw
They were interested in modelling neurons as logical elements rather than networks that could be used for information storage. But didn't Minsky do a PhD? His PhD thesis was about neural networks, wasn't it?

Vassilis Galanos
And how did your approach especially in Edinburgh differ to the Perceptrons approach? You said it was quite different.

David Willshaw
Well the history was different; also, the actual functioning of the network was different. The Perceptron was really the forerunner of deep learning networks and other types of current networks.
These networks learn by error correction to assign data to specific categories. The simplest architecture is a set of so-called input units where the data is presented which are connected to a single output unit where the response is given, with adjustable 'synaptic strengths' or 'weights' between input and output.

The big stepping stone from what Rosenblatt did, and what people did in the 1980s, was that you need intermediate layers of computing units. At the time there was no known algorithm to change the weights of| the intermediate units, until this was solved by the so-called backpropagation algorithm. Deep learning is that idea extended to lots and lots of layers of units. It's not particularly theoretically more interesting - just an extension. In our networks, it's a one trial learning situation where the network associates information together, say name of person and telephone number. Here's the name, here's a telephone number, store it and then repeat many, many times with other names and telephone numbers. And if you don't put too many associations in, then a phone number can be retrieved from a name or vice versa. You can't do this forever, because after many associations, it starts making mistakes, muddling them up And we worked out what the limits to storage and retrieval are. The Hopfield Net, discovered many years later, which comes from the physics community, does a similar sort of thing. Being interested primarily in neurobiology, we always felt that our system was more biologically plausible than a Hopfield Net. Probably it is but a still a long way from the neurobiology!

Vassilis Galanos
Some contemporary scholars have started building theories around how these different approaches to doing neural nets or AI or machine learning, have implications about how we see the world or how we understand the world. And myself studying the early foundations of AI, sometimes the question is, I do AI in order to understand the mind, or I take inspiration from the mind in order to do AI. What do you think your approach, how do you how do you think its stance positions itself within that?

David Willshaw
I think that it all depends on what your goal is – there is no general rule. Insofar as the neural network derived versions of AI are based ultimately on inspiration from neuroscience, it would be a bit circular to then use them to understand the brain or the mind. My only experience on working in the AI/Computer Science world is through my idea of a novel approach to solving the Travelling Salesman Problem (TSP), which we have already discussed. Here the inspiration was that in the brain there are many topological maps between two structures. For example, there is a 2D map of visual space through the retina onto many brain regions – a 2D X 2D map in which neighbouring points in visual field project to neighbouring points on the brain region. With colleagues I had developed a series of models for the development of these maps encapsulating the idea that neighbours project to neighbours. In the TSP, the mapping is now of a 1D structure (the tour) to a 2D structure (the space of cities) with the neighbourhood property preserved. My insight was to adapt the same algorithm developed for the neurobiological case to the TSP and this is what we did. Of course the constraints are different. For example, for the TSP you don't have to worry about biological plausibility. Like all TSP algorithms, the Elastic Net gives only approximate solutions, and the algorithm is parallisable. I give it here as an example of how looking at a problem in one domain gives inspiration for a novel approach to a problem in an entirely different domain.

Vassilis Galanos
But you need to have access to lots of, I guess, mapping information, architectural information to know what the conditions and the context is in order to find and sometimes that goes beyond what is controllable from the outset. Very interesting. Wow, we have the room for another, 17 minutes. I'm conscious

of time. And I know that there is a lecture or something. Have you got any other anecdotes to share? You've been thinking on your way here, or any other thoughts in general?

David Willshaw
Just trying to think of any more anecdotes. The only one I can think of as a general point, which can be illustrated by a small anecdote is that when I was here as a PhD student, certainly from my perspective, everything was more relaxed. Perhaps it was because we weren't looking at our mobile phones every half an hour or half a minute or so. On the contrary, some of the departmental bureaucracy was very bureaucratic - Donald Michie was a very bureaucratic type of person. He used to send around Civil Service type memos, with three boxes on them labelled 'R', 'S' and 'T' which he ticked to indicate whether they were destined for Research, Secretarial or Technical people. I remember two occasions when we tweaked his tail by sending out spoof memos. I must admit that I wrote one of those. It must have been a success as I heard afterwards that Donald had tried to suppress its distribution by intercepting the Servitors who were tasked with delivering the mail. Very much later on, in the early days of Informatics, I sent around an April Fool about Informatics spreading throughout Scotland – the (fake) link between computer languages and Gaelic and a new building with the use of new single office accommodation with excellent video facilities (actually a decommissioned prison) etc., about which I was very proud – particularly because of the large number of people who believed it for a while.

Vassilis Galanos
Yeah, I wonder if this kind of playful culture also played a role in Edinburgh's position as a cradle of AI and computer science?

David Willshaw
I don't know whether this extended to other places. Perhaps the speed of life was not so fast. If you communicated with people, it was usually by phone or letter.

Vassilis Galanos
Yeah. Do you think contemporary technology culture is influenced in a negative way because of the speed of life? Do you feel nostalgic about those days? Do you feel optimistic about contemporary situations? Or a mix of both?

David Willshaw
Yeah, well, I don't use social media. I would find that too much. So on that basis, I think the speed of life is a negative, I mean, current speed of life is too fast. A slower speed gives more time for reflection. So probably this does have a negative impact. I suppose there are benefits. If you have got an urgent thing that can't wait, you call somebody, text somebody, and you get an instant answer.

Vassilis Galanos
It's a tradeoff. Yeah.

David Willshaw
Do people read these days.? Do you need to read?

Vassilis Galanos
You get generative AI producing summaries, dubious summaries of texts.

David Willshaw
That's worrying. I mean, obviously, that's something that Informatics is interested in from the policy development side. That's a question of, can you make guidelines for this, it's going to be very difficult.

Vassilis Galanos
In general? You can write a list of recommendations, but will people take them into account? That's a different story.

David Willshaw
I don't know enough. It's very difficult to find out what are the sort of questions you need to ask and what is the guidance needed? Should it be guidance or imposition? And if it's an imposition, it has to be the right imposition.

Well, time is running. I see people might be waiting outside. I'd like to thank you for sharing your experiences, your insights. If you have anything more to share, feel free to send me an email say, you know, you want to share a bit more or we'll be in touch with the transcript and the proofs for your pending your approval. It has been a real pleasure. Thank you very much.

David Willshaw
My pleasure. I hope that some small proportion of what I have said is useful.