Case No: CA-2024-000036

IN THE COURT OF APPEAL (CIVIL DIVISION)
ON APPEAL FROM THE HIGH COURT OF JUSTICE
CHANCERY DIVISION
BUSINESS AND PROPERTY COURTS OF ENGLAND AND WALES
ON APPEAL FROM THE UK INTELLECTUAL PROPERTY OFFICE
Sir Anthony Mann (sitting as a Judge of the High Court)
[2023] EWHC 2948 (Ch)

Royal Courts of Justice
Strand, London, WC2A 2LL

Date: 19/07/2024

Before :

LADY JUSTICE NICOLA DAVIES
LORD JUSTICE ARNOLD
and
LORD JUSTICE BIRSS

- - - - - - - - - - - - - - - - - - - -

Between :

Comptroller - General of Patents, Designs and Trade Marks	**Appellant**
- and -	
Emotional Perception AI Limited	**Respondent**

- - - - - - - - - - - - - - - - - - - -
- - - - - - - - - - - - - - - - - - - -

Brian Nicholson KC, Anna Edwards-Stuart KC (instructed by **Government Legal Department**) for the **Appellant**
Mark Chacksfield KC, Edmund Eustace (instructed by **Hepworth Browne**) for the **Respondent**

Hearing dates: 14th & 15th May 2024
- - - - - - - - - - - - - - - - - - - -

Approved Judgment

This judgment was handed down remotely at 10.30am on [date] by circulation to the parties or their representatives by e-mail and by release to the National Archives.

....

Lord Justice Birss :

1.   The first question in this appeal is whether the exclusion from patentability of a program for a computer "as such" by s1(2) of the Patents Act 1977 has any application to artificial neural networks. These networks are the backbone of the machine learning systems on which modern artificial intelligence systems are based. If artificial neural networks do engage s1(2) then the second question arises. This concerns how that exclusion would apply to the particular patent application in this case. By decision BL O/542/22 on 22 June 2022 of the Hearing Officer, Deputy Director Phil Thorpe for the Comptroller, the patent application was rejected on s1(2) grounds. Sir Anthony Mann, sitting as a judge of the High Court, allowed the appeal on 21 November 2023 ([2023] EWHC 2948 (Ch)). The judge held that no computer program was involved at all, at least for a hardware implemented artificial neural network and therefore the exclusion had no application. The judge also held that even if the provisions did apply, the subject matter was not excluded. The Comptroller appeals to this court with leave from the judge.

2.   The patent application is in the name of the respondent Emotional Perception AI Ltd ("EPL"). The invention is a system for providing media file recommendations to a user. A typical example of its use could be on a music website, where a user may be interested in listening to music similar to another track which they are already aware of. Existing websites are capable of offering similar pieces in the same category (such as rock, folk, classical etc.), but the categorisation tends to be limited to types of music. The existing approach depends on the classification of individual tracks into categories by people, i.e. human beings. The advantage of the invention is said to be that it is able to offer suggestions of similar music in terms of human perception and emotion, irrespective of the genre of music and the apparently similar tastes of other people. The invention arrives at these suggestions by passing music through a trained artificial neural network.

3.   The two relevant claims are set out in an appendix to this judgment. Claim 1 is a claim to a system for providing file recommendations and claim 4 is to a method for providing the same recommendations. It is convenient to think about the claims as if they relate to music but in fact the invention is applicable to other media and the claims are not so limited.

4.   In order to understand how the invention works the starting point is to explain the nature of artificial neural networks, which from now on I will refer to as ANNs. The explanation which follows is not intended to be contentious. Much of it comes from the decisions below, the patent application and the expert report of Professor Pardoe which EPL filed before the Comptroller.

5.   An ANN is a machine built, as the name suggests, as a network of things called artificial neurons. These artificial neurons are akin to the neurons in the brain. In an ANN they are arranged in layers. Each neuron is connected to other neurons. Each neuron is capable of processing inputs and producing an output which is passed on to other neurons in other layers. The first layer receives inputs from outside the ANN system and the last layer produces an output from the system. These features of a conventional ANN are depicted in figure 6 of the patent application, as follows:
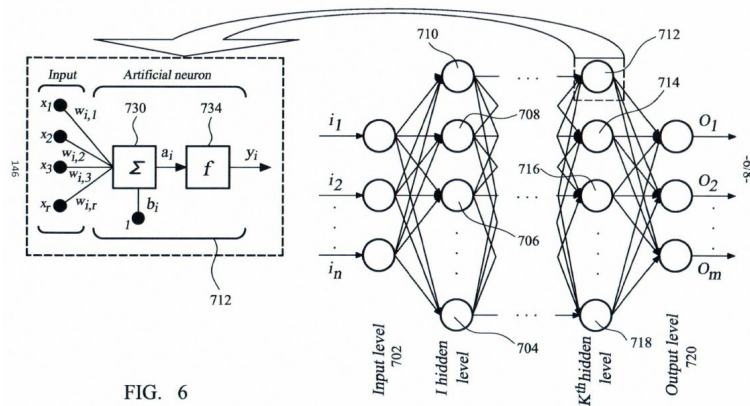
FIG. 6

6.   The right hand side of the image depicts the network of neurons. Each circle is a neuron and the connections are shown as lines. In the diagram the layers are called levels. The input level is on the left (702) and the output level is on the right (720). In between are a number of "hidden" levels. In the diagram there are n input signals, and m output signals.

7.   Given the current interest in AI it is worth bearing in mind, as was common ground, that ANNs as such are not new. A well known early example was a machine called the Mark 1 Perceptron which was built in the 1950s. It had three layers.

8.   Coming back to the diagram above, the box on the left depicts a single artificial neuron (i). The neuron takes in a number of inputs (x_1, x_2, x_3 to x_r), applies a specific weight to each input ($w_{i,1}$, $w_{i,2}$, $w_{i,3}$ to $w_{i,r}$), then adds these weighted values all together (the symbol Σ in box 730 refers to this summing function). A further single weight called a bias (b_i) is given to the result of adding up the weighted inputs and then a function (f) is applied to the output. That function converts the result (a_i) of the biased sum into an overall output (y_i). It is usually a non-linear activation function such as a continuous sigmoid. That sort of function, in effect, sets overall output (y_i) to zero if (a_i) is below a threshold, it sets overall output (y_i) to unity (1) if (a_i) is above a higher threshold, and if (a_i) is between the two thresholds it sets overall output (y_i) to a value ranging between 0 and 1, which scales with input (a_i).

9.   Thus an ANN is a machine which processes information. It takes a set of input information (on the left) and produces output information (on the right). Each neuron acts according to the aggregate set of its weights and biases. The effect of the ANN as a whole is the result of the overall network structure – the number of neurons in each layer, the number of layers, the links, the function and all the weights, biases. As the Professor explained the weights and biases are parameters which are adjustable during training (see below). Like the Professor ([22]), from now on I will simply use the term weights. There is a point of detail that in some systems links can be "pruned" too, but that does not matter.

10.   So, for example, once it is set up in the right way, if the ANN is given data representing an image, the ANN's output might be a signal (1) on output O_1 if the image is a picture of a flower and no signal on the other outputs, whereas if the image is a dog there may be a 1 at output O_2 with the others at zero, and so on. As Professor Pardoe explained (at [33] of his report), another kind of output from an ANN is a multidimensional numerical vector which describes the input in an abstract manner. The idea would be

that similar inputs would produce numerically similar outputs. Applying this concept to a large amount of data populates a multi-dimensional space with discrete data points within that so called "embedded space".

11. So far, I have not described how an ANN is set up to do any particular task, in other words to answer the question - how does an ANN acquire the weights to do something specific? The answer is that these are found by a process referred to as training. In the training process the weights are adjusted iteratively in order that the ANN produces a given output in given circumstances.

12. The judge at [16] drew attention to the explanation of training given by Professor Pardoe. The operator needs to have a training dataset, a validation dataset and a loss function. Conceptually one can imagine starting with a naïve ANN in which all the weights have their default settings. The training dataset consists of sets of potential input data and an indication of the desired output (the target). For example in a group of images, the pictures of flowers, dogs and the images with neither are each marked accordingly. Data from the training dataset is presented to the ANN and the output is examined. The difference between the actual output and the target is called the error. The job of the loss function is to determine this error. The training process then applies small changes to the network parameters. The training dataset is applied again and the output examined again. The idea is to reduce the error. This process of feeding back from the output to adjust the weights is done repeatedly using every example in the training dataset every time, in order to reduce the error. Another term used to refer to one version of this process is backpropagation. Every now and again the validation dataset can be used to see how well the ANN is doing at correctly classifying data it has never encountered before. The validation dataset is not used to modify the network parameters.

13. Once an ANN has been trained one can move to the second stage and use it to classify new data it has never seen before. From now on the network topology and parameters remain frozen. Used at this second stage the ANN is sometimes referred to as a pattern recognition machine or inference engine. There are some internal parameters which are used at the training stage which influence the learning process but are not used in the inference stage.

14. An important point of detail was explained by Professor Pardoe at [41] of his report. Once the network topology and parameters are frozen (static) this allows the specific implementation of this pattern recognition machine to be implemented in a range of ways and forms.

15. To expand on that briefly, the point is that once an ANN has been trained, its weights can be extracted and used to set up other ANNs of the same kind in order for them to perform that task. By the same kind I mean the same fixed topology of neurons, links and layers. To convert a naïve ANN (call it machine A) into an ANN which can perform a useful task, one does not need to train that particular ANN machine A, one can simply transfer into it the adjustable parameters derived from training another ANN machine B of the same kind.

16. So far this description of an ANN is at the level of the functions of its components, the neurons and links. Conceptually there are two ways in which ANNs can be built in practice. They are referred to as hardware ANNs and software ANNs. In fact there is

a spectrum between these two but for present purposes that can be ignored. A software ANN can also be thought of as a software emulation of an ANN. In other words there is a conventional computer system in which all the components of the ANN I have described exist only as software. The neurons, links, layers, weights, biases and so on are only what one might call virtual entities.

17.   A hardware ANN is, as the judge put it ([14]), a physical box with electronics in it. Putting it crudely, all the neurons are made of nothing more than some electronic components such as resistors and transistors, the links are just wires, and the layers exist because of the way the link wires are arranged. More realistically the hardware ANNs are implemented on so called digital neuromorphic hardware or field programmable gate arrays. An advantage of this approach is that once a hardware ANN has the right network parameters for a given classification task, it can perform the classification task faster than the same ANN running as a software ANN on a conventional computer. They can also undertake the training faster.

18.   The same kind of ANN, in terms of the set of links, layers, weights and so on, can be implemented as a hardware ANN or a software ANN. As ANNs, the two are identical. The software and hardware implementations are the same in terms of architecture, weights and so on.

The claims of the patent application

19.   Turning to the claims, so far the case has focussed on claim 4 and I will do the same. The claim relates to a method. The following is based heavily on the judge's paragraphs 8 to 13. It is also worth repeating that in some aspects this description is more specific than the generalised claim language, in order to explain what is going on.

20.   A pair of music files is taken, each of which is accompanied by a text description of some sort of the type of music in the file. The text describes how that music is perceived by a human. At its simplest the music might be described as happy, sad, jazz, rock or anything else. The descriptions are in word form (hence the use of the word "semantic" in the claims). These descriptions are to be analysed by using the natural language processing (NLP) system. These NLP systems are in fact ANNs but the term ANN as it appears in claim 4 does not refer to that NLP ANN. By contrast I will use the term "EPL ANN" to refer to the ANN named as such in claim 4.

21.   The NLP system takes these existing characterisations of the tracks and produces a vector in an embedded space called the semantic space. So a music track which was happy and exciting might have one vector whereas a track which was sad and relaxing might have another. The similarity or difference between the semantic types of music is reflected by the distance between those two vectors in the semantic space. In the claim this distance is called the separation distance between the two files. It is what the claim calls a "first independently derived separation distance", which is a "measure of the relative distance between a first pair of training data files in semantic embedding space". The NLP ANN is not being trained as part of this process. It is being used to derive the separation distance in semantic space. Bearing in mind the name of the respondent company, this space is in effect a map of what one might call the "emotional perceptions" of human beings of the music tracks as those perceptions are expressed in the text descriptions.

22.　　The same two tracks are independently analysed in the ELP ANN for what are described in the claim as their "properties" in other words their physical properties, such as tone, timbre, speed, loudness etc. These are properties which a machine can measure (hence the reference to "measurable properties" in claim 4 a). The analysis produces vectors in a notional "property space" or "property embedding space", again with the differences or similarities in the music thus assessed reflected in the proximity of the co-ordinates. This part of the process creates what the claim calls a "second independently derived separation distance" for the same two tracks. The second independently derived separation distance is a measure of the relative distance between the two files in property space. This is in effect a map of the tracks by reference to their physical properties which can be measured by a machine.

23.　　The next step is the significant "trick" in the invention (judgment [10]). The EPL ANN is then trained to make the distances between pairs of the property co-ordinates converge or diverge in alignment with the distancing between that pair in the semantic space. Thus if the initial property space co-ordinates are farther apart than those in the semantic space, they are moved closer together, and conversely if the initial distance is close together in the property space the distance is increased to reflect semantic dissimilarity. This pairwise comparison based training is achieved by backpropagation, in which the error in the property space is corrected in order to make the results coincide with the training objectives. The correction is achieved by the adjustment of the network parameters in the EPL ANN. This is step (b) of claim 4.

24.　　The process is done repeatedly and the result, as the judge explained in [11], is that the system can now provide a single vector in property space for any given track of music which will have a degree of semantic similarity to other music tracks which is reflected in their relative property vector proximity. Similar semantic styles will be reflected in the property vectors being closer together; dissimilar styles will be reflected in the vectors being farther apart. The EPL ANN has learned how to discern semantic (dis)similarity from physical properties.

25.　　The EPL ANN is now ready to take any given track of music provided or proposed by a remote user, determine its physical properties and attribute a property or physical vector to it. It can then relate that vector to the vectors of files in an overall database from which it is to make recommendations. The effect is to identify music which is semantically similar by looking for tracks with proximate physical vectors, and make a recommendation of a similar track from those nearby vectors. It completes the task by sending a message and a file to the remote user.

26.　　The advantage of this over other systems for providing recommendations of similar music to users is described in the Hearing Officer's Decision in the following terms, which are not disputed on appeal here or below:

　　　　"49. At this point it is helpful to turn to the main piece of prior art identified by the examiner on the basis of the searching conducted so far, US 2018/0349492 A1. Much discussion of this document was provided in the skeleton arguments, in Professor Pardoe's report, and again at the hearing. It generally discloses training an ANN-based system to label media items with relevant contexts which can be used to generate playlists themed around those contexts. Several differences between this

document and the claimed invention are identified by the applicant, not least of which is the lack of pairwise comparisons of the property and semantic vectors of files to provide convergence of semantically similar files in property space during the ANN training stage. Further, the prior art requires a larger number of ANNs in both the training and inference stages as compared to the claimed invention. The claimed invention is said to be simpler and faster as a result. I am willing to accept these alleged differences and advantages over the prior art."

27.   It is useful to see how these advantages are described although it is worth noting that this is not a case concerned with inventive step.

The law in outline

28.   Section 1(2) of the Patents Act 1977 provides as follows:

It is hereby declared that the following (among other things) are not inventions for the purposes of this Act, that is to say, anything which consists of—

(a) a discovery, scientific theory or mathematical method;

(b) a literary, dramatic, musical or artistic work or any other aesthetic creation whatsoever;

(c) a scheme, rule or method for performing a mental act, playing a game or doing business, or a program for a computer;

(d) the presentation of information;

but the foregoing provision shall prevent anything from being treated as an invention for the purposes of this Act only to the extent that a patent or application for a patent relates to that thing as such.

29.   Sections 1(1) and (2) of the Act implement Article 52 of the European Patent Convention (EPC). The drafting differs a little between the two provisions but they have the same effect.

30.   A full consideration of the law on how to apply these provisions both here and by reference to EPO and other states' law is not necessary for the purposes of this appeal. The two cases generally cited at this stage are *Aerotel Ltd v Telco Holdings Ltd* [2007] RPC 7 and *AT&T Knowledge Ventures v Comptroller* [2009] EWHC 343 (Pat).

31.   The legislative history such as it is, the position elsewhere and the UK cases up to 2006 were reviewed in detail by the Court of Appeal in *Aerotel* (see the judgment of the court given by Jacob LJ at [6] to [49] and also the case law appendix from [78]). The Court of Appeal there set out a four stage approach to the application of these exclusions. The four *Aerotel* stages are:

(1) Properly construe the claim.

(2) Identify the actual contribution (although at the application stage this might have to be the alleged contribution).

(3) Ask whether it falls solely within the excluded matter.

(4) If the third step has not covered it, check whether the actual or alleged contribution is actually technical.

32.   The approach can be summarised loosely as being to work out if the claimed invention makes a contribution which is technical in nature. In saying this one has to recognise: first that in this context the "contribution" is not the same thing as the inventive step; and second that the mere fact that computers are involved (which are as technical in nature as one could ever imagine) does not make the contribution technical.

33.   At the heart of the law is the consistent principle that an inventor must make a contribution to the art (that is to say the invention must be new and not obvious) and that contribution must be technical in nature (susceptible of industrial application and not within one of the areas excluded by Art.52(2)). For the provenance of these words, see Kitchin J in *Crawford's Application* [2006] RPC 11 and then a decision of mine in *Re Halliburton Energy Services* [2011] EWHC 2508 (Pat) at [27]. I believe they are an accurate statement of the law.

34.   In *AT&T Knowledge Ventures* Lewison J had identified five signposts to use when considering whether a computer program makes a technical contribution. While that decision was in the High Court, the signposts were endorsed (and recast very slightly) by the Court of Appeal in *HTC v Apple* [2013] EWCA Civ 451. In the right cases – which are often difficult ones and are generally known as "better computer" cases because the invention is said to improve the functioning of the computer system itself – these signposts are very helpful.

35.   Finally at this stage it is relevant to note that it was not suggested that we could or should depart from the existing English case law in the light of decision G1/19 of the Enlarged Board of Appeal of the EPO on 10th March 2021.

36.   I will come back to the law below when addressing the grounds of appeal.

The course of the proceedings

37.   Both the judge and the Hearing Officer recognised the four stage *Aerotel* approach, sought to apply it and also noted the five signposts of *AT&T Knowledge Ventures*. The issues of construction at *Aerotel* stage 1 were all minor matters before the Hearing Officer. As the judge held at [22], the Hearing officer had determined (at HO[41]) that the ANN used in the method of claim 4, which I have called the EPL ANN, can be implemented in software or hardware.

38.   At *Aerotel* stage 2, again as the judge recognised at [24] the Hearing Officer identified the contribution at HO[53] in the following way:

> 53. "...the invention of the Application is an ANN-based system
> for providing improved file recommendations. The invention
> may be hardware or software implemented. The fundamental
> insight is in the training of the ANN which analyses the physical
> properties of the file by pairwise comparisons of training files.
> In these pairwise comparisons the distance in property space

between the output (property) vectors of the ANN is converged
to reflect the differences in semantic space between the semantic
vectors of each pair of files. The result is that in the trained ANN,
files clustered close together in property space will in fact have
similar semantic characteristics, and those far apart in property
space will have dissimilar semantic characteristics. Once trained
the trained ANN can then be used to identify, swiftly and
accurately, files from a database which correspond semantically
to a target file, and to provide - against [sic] swiftly and
accurately - file recommendations to a user device (over a
communication network)."

39.   That statement of the contribution was also common ground before this court.

40.   The Hearing Officer considered *Aerotel* stages 3 and 4 together. At that point in the
        proceedings the applicant EPL advanced three arguments: one starting from a hardware
        ANN, a second based on *Halliburton* (see below), and a third based on *Protecting Kids*
        (see also below)

41.   The first argument had three stages. First, it was said that a hardware implementation
        of an ANN would be outside the computer program exclusion altogether. Second, the
        applicant made the point that the computer program exclusion is not supposed to
        operate to exclude inventions which would otherwise be patentable but for their
        implementation as software. Third, therefore, it was said, a software implemented
        ANN ought not to be excluded by the computer program exclusion either.

42.   On this basis the applicant submitted that the claim, which covers both hardware and
        software ANNs, was not affected by the computer program exclusion. The argument
        sought to draw a distinction between the choices made by a human programmer to
        define the problem and the training method, as opposed to the generation of the network
        parameters which happens in the training process. The latter was described by
        Professor Pardoe as the creation of an internal model independent of the software
        programmer and independent of both the expression or language chosen by the
        programmer.

43.   The Hearing Officer rejected this first argument from HO[61] onwards, rejecting the
        idea that the training is a process entirely independent of any instruction from a
        programmer. The Hearing Officer also held at HO[63] that even if a software
        implemented ANN could be decoupled, for the purposes of applying the law, from the
        platform which supports it, nevertheless the mathematical method exclusion, which is
        also in s1(2), would lead to the exclusion of this claim from patentability at step 3 or
        step 4, albeit that in a later passage (HO[78]) the Hearing Officer does appear to indicate
        that he thought the mathematical method exclusion would not defeat the application
        (addressed below).

44.   The second argument was based on posing the question asked in *Halliburton*, at [38] of
        that judgment, which drew attention to the utility of asking: what task is it that is
        performed by the program? If that task is specific and external to the computer and is
        not within the excluded areas, that is an indication (although not determinative) that the
        invention is likely to be patentable.

45.   On this second argument the focus was on the step in claim 4 of actually sending a file (which is the recommended music track) over a network. This could be said to be a concrete task which the system performed. However the Hearing Officer held that while this was more than the standard transmission of a file in a network, what distinguished it from standard file transmission was that the file was a better recommendation, e.g. a song which the user was likely to enjoy. The problem for the applicant was, as held in the last sentence of at HO[69] that "… such a beneficial effect is of a subjective and cognitive nature and does not suggest there is any technical effect over and above the running of a program on a computer." Therefore EPL's second argument foundered.

46.   The third argument was also focussed on the sending of a better recommendation message and sought to draw an analogy with a judgment of Floyd J in the High Court in *Protecting Kids the World Over* [2011] EWHC 2720 (Pat). There Floyd J had identified an effect of the invention as the technically superior monitoring of electronic communications, and held the invention was patentable. However the Hearing Officer decided that the similarity between that and the present case was superficial only and (HO[72]) there was nothing like that here. The improved identification and recommendation of files here was "based on their semantic similarity, which is not a relevant technical effect." The Hearing Officer also considered and rejected a similar argument based on an analogy with *Gemstar v Virgin Media* [2009] EWHC 3068 (Ch) at HO[73].

47.   Finally the Hearing Officer at HO[75] – HO[77] considered the *AT&T* signposts but found no reason from there to see a technical contribution in the present case. He concluded at HO[79] that the contribution in the present case falls solely within the computer program exclusion and that "the ANN-based system for providing semantically similar file recommendations is not technical in nature".

48.   On appeal to the High Court, the judge, Sir Anthony Mann, understood that before him the Comptroller was arguing that a hardware ANN was a computer but it was a computer with no program. Therefore there was no relevant computer program to which the exclusion applied and so the Comptroller was accepting that if the patent application had been confined to a hardware implemented ANN, then it could not have been excluded from patentability (see e.g. [36] in which various dictionary definitions of "computer program" were noted and see the conclusion at [43]). Jumping ahead, in this court there was a submission by the Comptroller that the judge had here misunderstood part of the Comptroller's submissions below. I have my doubts that the judge did misunderstand what was put to him but there is no need to examine that matter now because EPL helpfully did not object to the manner in which counsel for the Comptroller put the Comptroller's case before this court.

49.   Turning back to the High Court judgment, having focussed on the hardware ANN the judgment then focussed on a software ANN. The judge held (at [49]) that a computer program is involved at the training stage (and noted that counsel for EPL had accepted that). Then focussing on the operation of the trained software emulated ANN, the judge held at [54] that what he understood the Comptroller's concession about hardware ANNs to mean was that the operation of a trained hardware ANN did not involve a computer program, because it was not implementing a series of instructions pre-ordained by a human. By contrast it was operating according to something it had learned for itself. In this paragraph the judge indicated he could not see why the same

should not apply to a software emulated ANN (a point I sympathise with when it is put that way). At [56] the judge decided that it is appropriate to look at a software ANN as in substance operating at a different level (metaphorically) from the underlying software on the computer and so, if a hardware ANN is not operating a program then neither is a software ANN working in the same way. Therefore [58] holds that a software ANN is not a program for a computer at all.

50.   The end of this part of the judgment addressed the training stage, which was accepted to involve some programming activity. The judge held that that program, such as it is, is a subsidiary part of the claim and is not what is claimed. Therefore the exclusion is not invoked as a result of the training process ([60]-[61]).

51.   The judgment then turned at [63] to apply the technical contribution approach, in case the conclusion that the computer program exclusion is not invoked at all is wrong. The issue was approached focussing on the sending of an improved recommendation message, asking whether it involved a technical effect, acknowledging ([69]) that the meaning of the word technical is elusive and the boundaries imprecise. The judgment refers to a number of familiar cases (including *Vicom*, *Symbian*, *Halliburton*, *Gemstar* and *Protecting Kids*) and decides the matter at [76]. The findings are that the Hearing Officer was right to identify that the transmission of the file was an effect external to the computer but wrong to hold that a subjective appreciation of the output of the system "was just that, subjective and in the user, and therefore not a technical effect". The output file is a file identified as being semantically similar by the application of technical criteria which the system worked out for itself. It is "not just any old file", the output is a technical effect outside the computer and when coupled with the purpose and method of selection it fulfils the requirement of a technical effect to escape the exclusion. The (music) file goes on to have an effect on the user if the thing works at all, but it would not matter if the user never listened to the file. The file with its similarity characteristics is still produced. Therefore the system is not excluded.

52.   Finally the judge dealt with two points. First he considered the matter on the footing that the relevant computer program which engaged the exclusion is the training program or training stage. There is no need to deal with that because that question does not arise on appeal. It cannot help either way, in other words, put shortly, if the trained ANN is excluded by s1(2) then this argument cannot save it, while if the trained ANN is not excluded then this argument cannot lead to it being excluded.

53.   Second the judge declined to address the mathematical method point because the Hearing Officer at HO[78] rejected it and the Comptroller had not filed a respondent's notice ([79]-[83]).

On appeal to this court

54.   The Comptroller appeals to this court, with the judge's leave, on four grounds:

   i)   Ground 1: the Judge erred in holding that the exclusion from patent protection for "a program for a computer … as such" was not engaged;

   ii)   Ground 2: the Judge was wrong to rely on the Appellant's 'concession' that a hardware ANN was a computer but it was a computer with no program, or words to that effect;

iii)   Ground 3: the Judge was wrong to exclude the consideration of the mathematical model exclusion; and

iv)   Ground 4: the Judge was wrong to hold that the claimed invention involves a substantive technical contribution.

55.   Given the way the case has developed in this court, the convenient way to deal with Grounds 1 and 2 is to deal with them by answering what I called the first question. That involves asking what a computer program is and whether there is a computer program in an ANN. If there is then the second question arises – which is addressed as Ground 4. Finally Ground 3, the mathematical method exclusion, only arises if the Comptroller has lost on Grounds 1 and 2. The respondent contests Ground 3 on its merits but does not advance a procedural objection.

The first question (Grounds 1 and 2)

56.   An ANN is unlike what I have already called a conventional computer, by which I mean a computer of the normal sort most people are familiar with. A technical person might say one needs to get into questions of Von Neumann architecture or to draw a distinction between neural networks and Turing machines, but neither party did that and there is no reason to do so here. The Comptroller's submission is that an ANN is a computer, albeit of a relatively unfamiliar kind.

57.   The Comptroller then submits that in order to customise an ANN for a particular task the set of weights and biases (which again I will refer to simply as "weights") have to be configured appropriately, and it is that set of weights which forms the program of this kind of computer. The Comptroller contends that this submission, that ANN weights and biases are a computer program, accords with the definitions of that term which were quoted by the judge at [36]. The four definitions in the Cambridge, Collins English, Macmillan and The Free dictionaries were:

> "a set of instructions that makes a computer do a particular thing"; [Cambridge Dictionary]

> "a set of instructions for a computer to perform some task;" [Collins English Dictionary]

> "a set of instructions stored inside a computer that allows the user to do a particular thing …" [Macmillan Dictionary]

> "(computer science) a sequence of instructions that a computer can interpret and execute;" [The Free Dictionary]

58.   The respondent agrees with these definitions (albeit arguing that the Comptroller's counsel's oral submissions moved away from them) and submits that an ANN is not a computer program and, more specifically the weights of an ANN (whether hardware or software) are not a computer program. The respondent's case is that a computer program takes the form of serial logical 'if-then' type statements defined by a human programmer and which define exactly what it is that the programmed computer does. Therefore, it is said, the weights of an ANN are not a computer program. Paragraph 17

of the respondent's appeal skeleton sets out more details of the respondent's case as follows:

> "17. The core utility of ANNs (including the ANNs of the Application) lies in their ability to address problems which would be intractable to computer programming. To write a computer program requires the programmer to understand the problem at hand and the manner of its solution, from which to formulate a series of logical commands for the computer to follow. Where the problem is itself intractable, then a computer programmer (and computer program) cannot help – how, rhetorically, can a programmer write a program when the solution to the problem is not even understood? By contrast an ANN is a machine-based system which is able through iterative training on a (usually very extensive) data set to create for itself an internal structure which solves the otherwise intractable problem. The solution to the problem is embedded into the structure of the ANN, i.e., in its links, nodes, weights and biases; that structure being the result of iterative changes made during the training process. In many cases, certainly including those of the Application, there may be enormous insight in how the training objectives and the training data are used to cause the ANN in question to evolve during training; but even then the computer scientist does not know in advance what structure the ANN will adopt nor what patterns and relationships the ANN will ultimately pick up in the data, and indeed it is normally impossible even once the ANN is trained to understand how it is approaching the problem in question to produce the answers given."

Assessment

59.   The meaning of "program for a computer" as it appears in s1(2) of the 1977 Act is a question of law. It is answered bearing in mind that this provision is intended to correspond to Art 52 EPC. Neither party suggested that any preparatory material either for the EPC or the Act, illuminated the meaning of that term. The court in *Aerotel* noted the difficulty in finding anything useful in the travaux which shed light on this exclusion or which identified an underlying purpose against which to interpret the law. Also relevant is Art 27 of the TRIPS treaty which provides that patents are to be granted in "all fields of technology", but neither party suggested that this made any difference in the present case.

60.   The EPC was signed in October 1973 to come into force in 1977. The United Kingdom was one of the six original signatories. Although at times in the argument there seemed to be a suggestion that ANNs were new and so could not have been in the legislator's mind in 1973, neither side actually made any submission that the term as understood in 1973 might be different from the way it would be understood today. I mention this because there was no submission that the doctrine that a statute is "always speaking" helped in this case. I refer to *News Corp UK & Ireland Ltd v HMRC* [2023] UKSC 7 in which the Supreme Court identified the doctrine when it grappled with a problem of interpreting a statute from some time ago when technology had moved on. The case

concerned the question how to apply a VAT statute from 1994 to digital newspapers, when newspapers in 1994 were things made of paper. As I say neither party submitted that the "always speaking" principle would help to answer the question in the present case. Nevertheless I have found assistance in the distinction between meaning and reference drawn by Lord Leggatt in his concurring judgment in *News Corp* at [94]-[95]. The point is that the meaning of a statutory expression does not change whereas the class of things which it covers may do so. Something which did not exist when an Act was passed and therefore could not have been identified as being within the Act at the time, may still be covered by that Act today once its meaning has been understood.

61.   I start with the term computer. I would hold that a computer is a machine which processes information. Neither party came up with a better definition and I believe that is a useful one. Turning to computer program, (which is the same thing as a "program for a computer"), in terms of the meaning of a statute, dictionary definitions are not determinative but in this case I think the definitions are helpful. I would hold that a computer program is a set of instructions for a computer to do something. These two definitions work together, so one can say that a computer is a machine which does something, and that thing it does is to process information in a particular way. The program is the set of instructions which cause the machine to process the information in that particular way, rather than in another way.

62.   This focus on a program as instructions is consistent with the approach of the Court of Appeal in *Gale's Application* [1991] RPC 305, at p321 (ln 13-19), in which Nicholls LJ, who was considering what a program was in the context of a case about a conventional sort of computer, noted that "program" was a flexible term and that "a sequence of instructions" was called a program.

63.   It is also consistent with the approach of the Court of Appeal in *Aerotel* at [31], in which Jacob LJ, giving the judgment of the court, described a computer program as a "set of instructions". This was in the context of a debate whether the term was limited to the set of instructions in the abstract or included the instructions on some form of media (referring back to *Gale*) and preferring the latter.

64.   Much of EPL's argument here sought to add various limitations into the definition. The first limitation related to the involvement of a human computer programmer. I do not believe that referring to a human programmer is relevant or helpful. I can think of no principle which would justify that as a necessary aspect of the definition and the authorities in this area have never drawn a distinction of that kind. The code which human programmers write for conventional computers is written in a form which is sometimes called a high level programming language. That is a form which human programmers can understand and grapple with. However, as the Comptroller submitted, ordinary computers work by running machine code, which is different and hard for humans to understand. The machine code is derived by a computer system (normally what is called a compiler program) under the direction of a human programmer. There is no justification for drawing a distinction in law between instructions created by a computer and those created by a human.

65.   Nor do I accept that focussing on the characteristics of the problem the programmer wants to solve (tractable or intractable) is relevant or helpful either. The fact that ANNs aim to solve problems which are not easy to solve with conventional computers is

irrelevant. Both conventional computers and ANNs can (aim to) solve problems which are difficult for humans to solve unaided.

66.   The respondent puts weight on the fact that the particular values for the weights are produced by a training process in which the machine learns for itself, but I do not see how that can be relevant either. This argument is related to the two previous arguments in that it is focussed on the manner in which the instructions are produced. As I have said I do not accept there is justification for that either in principle or in the Act (or the international conventions: EPC or TRIPS). How the program came into being is irrelevant.

67.   Another distinction which I believe is irrelevant relates to permanence. There are some computers with programs which cannot be changed – e.g. the chips embedded in a payment card or a washing machine – but it remains meaningful to draw the same distinction between the program in that case and the computer itself. Whether the program for a given computer is fixed in a permanent form or not does not, in my judgment, alter the fact that the program represents a set of instructions for a computer to do something. The result in *Gale*, which involved rejecting a distinction between the permanence of instructions in ROM circuitry as opposed to those stored in other media would have been quite different if this distinction was relevant.

68.   Turning to an ANN, the first point to make is that however it is implemented, such a machine is clearly a computer – it is a machine for processing information. Focussing on the weights of an ANN, in my judgment irrespective of the manner in which an ANN is implemented (hardware or software), the Comptroller is right that these weights are a computer program. They are a set of instructions for a computer to do something. For a given machine, a different set of weights will cause the machine to process information in a different way. The fact the set does not take the form of a logical series of 'if-then' type statements is irrelevant. The weights for a given artificial neuron are what cause the neuron, if the inputs are of a given type, to then produce an output of a given type. Aggregated up to the ANN as a whole, these weights work that way in parallel with one another to a significant extent and not just in a logical series, but that is not a relevant distinction. The set of weights as a whole instruct the machine to process information it is presented with in a particular way.

69.   It is notable that the Technical Boards of Appeal of the EPO take the same approach: see decision T 702/20 *Mitsubishi/Sparsely connected neural network* at [10] and [11]. Here the Board of Appeal applied exactly the same approach to a case about an ANN as it applies to other computer implemented inventions. At [10] the Board held explicitly that since "a neural network relates to both programs for computers and to mathematical methods", the question was whether it related only to such subject-matter "as such" or whether there was something more, i.e. something that can fulfil the patentability conditions of the EPC.

70.   Therefore the exclusion from patentability of a program for a computer as such in s1(2) of the 1977 Act is engaged in this case. Nor is there any difference for this purpose between a hardware ANN and a software ANN. However it is implemented, the weights (by which I mean weights and biases) of the ANN are a program for a computer and therefore within the purview of the exclusion.

The second question (Ground 4)

71.   It is worth emphasising that the fact that s1(2) of the Act is engaged in a case of an ANN implemented invention, as much as it would be in any computer implemented invention, does not mean it is unpatentable. Very many computer implemented inventions are outside the exclusion and are patentable as a result. A computer implemented method controlling an X-ray machine was patentable in *Koch v Sterzel* (T26/86), as was a computer system for designing drill bits in *Halliburton*, and a system presenting a new interface to application programmers writing software for multi-touch devices in *HTC v Apple*. Each of these would have been just as patentable if the computer involved had been or used an ANN. Conversely the conclusions that a computer implemented financial trading system (*Merrill Lynch*) was excluded or a computer set up to produce the documents needed to form a company (the *Macrossan* case decided in *Aerotel*), would also be the same if an ANN was involved. The fact the exclusion is engaged as a result of the first part of this appeal, simply means that ANN implemented inventions are in no better and no worse position than other computer implemented inventions.

72.   To recap on the case itself, the Hearing Officer held that the claims were excluded by s1(2) while the judge, in case he was wrong that s1(2) was not engaged at all, addressed that question and held that if the section was engaged, the claims were not excluded.

73.   No issues of claim construction arise (*Aerotel* step 1), the claim clearly covers both hardware and software ANNs. Thus the analysis can turn to the contribution (*Aerotel* step 2). That is not in dispute and is set out at [37] above.

74.   As in many cases *Aerotel* steps 3 and 4 can be taken together. As the judge did, I regard the training activity, which is set out in the claim, as subsidiary in nature and irrelevant. In saying that I do not mean to downplay the importance overall of the "trick", the pairwise comparison technique which is used in the training phase to produce a useful system. It is clearly part of the contribution. However for the purpose of analysing the patentability of either claim, the training aspect makes no difference. The training is, in effect, part of the creation of the program.

75.   Subject only to the step of sending the recommended file to a user, the whole of the remainder of the contribution consists of a program for a computer, and the mere involvement of a computer does not help. Therefore the focus, both before the Hearing Officer and the judge, was on that step of sending a recommended file or, putting it in different words, a recommendation message. I agree with that approach although one does also need to have in mind that the provision of a recommendation message is the presentation of information - which is also unpatentable subject matter, unless it involves a technical contribution.

76.   Put simply, the program here provides improved file recommendations. That is what it does. The correct characterisation of that function should, in this case, provide the answer to the question of patentability. This also illustrates why the present case is not concerned with improvements inside the running of a computer. It is not a better computer case.

77.   The mere fact a file is actually sent does not help. It is a concrete task which the system performed but, as the Hearing Officer held, while that is more than the standard transmission of a file in a network, what distinguishes it from standard file transmission is that the file represents a better recommendation, e.g. a song which the user was likely

to enjoy. In other words, one comes back again to the correct characterisation of the function of this computer program in this case.

78.   The judge also approached this issue in the same way (see [63], [68], [74] and [76]). The issue boils down to whether the Hearing Officer was right to find the exclusion applied because the beneficial effect was of a subjective and cognitive nature (HO[69]) or whether the judge was right to hold as he did in [76] that the exclusion did not apply because even though what made the file recommendation better was not technical criteria (because the semantic similarity is a subjective matter) the ANN had reached that result by "going about its analysis and selection in a technical way". As the judge put it in that paragraph: "It is not just any old file; it is a file identified as being semantically similar by the application of technical criteria [...]". The sentence ended making the point that the system had worked out the criteria for itself but that aspect cannot now be relevant given that ANNs, which are set up by training, are within the ambit of the exclusion.

79.   In my judgment the Hearing Officer's conclusion is the right one. What makes the recommended file worth recommending are its semantic qualities. This is a matter of aesthetics or, in the language used by the Hearing Officer, they are subjective and cognitive in nature. They are not technical and do not turn this into a system which produces a technical effect outside the excluded subject matter. I note that the same view was expressed by the Technical Board of Appeal of the EPO in *Yahoo* T 0306/10, at paragraph 5.2 in holding whether song recommendations are "good" or "bad" does not amount to a technical effect. EPL make the point that this case was concerned with inventive step but that is only an artefact of the difference in the way the EPO approaches patentability from the manner in which it is approached in this jurisdiction. It does not undermine the relevance of the Board's observation.

80.   It is true that as the judge said, the system has gone about its analysis and selection in a technical way but that is because it is an ANN, i.e. a computer. The fact the computer is using properties it can measure to make this semantic recommendation makes no difference. I think the flaw is that this approach imports the undoubtedly technical nature of computer systems (including ANNs) into the analysis. If that was appropriate then the same could be said of the other cases of excluded matter such as the computer implemented financial trading system of *Merrill Lynch*.

81.   It is the semantic similarity of the files here which gives rise to their recommendation but that is not a technical matter at all. Putting it another way the similarity or difference between the two files is semantic in nature and not technical. I agree with the Hearing Officer that the similarity between this case and the one addressed by Floyd J in *Protecting Kids* is superficial only and also that no useful analogy can be drawn from the patent in *Gemstar* which was held not to be excluded. The fact that in the present case there is what one might call an external transfer of data (the file recommendation) does not help for the same reason. What matters is the correct characterisation of the data being transferred and that brings the issue back to the aesthetic and therefore non-technical quality of this aspect of the contribution.

82.   Finally, in my judgment, considering the signposts in this case do not assist EPL.

83.   I would allow this appeal on Ground 4 and uphold the decision of the Hearing Officer that this application is excluded from patentability.

The mathematical method exclusion (Ground 3)

84.   There is no need to consider that aspect of this appeal. I will only say that I think this objection might well have had traction if the conclusion was that weights and biases of an ANN were not a computer program. It is hard to see why even if they are not to be regarded as a computer program for some reason, they are not in any case a mathematical method and so the very same analysis based on the *Aerotel* approach would apply with the same result. I note that the EPO in *Mitsubishi* also took the same view and regarded the mathematical method exclusion as relevant.

Lord Justice Arnold:

85.   I agree.

Lady Justice Nicola Davies:

86.   I also agree.

Appendix - the principal claims

Claim 1

1 . A system for providing semantically relevant file recommendations, the system containing:

a) an artificial neural network "ANN" having an output capable of generating a property vector in property space, the ANN trained by subjecting the ANN to a multiplicity of pairs of training data files sharing a content modality and where for each pair of training data files there are two independently derived separation distances, namely:

a first independently derived separation distance that expresses a measure of relative distance between a first pair of training data files in semantic embedding space, where the first independently derived separation distance is obtained from natural language processing "NLP" of a semantic description of the nature of the data associated with each one of the first pair of training data files; and

a second independently derived separation distance that expresses a measure of relative distance similarity between the first pair of training data files in property embedding space, where the second independently derived separation distance is a property distance derived from measurable properties extracted from each one of the first pair of training data files, and

wherein training of the ANN by a backpropagation process uses output vectors generated at the output of the ANN from processing of said multiplicity of pairs to adjust weighting factors to adapt the ANN during training to converge distances of generated output vectors, in property embedding space, towards corresponding pairwise semantic distances in semantic space, and

wherein shared content modality is: (i) video data files; or alternatively (ii) audio data files; or alternatively (iii) static image files; or alternatively (iv) text files; and

b) a database in which is stored a multiplicity of reference data files with content modality with target data and a stored association between each reference data file and a related individual property vector, wherein each related individual property vector is obtained from processing, within the trained ANN, of file properties extracted from its respective reference data file and each related individual property vector encodes the semantic description of its respective reference data file;

c) a communications network;

d) a network-connected user device coupled to the communications network;

e) processing intelligence arranged:

in response to the trained ANN receiving target data as an input and for which target data an assessment of relative semantic similarity of its content is to be made, and the ANN producing a file vector (Vpiie) in property space for the target data based on processing within the trained ANN of file properties extracted from the target data;

to access the database;

to compare the file vector of the target data with individual property vectors of the multiplicity of reference data files in the database to produce an ordered list which identifies relevant reference data files that have property vectors measurably similar to the property vector and thus to identify relevant reference files that are semantically similar to the target data; and

to send, over the communications network, relevant reference files to the user device;

wherein

the user device is arranged to receive the relevant reference files and to output the content thereof.

Claim 4

4. A method of providing semantically relevant file recommendations in a system including an artificial neutral network "ANN" having an output capable of generating a property vector in property space, the method comprising:

a) training the ANN by subjecting the ANN to a multiplicity of pairs of training data files sharing a content modality and where for each pair of training data files there are two independently derived separation distances, namely:

a first independently derived separation distance that expresses a measure of relative distance between a first pair of training data files in semantic embedding space, where the first independently derived separation distance is obtained from natural language processing "NLP" of a semantic description of the nature of the data associated with each one of the first pair of training data files; and

a second independently derived separation distance that expresses a measure of relative distance similarity between the first pair of training data files in property embedding space, where the second

independently derived separation distance is a property distance derived from measurable properties extracted from each one of the first pair of training data files,

and wherein shared content modality is: (i) video data files; or alternatively (ii) audio data files; or alternatively (iii) static image files; or alternatively (iv) text files;

b) in a backpropagation process in the ANN, using output vectors generated at the output of the ANN from processing of said multiplicity of pairs to adjust weighting factors in the ANN, thereby adapting the ANN during training to converge distances of generated output vectors, in property embedding space, towards corresponding pairwise semantic distances in semantic space, and

b) storing, in a database, a multiplicity of reference data files with content modality with target data and a stored association between each reference data file and a related individual property vector, wherein each related individual property vector is obtained from processing,

within the trained ANN, of file properties extracted from its respective reference data file and each related individual property vector encodes the semantic description of its respective reference data file;

c) in response to the trained ANN receiving target data as an input and for which target data an assessment of relative semantic similarity of its content is to be made, and the ANN producing a file vector (Vnie) in property space for the target data based on processing within the trained ANN of file properties extracted from the target data;

d) comparing the file vector of the target data with individual property vectors of the multiplicity of reference data files in the database to produce an ordered list which identifies relevant reference files that are measurably similar to the property vector and thus identifying relevant reference data files that are semantically similar to the target data;

e) sending, over the communications network, relevant reference data files to the user device; and

f) at the user device, receiving the relevant reference files and outputting the content thereof