Intellectual Property Office

BL O/542/22

22 June 2022

PATENTS ACT 1977

APPLICANT	Emotional Perception AI Limited
ISSUE	Whether patent application GB1904713.3 complies with section 1(2), section 14(3) and section 76(2) of the Patents Act 1977
HEARING OFFICER	Phil Thorpe

DECISION

Introduction

1 Patent application GB1904713.3 entitled "Method of training a neural network to reflect emotional perception and related system and method for categorizing and finding associated content" was filed on 3rd April 2019 in the name of Mashtraxx Limited. The application was published on 4th November 2020 as GB 2583455 A. It has since been assigned to Emotional Perception AI Limited.

2 Despite amendments to the application and several rounds of correspondence, the applicant has not been able to satisfy the examiner that the application meets the requirements of the Act, and so a pre-hearing report was issued on 28th June 2021 in respect of the claims dated 26th May 2021.

3 The matter came before me at a hearing by video conference on 15th March 2022. In the run up to the hearing, an amended set of claims were formally received on 2nd March 2022, alongside supporting arguments and an expert report by Professor Andrew Pardoe. Skeleton arguments and a bundle of authorities were received on 8th March 2022.

4 At the hearing, the applicant was represented by Mr Mark Chacksfield QC and Mr Henry Edwards, instructed by Hepworth Browne. The applicant's attorney, Mr Bruce Dearling, was also in attendance.

The issues to be decided

5 The principal matter to be decided, and that which was the focus of the hearing, is that of exclusion under section 1(2) of the Patents Act 1977. The examiner has maintained throughout that the invention relates to a mathematical method and a program for a computer as such.

6 The examiner's pre-hearing report sets out further objections to added subject matter, clarity and sufficiency. In brief, the applicant's position is that the subsequent amendments to the claims have largely overcome these issues, save for one objection to added subject matter for which additional arguments have been provided. These matters were not discussed in any detail at the hearing, with Mr Chacksfield indicating that a view could be taken on the basis of the arguments in the skeleton. To the extent that these issues are still relevant to the claims under consideration for this decision, I will decide them below.

7 In the skeleton arguments, and again at the hearing, it was suggested by Mr Chacksfield that I also decide the question of novelty and inventive step (at least with respect to the main prior art document identified). As I explained at the hearing, the examiner has indicated that the search is incomplete and has deferred further assessment of novelty and inventive step. An impasse has not been reached on these issues and I will make no finding on them.

The invention

8 The application relates generally to the field of artificial neural networks (ANNs). The invention is concerned with training an ANN to perceive semantic similarity or dissimilarity between media files and using the trained ANN to recommend a file which is semantically similar to a given input. The files may be audio files, video files, static image files, or text files. Much of the specification relates to the audio embodiment, particularly to the field of music recommendation, and so it is convenient to illustrate how the invention works in this context.

9 The process of training the ANN involves making pairwise comparisons between music tracks in the training dataset. For a given pair of tracks, textual descriptions of each track are analysed by natural language processing (NLP) to provide a vector in semantic space. These semantic vectors encode the semantic properties of the tracks, i.e. a representation of the subjective, emotional response to the music as perceived by the human listener. Semantically similar tracks will have semantic vectors that are closer together in semantic space than semantically dissimilar tracks. Each track of the pair is also analysed in terms of its measurable physical properties, such as rhythm, tonality, timbre and/or musical texture. This produces a property vector in property space.

10 During backpropagation, the weights and biases of the ANN are adjusted so that a pair of tracks which are close together in semantic space become close together in property space. Essentially, this process forcibly converges the distance between the generated property vectors towards the distance between the corresponding semantic vectors for the tracks. This results in a system which can identify the semantic similarity between a given pair of music tracks based on an analysis of their measurable physical properties. It ultimately enables a semantically similar music track to be recommended to a user.

11 The claims under consideration at the hearing are those filed on 2 March 2022. The independent claims are as follows:

Claim 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 (V_{File}) 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

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) [sic; there are two parts labelled 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 (V_{File}) 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.

Added subject matter, sufficiency and clarity

12 In their pre-hearing report, the examiner raises two objections to added matter against the previous version of the claims. Section 76(2) states:

> *No amendment of an application for a patent shall be allowed under section 15A(6), 18(3) or 19(1) if it results in the application disclosing matter extending beyond that disclosed in the application as filed.*

13 Guidance on section 76 has been provided in Richardson-Vicks Inc's Patent[1] where Jacob J (as he was then) noted:

> *"The test of added matter is whether a skilled man would, upon looking at the amended specification, learn anything about the invention which he could not learn from the unamended specification."*

and in *Bonzel and Schneider*,[2] where Aldous J (as he was then) stated:

> *"The decision as to whether there was extension of disclosure must be made on a comparison of the two documents read through the eyes of a skilled addressee. The task of the court is threefold:*
> *(1) To ascertain through the eyes of the skilled addressee what is disclosed, both explicitly and implicitly in the application.*
> *(2) To do the same in respect of the patent as granted.*
> *(3) To compare the two disclosures and decide whether any subject matter relevant to the invention has been added whether by deletion or addition. The comparison is strict in the sense that subject matter will be added unless such*

[1] *Richardson-Vicks Inc.'s Patent* [1995] RPC 568
[2] *Bonzel and Schneider (Europe) AG v Intervention Ltd* [1991] RPC 553

matter is clearly and unambiguously disclosed in the application either explicitly or implicitly."

14 The examiner's objections can be summarised as follows:

i) Claim 1 requires that the "two data files (302, 304) shar[e] the same content modality" whereas sharing the same content modality is only disclosed in the application as filed for the specific case of audio files, rather than for other file types. Introducing this feature into claim 1 whilst omitting a limitation to audio files therefore adds matter by intermediate generalisation.

ii) Claim 1 requires that "the subjectively-derived semantic vector [is] generated externally from the ANN" whereas the application as filed was silent as to whether the NLP used for generating the semantic vector occurs within or without an ANN.

15 I can address the second of these quickly as the allegedly offending passage is no longer present in the current claims. I am satisfied that this matter has been resolved.

16 The first is still relevant to the current claim set, albeit the wording has been amended slightly. Claims 1 & 4 both include reference to "pairs of training data files sharing a content modality" and further specify that the "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". In the skeleton arguments, page 16 lines 25-27 of the specification as filed are put forward as providing an explicit teaching of this feature. This passage is preceded by some discussion of an embodiment relating to a music track recommendation tool, and reads:

"The same principles apply to the identification of other contextually describable subjective works that act as a source for computer-implemented data analysis, including images, text and/or video."

17 I cannot agree that this passage alone explicitly discloses the general principle of pairs of training data files sharing a content modality. However, this passage read in context does teach in a general sense that the audio file embodiment can be generalised to other media file types (namely, images, text and/or video which are the subject of the claims). Page 51 line 27 to page 52 line 2 makes a similar teaching about the generalisation to different content modalities:

"It will, of course, be appreciated that the above description has been given by way of example only and that modifications in detail may be made within the scope of the present invention. For example, the principle by which the neural network is trained and how semantically-assessed qualities, indicated by scaled distances, in a sematic vector space can be mapped to an objectively-generated (typically Euclidean) vector in property space can be applied to multiple forms of searchable data, including audio, visual and/or film, literature and scientific reports (such as medical reports requiring cross-referencing for trend analysis)."

18 Further, as noted in the skeleton, claims 38 & 39 as filed disclose, in a general sense, a training process whereby semantic and property vectors are derived from pairs of source files, those source files including at least one of digital audio, image

data, and contextual literary data. In my view, the idea that pairs of training data files share a content modality, which is not necessarily limited to audio, is at least implicitly disclosed by the application as filed. What is now claimed is therefore a fair generalisation to make and does not constitute added matter.

19 Therefore, I do not consider there to be any added subject matter based on the specific objections set out in the examiner's pre-hearing report.

20 The examiner has also raised an objection to sufficiency. The requirement for sufficiency is set out in section 14(3), which reads:

> *The specification of an application shall disclose the invention in a manner which is clear enough and complete enough for the invention to be performed by a person skilled in the art.*

21 The type of sufficiency objection at issue here is that of insufficiency by excessive claim breadth.[3] An application must be sufficient to enable the invention to be worked across the full width by which it is claimed.

22 In summary, the examiner's objection sets out that claim 1 (as it then was) covers the comparison of disparate property vectors. For example, the pair of files may both consist of a video set to music, where the property vector for one file is derived solely from the image content of the video and the property vector for another is derived solely from the audio content of the video. It is argued that comparing such disparate property vectors, i.e. to determine the distance between them in property space, is not enabled by the disclosure and thus the invention is not enabled across the full scope of the claim.

23 The relevant wording of current claims 1 & 4 is as follows:

> *"…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…"*

24 Whilst this wording does not *explicitly* state that the same measurable properties are extracted from each file of the pair, the suggestion that these claims *encompass* the comparison of disparate property vectors generated by extraction of dissimilar properties from each file relies in my opinion on an overly broad reading of the claims. In particular, the person skilled in the art would not construe current claims 1 & 4 in such a broad manner in light of the description and figures.

25 In any case, I am not convinced that the current claims *do* cover an embodiment where the file is a multimedia file, e.g. a video set to music, and disparate properties are extracted from disparate parts of each file. In their current form, the claims specify that the shared content modality between the pair of files is one of video or audio or static image or text. Page 1 lines 7-10 refer to "image data (whether static or video)", which rather suggests that a video should be construed as a subset of image data rather than as a multimedia file including both image and audio content.

---

[3] Also known as Biogen insufficiency after *Biogen Inc v Medeva plc* [1997] RPC 1

Therefore, I am satisfied that the invention of claims 1 & 4 is enabled across its full scope.

26 A number of clarity objections were also raised in sections 9-16 of the examiner's pre-hearing report. These were based on features of the claims that have since been removed by amendment. Therefore, the specific objections raised have fallen away and I make no findings on the clarity of the claims.

Excluded matter – the law

27 The examiner has raised an objection under section 1(2) of the Act that the invention is not patentable because it relates to a category of excluded matter, namely a mathematical method and a program for a computer as such. Section 1(2) of the Act reads 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.*

28 The generally accepted starting point for determining whether an invention falls within the exclusions of section 1(2) is the judgment of the Court of Appeal in *Aerotel/Macrossan*[4], which provides a structured approach in four stages.

29 The interpretation of section 1(2) has been considered by the Court of Appeal in *Symbian*[5]. *Symbian* arose under the computer program exclusion, but as with its previous decision in *Aerotel* the Court gave general guidance on section 1(2). Although the Court approached the question of excluded matter primarily on the basis of whether there was a technical contribution, it nevertheless (at paragraph 59) considered its conclusion in the light of the *Aerotel* approach. The Court was quite clear (see paragraphs 8-15) that the structured four-step approach to the question in *Aerotel* was never intended to be a new departure in domestic law; that it remained bound by its previous decisions, particularly *Merrill Lynch*[6] which rested on whether

---

[4] *Aerotel Ltd v Telco Holdings Ltd and Macrossan's Application* [2006] EWCA Civ 1371; [2007] RPC 7
[5] *Symbian Ltd v Comptroller-General of Patents* [2009] RPC 1
[6] *Merrill Lynch's Application* [1989] RPC 561

the contribution was technical; and that any differences in the two approaches should affect neither the applicable principles nor the outcome in any particular case.

30 The applicant stressed that following the *Aerotel* approach is not obligatory and that, for an invention concerning an ANN, it is important to understand the context in which the *Aerotel* questions are being asked. In support, they cite *HTC v Apple* (at paragraph 44), which sets out that it remains appropriate to follow the *Aerotel* approach but that it is not strictly necessary, and *Symbian* (at paragraph 16), which provides a general warning against blindly following staged approaches. At the hearing, Mr Chacksfield clarified that he was not suggesting that I cannot apply *Aerotel* or that *Aerotel* is not capable of dealing with ANN inventions.

31 The case law above is clear that it remains appropriate to follow the *Aerotel* approach. Whilst the approach may not always be necessary and should not be followed blindly, I think it is appropriate, in this case, to answer the question as to whether the claimed invention makes a technical contribution on the basis of the four-step approach explained at paragraphs 40–48 of *Aerotel*, namely:

(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 Lewison J (as he then was) set out five signposts *AT&T/CVON*[7] that he considered to be helpful when considering whether a computer program makes a technical contribution. In *HTC*[8] the signposts were reformulated slightly in light of the decision in *Gemstar*.[9] The signposts are:

> i. Whether the claimed technical effect has a technical effect on a process which is carried on outside the computer.
>
> ii. Whether the claimed technical effect operates at the level of the architecture of the computer; that is to say whether the effect is produced irrespective of the data being processed or the applications being run.
>
> iii. Whether the claimed technical effect results in the computer being made to operate in a new way.
>
> iv. Whether the program makes the computer a better computer in the sense of running more efficiently and effectively as a computer.
>
> v. Whether the perceived problem is overcome by the claimed invention as opposed to merely being circumvented.

33 It is important to stress that these signposts are just that. They are not barriers or hurdles that need to be individually or collectively overcome by the applicant. They

---

[7] *AT&T Knowledge Venture/CVON Innovations v Comptroller General of Patents* [2009] EWHC 343 (Pat); [2009] FSR 19
[8] *HTC v Apple* [2013] EWCA Civ 451
[9] *Gemstar-TV Guide International Inc v Virgin Media Ltd* [2009] EWHC 3068 (Pat); [2010] RPC 10

are rather a non-exhaustive list of some of the factors that can indicate in some cases whether a particular contribution may be technical.

Excluded matter – analysis

Step 1 – Properly construe the claim

34 There are two independent claims under consideration; system claim 1 and method claim 4. One major point of construction arises on system claim 1. It is a product by process claim in the sense that part a) defines the ANN of the system (a product) in terms of the method by which it has been trained (a process). A claim to a product obtained by a particular process is construed as a claim to the product *per se*. In this case, a claim to an ANN which has been trained by a particular process is no more than a claim to the trained ANN itself, irrespective of the detailed training method by which the trained ANN has been arrived at.

35 At the hearing, Mr Chacksfield acknowledged that claim 1 was a product by process claim but nevertheless felt that it "properly catches the invention". He suggested that if any new point arose as a result of the form of the claim, then method claim 4 could be considered instead. I am minded to proceed with this analysis on the basis of method claim 4, where there is no doubt that the steps of training the ANN are properly captured.

36 The only other point of construction raised at the hearing relates to the step of sending of relevant reference data files to a user device over a network (see part e) of claim 4). I queried whether this involves sending a list of files to the user device or whether it is the files themselves that are sent. Mr Chacksfield submitted that the claim certainly covers sending the file itself but that it may also cover a link to the file. I cannot identify any support for the latter. I have construed this as the sending of the actual file to the device, e.g. the sending of a music track.

37 There is a minor clarity issue within part d) of claim 4, which reads:

> *"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"* (emphasis added).

38 Firstly, it appears that the bolded reference to "the property vector" is actually a reference to "the file vector" of the target data, as distinct from the property vectors of the reference data files. Secondly, it is not the reference files themselves that are "measurably similar" but their property vectors (cf. the wording of part e) of claim 1). As I understand it, this passage intends to set out how the file vector in property space of the target data is compared to the property vectors of the reference files in order to identify reference files that are semantically similar to the target data.

39 The method is clearly computer-implemented, involving a database and a user device in communication over a network. The applicant argues that the skilled person would understand that the ANN could be implemented as a hardware

solution or by way of a software emulation and points to page 16 lines 29-31 in support of this. It reads:

> *"Various aspects and embodiments of the invention as outlined in the appended claims and the following description can be implemented as a hardware solution and/or as software, including downloadable code or a web-based app."*

40 This passage, which is the sole mention of the term "hardware" in the application, is somewhat vague. Figure 8 and the associated section of the description from page 49 line 27 to page 51 line 2 provide further details of the system. They disclose a server 904 which supports the ANN and the system/processing intelligence, and which is in communication with a database and various computing devices over a network such as the internet. The system intelligence may also be distributed or based in the cloud.

41 Thus, the only concrete example given is of an ANN that is seemingly hosted at a server (or servers). The expert report of Professor Pardoe, in setting out the common general knowledge of the skilled person, notes that ANNs can be implemented in software as well as in specialist hardware, including hardware architectures for producing the topologies of ANNs which are commercially available. Professor Pardoe does not comment on whether the ANN of the application would be construed as being implemented in hardware and/or software by the skilled person. In my view, the skilled person would recognise that the ANN could be implemented in whatever technologically conventional way was desired, i.e. certainly as software but perhaps also in hardware. In summary, the method of claim 4 is computer-implemented and the ANN can be implemented in software or hardware as is conventional in the art.

42 Claim 4 can otherwise be construed straightforwardly. It defines a method in which an ANN is trained on pairs of modally identical files (for example pairs of songs) in order to map the distance between their property vectors in property space towards the distance between their semantic vectors in semantic space. When an input (for example representative of a user selected song) is presented to the ANN, it generates a file vector in property space for said input. This file vector is then compared to the property vectors of reference files (representative of a library of songs, say) to identify those files having similar vectors in property space to that of the input. Such files will be semantically similar to the input by virtue of the ANN having been trained to map/converge distances in property space towards distances in semantic space (i.e. so that it generates similar property vectors for semantically similar files). A file can then be sent to, and output by, a user device. In this way, it provides a tool for recommending semantically similar files.

Step 2 – Identify the actual contribution

43 Jacob LJ addressed this step in *Aerotel/Macrossan* where he noted:

> "*43. The second step — identify the contribution — is said to be more problematical. How do you assess the contribution? Mr Birss submits the test is workable — it is an exercise in judgment probably involving the problem said to be solved, how the invention works, what its advantages are. What has the inventor really added to human knowledge perhaps best sums up the exercise."*

44 Jacob LJ goes on to say that in the end:

> *"the test must be what contribution has actually been made, not what the inventor says he has made".*

45 In terms of the problem said to be solved, this principally relates to an alleged inability of prior art artificial intelligence systems to reflect human subjective and emotional responses to stimuli. Across page 1 line 23 to page 3 line 6, the prior art is said to be constrained to absolute similarities, based on measured parameters, rather than semantically perceived similarities. Such models are said to generate inconsistent and/or spurious results and to "overlook semantic similarities whilst accepting or suggesting that perceptually-distinct dissimilarities are closely related". In the skeleton arguments, this is summarised as "the prior art does not give what the human user wishes to receive". To the extent that the human user wishes to receive recommendations based on a more subjective semantic similarity rather than a more objective analysis of measurable properties, this is a fair summary of the main problem.

46 Numerous other problems are discussed across pages 4-7, including the deficiencies of prior art "collaborative filtering" systems which make recommendations based on user data; the issue of "cold-start" where new artists fail to gain traction when recommendations depend on their playlist presence; and the need for fast and reliable recommendations to provide a track finding tool that will not be disregarded or discounted by users.

47 I have already covered how the invention works at the beginning of this decision and again at step 1. Key to this is how the ANN is trained, using the claimed pairwise comparisons, to force distances in property space towards distances in semantic space. It provides a trained system that can recommend semantically similar files by reference to extracted measurable properties.

48 Therefore, it advantageously reflects the subjective, emotional perception of similarity between files. Music tracks in different genres, which by a mechanistic analysis of their measurable properties alone (rhythm, timbre, etc.) appear to be dissimilar, may in fact have semantic similarities as judged by the human listener. The invention is better at recognising these semantic similarities and provides better recommendations as a result.

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.

50 As will be seen, a key pillar of the applicant's arguments relies on the fact that the ANN (and ANNs in general) can be implemented in hardware or in software. I have construed the ANN this way at step 1. However, I reiterate that the application is silent as to any specific details of a hardware implementation. This is not a case where the contribution resides in the hardware *per se*.

51 Further, I note that the contribution does not reside in NLP or the extraction of measurable properties from files *per se*. The applicant freely admits that off-the-shelf software is used for NLP to generate the semantic vectors. The description also suggests the use of off-the-shelf software for performing feature extraction from audio files to determine rhythm, tonality, timbre, etc. (see page 22 line 25 to page 31 line 6). For feature extraction from other media file types, the application is much less detailed such that the skilled person would have to rely heavily on their common general knowledge to implement this (see e.g. page 23 lines 20-21).

52 The applicant has proposed the following statement of the contribution:

> "an improved system for providing a file recommendation, including an improved file recommendation message. The system and message are simpler and faster than the prior art, and provide better recommendations."

53 This suffers from a lack of detail, in particular by missing the key insight of how the ANN is trained. At the hearing, Mr Chacksfield urged me to consider the following expanded summary:

> "…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)."

54 This is a fair way of putting the contribution. The reference to accuracy captures that the system is better at identifying and recommending files to the user based on their semantic similarity.

<u>Steps 3 and 4 – Ask whether it falls solely within the excluded matter and check whether the actual or alleged contribution is actually technical</u>

55 I will consider steps 3 and 4 together.

56 The applicant makes three main arguments at different levels of generality as to why the invention is not excluded. The first and broadest argument is that the computer program exclusion is not even *prima facie* engaged for many ANN claims, including the present. At the hearing, Mr Chacksfield put it as follows:

> *"…you can have an ANN that is hardware implemented or you can have an ANN which is software emulated but it is the same thing, albeit one of them is made of bits of metal and one of them is emulated by a platform. What we say is that in accordance with conventional patentability criteria the software supported version should not be excluded from patentability by the mere fact that for its implementation modern technical means, in the form of a computer program in this case, are used to support it."*

57 It is submitted that a hardware implementation would clearly be outside the computer program exclusion. Thus, since the program exclusion does not operate to exclude inventions that would otherwise be patentable but for their implementation as software, it follows that implementation of the ANN as a software emulation is also non-excluded. This position is summarised as "if the hardware ANNs are patentable, then the related software emulations will also be patentable".

58 As I noted at step 2, this is not a case where there is a contribution to hardware *per se*; standard computer hardware and networks may be used. A contribution to new hardware would likely make a technical contribution, but that is not the case here. The ANN may be implementable in conventional hardware but that does not give the invention the necessary technical character.

59 The applicant's argument, however, is somewhat more nuanced. A distinction is made between so called "classical" computer programs and ANNs. Mr Chacksfield submitted that whereas classical software relies on a series of "if-then" type logical statements defined by a human programmer, an ANN does not. Instead, it uses training data to learn the logic to solve a problem and reconfigures itself accordingly. It is said to provide an internal model which is "independent both of a software programmer and of any instructions a programmer has given the system in code" (emphasis added). This argument has some basis in the expert report of Professor Pardoe, the relevant passage of which reads:

> *"22. … Machine learning eliminates the need to define complex hand-crafted rules that strictly follow a defined specification written by the programmer [as occurs in the development of computer programs] since the abstract machine in ML/AI technology is not processing data on a step-by-step instructional basis, but instead uses training data to learn the logic to solve a specific problem and thereby reconfigures the machine. Machine leaming does not therefore follow an 'if-then' statement approach. On a hardware abstraction level, each iteration is an emulation of a new and better structural topology defined at the end of the process and not the beginning. Rather, the ANN's functional capability is defined through the creation of an internal model independent of the software programmer, and independent of both the expression or language chosen by the software programmer…."*

60 Professor Pardoe does not go as far as to say that an ANN is independent of any instructions a programmer has given the system in code. Plainly, some input from the programmer is still needed to define the problem and the training method to be

used to solve it. Mr Chacksfield admitted as much at the hearing. What the programmer does not need to do is code all of the detailed logical steps of the trained model. Instead, the ANN is adjusted through its training to produce a model which satisfies the training objective that it has been set.

61 In terms of the present invention, the applicant's key insight involves training using pairwise comparisons of files and performing a backpropagation process to adjust weights and biases such that distances between output property vectors are converged towards the corresponding distances in semantic space. I do not believe that they are suggesting that this is a process performed entirely independently of any instruction from the programmer. The programmer defines the problem and the training approach, and the ANN operates within those boundaries to build a suitable model. This is still no more than a computer program in my opinion.

62 Mr Chacksfield freely admitted that there is a software platform underlying the ANN which is a program for a computer, but he made a distinction between an ANN (an emulated machine) and the software platform that supports it. He argued that the invention is not at the level of the software platform but at the level of the ANN above it. Even if this is so, key to the contribution is to specify the training method (pairwise comparison) and objective (converging distances), and this is no more than a computer programming activity.

63 I am not persuaded that the ANN can truly be decoupled from the software platform that supports it in the way Mr Chacksfield suggests. However, even if it can, it is important to consider what an ANN is at this level of generality. It is an abstract model which takes a numerical input, applies a series of mathematical operations (applying weights, biases and an activation function), and outputs a numerical result at successive layers. A claim to an ANN or the algorithm by which it is trained, in a general and abstract sense, relates wholly to a mathematical method and it fails at step 3. Even if there is something more than a mathematical method present, I cannot see how it is technical in nature and so it would not satisfy step 4.

64 Mr Chacksfield made an analogy between software emulation of an ANN and the technical simulation and design processes of *VLSI Chip Design*[10] and *Infineon*.[11] He argued that claims to simulation or design processes are allowable as long as there is an explicit (*VLSI*) or implicit (*Infineon*) link to a production step for the finalised product. For an emulation, he argued that the emulation itself is the final product and so an explicit or implicit link to a production step for the final product is not necessary, but would be sufficient, for it to define a patentable invention. Further, he compared the training of an ANN to an iterative design process in which the final outcome is the software emulation itself rather than a circuit or a silicon chip or the drill bit of *Halliburton*.[12]

65 In contradistinction to the design of a silicon chip or a drill bit and the simulation of a circuit subject to 1/f noise, training an ANN *per se* is an entirely abstract process and the resulting product – the trained ANN – is an abstract model. Emulating an ANN does not, in general, provide a relevant technical effect. An ANN or its training

[10] T 0453/91
[11] T 1227/05 (Circuit simulation I/Infineon Technologies)
[12] *Re Halliburton Energy Services Inc* [2011] EWHC 2508 (Pat)

method may have a relevant technical effect where it is applied to a technical process, however, an application to recommending semantically similar files is not a technical process (as will be discussed in more detail below).

66 The applicant's second strand of argument supposes that the computer program exclusion is relevant and considers the wider system to which the ANN is applied. The applicant sets out, as per *Halliburton*, that where the task carried out by a computer program is not itself something within the excluded categories then it is likely that the technical contribution has been revealed. It is argued that the claimed invention is to a wider technical system for providing file recommendations.

67 It is useful to consider paragraph 38 of *Halliburton* in its entirety:

> *"What if the task performed by the program represents something specific and external to the computer and does not fall within one of the excluded areas? Although it is clear that that is not the end of the enquiry, in my judgment that circumstance is likely to indicate that the invention is patentable. Put in other language, when the task carried out by the computer program is not itself something within the excluded categories then it is likely that the technical contribution has been revealed and thus the invention is patentable. I emphasise the word "likely" rather than "necessarily" because there are no doubt cases in which the task carried out is not within the excluded areas but nevertheless there is no technical contribution at all."*

68 What task then is the program performing? It performs an ANN training stage using pairwise comparisons of files, an ANN inference stage where an input file is analysed and semantically similar files are identified from a database, and it finishes by sending the file to the user over the network. At its core, this is a data analysis and information retrieval task which involves the processing of data within the computer or the computer network.

69 However, where it may represent something specific and external to the computer is in the provision to a user of an improved file recommendation. It is not the mere sending of a file over a network which makes this a process external to the computer – this transmission process is achieved in a standard fashion and operates wholly within the conventional computer network. It is external to the computer in the sense that there is a beneficial effect on the end user in being provided with a better recommendation, such as a song they are likely to enjoy. However, 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.

70 Mr Chacksfield submitted that the claimed invention is similar to that in *Vicom*[13] in the sense that the computer is doing "something clever" on an input to give a particular output. He acknowledged the difference that in *Vicom* the input was an image and the output was an improved image, and in the present case the input is a file and the output is a semantically similar file. In my opinion, this is a crucial difference. In *Vicom*, the program processed data representing an image (a physical entity), resulting in a change to that image. Here, the file itself is not altered in any way. The way that an assessment is made regarding a file's similarity to other files is

[13] T 208/84 (Computer-related invention)

changed (specifically, to value similarity in a more subjective and emotional sense) but this is not a technical process.

71 The applicant's third argument follows on from the second. It relates to the sending of an improved recommendation message over the network. It is submitted that the present invention is on all fours with claim 33 of *Protecting Kids the World Over* (*PKTWO*).[14] In *PKTWO*, the provision of an alarm message over a communication network to a remote terminal alerted a user to the fact that inappropriate content was being processed within the computer. At paragraph 34, Floyd J (as he then was) saw the effect of the invention as an improved monitoring of the content of electronic communications, in a technically superior way, to provide a technical contribution outside the computer. The relevant part of paragraph 34 reads:

> *"…I start with the proposition that the generation and transmission of an alert notification to the user/administrator is not a relevant technical process. I accept that in many cases this may be correct. Plainly it was correct in the case of two out of the three patents considered by Mann J in Gemstar, where information was simply displayed on a screen. But what is in play in the present case, namely an alarm alerting the user, at a remote terminal such as a mobile device, to the fact that inappropriate content is being processed within the computer, is in my judgment qualitatively different. First of all, the concept, although relating to the content of electronic communications, is undoubtedly a physical one rather than an abstract one. In that respect it was more akin to the third of the three patents considered by Mann J in Gemstar. Secondly, the contribution of claim 33 does not simply produce a different display, or merely rely on the output of the computer and its effect on the user. The effect here, viewed as a whole, is an improved monitoring of the content of electronic communications. The monitoring is said to be technically superior to that produced by the prior art. That seems to me to have the necessary characteristics of a technical contribution outside the computer itself."*

72 There is a superficial similarity between *PKTWO* and the present invention in that an alert of sorts (here, the notification of a recommended file) is sent over a network to a user device in response to an analytical task carried out at a computer (here, the identification of semantically similar files with the trained ANN), but that is as far as the similarities go. Floyd J acknowledged that in many cases the generation and transmission of an alert notification to a user is not a relevant technical process. In my opinion, this is one of those many cases. There is nothing at the level of improved monitoring of the content of electronic communications. There is only the improved identification and recommendation of files based on their semantic similarity, which is not a relevant technical effect.

73 In *Gemstar*, to which Floyd J refers, the so-called "transfer patent" involved a dedicated function within an electronic programme guide for controlling the transfer of recorded programme data to a secondary recorder. At paragraph 234, Mann J found the invention provided a relevant technical effect since the initiation of movement of data from one disk to another was a physical effect such that the computer program had an effect in the outside world. In the present case, although it involves the transfer of files to a networked user device, there is simply no contribution to how a data transfer process is controlled or initiated. It is

---

[14] *Protecting Kids the World Over (PKTWO) Ltd, Re* [2011] EWHC 2720 (Pat)

characterised by the content of the information, i.e. sending more semantically relevant file recommendations to the user for output at their device, rather than the provision of a different way to send those files between devices. To paraphrase Floyd J in *PKTWO* above, it merely relies on the output of the computer and its effect on the user.

74 Therefore, the three main arguments advanced by the applicant are not persuasive. The contribution relates wholly to a computer program.

75 Turning to the computer program signposts, the applicant's submissions are brief. They suggest that the signposts are primarily of relevance to the "better computer" cases and that, outside of such cases, at least signposts (ii) to (iv) are not relevant.[15] They submit that signposts (ii) to (iv) do not arise in the present case. To that effect, they have only put forward arguments in respect of the current claims on signposts (i) and (v). I agree that signposts (ii) to (iv) are not of particular relevance here and I am satisfied that they do not assist.

76 The applicant's position on signpost (i) is the same as their second and third arguments, which I have considered above and found not to be persuasive. The generation and transmission of an improved file recommendation has no technical effect on a process outside the computer regardless of whether the process is faster or more accurate. An effect on the end user by way of receiving a semantically similar file, such as a song they might enjoy, is not a relevant technical effect.

77 On signpost (v), Mr Chacksfield submitted that if the problem is providing better recommendations then the claimed invention overcomes it. As I set out at step 2, providing better recommendations (including speed and accuracy) is a fair summary of the main problem. However, it is not a technical problem and the invention derives no technical character from addressing it. I cannot identify any technical problem that has been overcome. Therefore, the signposts do not suggest that a technical contribution is present, consistent with my reasoning above.

78 On the issue of exclusion as a mathematical method, although an ANN and a method of training an ANN *per se* is no more than an abstract mathematical algorithm, its specific application here as part of a file recommendation engine is, in my opinion, enough to dispense with the mathematical method *as such* objection.

79 Finally, taking a step back and looking at the invention as a whole, I am satisfied that the contribution falls solely within the computer program exclusion. The ANN-based system for providing semantically similar file recommendations is not technical in nature.

Post-hearing submissions; amended claims

80 At the hearing, I invited the applicant to provide written submissions on *BDGB*,[16] which relates to a computerised method for classifying text documents using a mathematical algorithm. It is referenced in the section of the EPO Guidelines

[15] Relying on paragraph 23 of *Lenovo (Singapore) PTE Ltd v Comptroller General of Patents* [2020] EWHC 1706 (Pat)
[16] T 1358/09 (Classification/BDGB ENTERPRISE SOFTWARE)

covering artificial intelligence and machine learning, G-II, 3.3.1.[17] The Technical Board of Appeal held that classification of text documents in general does not qualify as a technical purpose and so the mathematical algorithm does not contribute to the technical character of the invention.

81 In their submission, the applicant argues that *BDGB*, as well as *Comptel*[18] and *SAP*[19] to which it refers, can be contrasted with the application in at least two ways. Firstly, in line with their broadest argument above, that an ANN is not a computer program at all but is a physical device or an emulation thereof and has technical character. Secondly, in line with at least their second argument above, that the present claim defines a specific and non-excluded purpose. This is said to be in contrast to the more general document classification methods that are the subject of these three decisions, which encompass non-technical use cases, namely business methods. I have addressed the substance of these arguments already in reaching my decision above. An ANN does not itself have technical character and its application as part of a system for recommending semantically similar files does not make it more than a program for a computer as such.

82 Alongside the written submissions, an amended set of claims were filed with the reference to “text files” having been removed from independent claims 1 & 4. My reasoning above is not predicated on the claimed invention covering the identification and recommendation of similar text files. Therefore, the removal of the reference to text files does not affect the outcome of my decision. Whether the shared content modality is video data files, audio data files, static image files or text files, there is no technical contribution present.

Conclusion

83 Having carefully considered the arguments, I am of the view that the invention falls solely within the matter excluded under section 1(2) as a program for a computer as such. I can see nothing in the specification that could reasonably be expected to form the basis of a valid claim. I therefore refuse this application under section 18(3).

Appeal

84 Any appeal must be lodged within 28 days after the date of this decision.

Phil Thorpe

Deputy Director, acting for the Comptroller

[17] https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_ii_3_3_1.htm
[18] T 1784/06 (Classification method/COMPTEL)
[19] T 1316/09