Intellectual Property Office

BL O/1015/26

20 November 2026

PATENTS ACT 1977

APPLICANT	Jio Platforms Limited
ISSUE	Whether application GB2212707.0 complies with sections 1(1),1(2) and 3 of the Patents Act 1977 (as amended)
HEARING OFFICER	Emma Porter

DECISION

1 Patent application GB2212707.0 ("the application") entitled "System and method for k-nugget discovery and retrofitting framework" was filed on 1st September 2022 claiming priority from Indian patent application IN202121039527, itself having a filing date of 1st September 2021. The application was published on 29th March 2023 as GB2611179 A.

2 A combined Search Report under Section 17(5)(b) and Abbreviated Examination Report under section 18(3) was issued on 18th January 2023. The report contained objections to all of the claims under section 1(2) as the claimed invention was considered to relate to a program for a computer, as such. The examiner did not search the application under section 17(5)(b) since the search would not serve a useful purpose.

3 Since this initial report was issued, there have been multiple rounds of argument, amendment and re-examination. Despite further limitations to the claims, the examiner maintains that the claimed invention is excluded from patentability under section 1(2).

4 The long-standing approach to assessing excluded matter is set out in the Court of Appeal's judgment in *Aerotel/Macrossan*[1]. However, on 11th February 2026, the Supreme Court judgment in *EPAI* [2] established that the test for excluded matter from *Aerotel/Macrossan* should no longer be used and instead set out a new approach for assessing claims that contain a mix of technical and non-technical features. In the

[1] *Aerotel Ltd v Telco Holdings Ltd & Macrossan's Application* [2006] EWCA Civ 1371
[2] *Emotional Perception AI Limited v Comptroller General of Patents, Designs and Trade Marks* [2026] UKSC 3

present case, the *Aerotel* approach had been used in all reports and responses until the first post-*EPAI* examination report, which was dated 10th April 2026.

5 In their examination reports issued after the *EPAI* judgment, the examiner objected that the claimed invention lacked an inventive step in light of prior art that was cited against an equivalent application before the European Patent Office (publication number EP 4396724 A1).

6 The applicant had previously submitted that the present invention had the required technical contribution to be considered non-excluded under the *Aerotel* approach. More recently, using the *EPAI* approach, the applicant submitted that both the technical character and the technical features of the invention are such that the present invention is non-obvious over the cited prior art.

7 With no agreement reached, a hearing was arranged. Under the *EPAI* approach, the matter before me is therefore whether the claimed invention has an inventive step over the prior art.

8 The matter came before me at a hearing on 29th September 2026, held via videoconference. I was assisted by Senior Patent Examiner Mr Andrew Hole. The applicant was represented by patent attorneys Drs Saiful Khan and Andrew Pearson both of Potter Clarkson. Mr Akhil Reddy attended for the applicant and Mr Ranjeev Gupta also attended on behalf of the applicant's Indian counsel.

9 Shortly after the hearing, on 1st October 2026, a UK IPO Hearing Officer issued their ex parte decision[3] regarding the remitted *EPAI* application. As this was only the second UK IPO decision after the new approach was set out by the Supreme Court, the applicant was given the opportunity to make written submissions after the hearing. No written submissions were received.

10 I confirm that in making my decision I have considered all of the documents on file. The initial application and full correspondence may be viewed via Search for a patent - GOV.UK.

The law

Section 130(7) says that:

> *…it is hereby declared that the following provisions of this Act, that is to say, sections 1(1) to (4), 2 to 6, 14(3), (5) and (6), 37(5), 54, 60, 69, 72(1) and (2), 74(4), 82, 83, 100 and 125, are so framed as to have, as nearly as practicable, the same effects in the United Kingdom as the corresponding provisions of the European Patent Convention,…*

11 Article 52 of the European Patent Convention reads:

> *(1) European patents shall be granted for any inventions, in all fields of technology, provided that they are new, involve an inventive step and are susceptible of industrial application.*

[3] https://www.ipo.gov.uk/p-challenge-decision-results/o089426.pdf

(2) The following in particular shall not be regarded as inventions within the meaning of paragraph 1:

(a) discoveries, scientific theories and mathematical methods;

(b) aesthetic creations;

(c) schemes, rules and methods for performing mental acts, playing games or doing business, and programs for computers;

(d) presentations of information.

(3) Paragraph 2 shall exclude the patentability of the subject-matter or activities referred to therein only to the extent to which a European patent application or European patent relates to such subject-matter or activities as such.

12 Section 1(1) sets out the requirement for invention to be patentable (my emphasis added):

A patent may be granted only for an invention in respect of which the following conditions are satisfied, that is to say –

(a) the invention is new;
(b) it involves an inventive step;
(c) it is capable of industrial application;
(d) the grant of a patent for it is not excluded by subsections (2)*and (3) or section 4A below;*

and references in this Act to a patentable invention shall be construed accordingly.

13 Section 1(2) reads (my emphasis added):

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.

14 Inventive step is addressed in Section 3:

An invention shall be taken to involve an inventive step if it is not obvious to a person skilled in the art, having regard to any matter which forms part of the state of the art by virtue only of section 2(2) above (and disregarding section 2(3) above).

15 As mentioned above, the Supreme Court in *EPAI* rejected the *Aerotel* approach and, at paragraph 66, said the approach which should be adopted by courts in the UK is that contained in the *Duns*[4] principles other than principle G. *Duns* principles A-F are:

> *(A) Article 52(1) EPC sets out four requirements to be fulfilled by a patentable invention: there must be an invention, and if there is an invention, it must satisfy the requirements of novelty, inventive step, and industrial applicability.*
>
> *(B) Having technical character is an implicit requisite of an 'invention' within the meaning of Article 52(1) EPC (requirement of 'technicality').*
>
> *(C) Article 52(2) EPC does not exclude from patentability any subject matter or activity having technical character, even if it is related to the items listed in this provision since these items are only excluded 'as such' (Article 52(3) EPC).*
>
> *(D) The four requirements - invention, novelty, inventive step, and susceptibility of industrial application - are essentially separate and independent criteria of patentability, which give rise to concurrent objections. Novelty, in particular, is not a requisite of an invention within the meaning of Article 52(1) EPC, but a separate requirement of patentability.*
>
> *(E) For examining patentability of an invention in respect of a claim, the claim must be construed to determine the technical features of the invention, i.e. the features which contribute to the technical character of the invention.*
>
> *(F) It is legitimate to have a mix of technical and 'non-technical' features appearing in a claim, in which the non-technical features may even form a dominating part of the claimed subject matter. Novelty and inventive step, however, can be based only on technical features, which thus have to be clearly defined in the claim. Non-technical features, to the extent that they do not interact with the technical subject matter of the claim for solving a technical problem, i.e. non-technical features 'as such', do not provide a technical contribution to the prior art and are thus ignored in assessing novelty and inventive step.*

16 In summary, any claim which has some technical character defines an 'invention' under Article 52 (and thus under Section 1 of the Act). This is described as the 'any hardware' test. Following this is a necessary intermediate step of determining which features contribute to the technical character of the invention as a whole. Non-technical features which do not contribute to, or interact with, the technical character

[4] T 0154/04 *(Estimating sales activity/DUNS LICENSING ASSOCIATES)*

of the invention are non-technical "as such" and are therefore filtered out before assessing novelty and inventive step.

17 The intermediate step is discussed in depth in paragraphs 100 to 111 of *EPAI* with reference to paragraphs 32, 33, 38 and 39 of *G 1/19*[5]. It is stated that the intermediate step begins with principle F of *Duns* above and is a necessary part of the process of applying articles 52 to 56 (at least) of the EPC (and hence sections 1 to 3 Patents Act 1977). The intermediate step entails a feature-by-feature analysis of the subject matter of an invention to ensure that only eligible features are considered in the assessment of inventive step. It is not relevant, at the stage of the intermediate step, whether the technical character of the invention appears to be novel or inventive.

18 Paragraph 65 of *EPAI* confirms that the approach to inventive step laid out in *Pozzoli*[6] remains legitimate. That approach is:

> *(1)(a) Identify the notional "person skilled in the art"*
> *(1)(b) Identify the relevant common general knowledge of that person;*
> *(2) Identify the inventive concept of the claim in question or if that cannot readily be done, construe it;*
> *(3) Identify what, if any, differences exist between the matter cited as forming part of the "state of the art" and the inventive concept of the claim or the claims as construed;*
> *(4) Viewed without any knowledge of the alleged invention as claimed, do those differences constitute steps which would have been obvious to the person skilled in the art or do they require any degree of invention?*

19 The Supreme Court in *EPAI* stated, at paragraph 106, that it is open to the UK IPO and to the UK courts to adopt any appropriate method of identifying the technical character of the invention, viewed as a whole, as the necessary first part of carrying out the intermediate step. Only when that has been identified can it be seen whether any particular feature of the invention does or does not contribute to it, or interact with it. The judgment goes on to say that an example of an appropriate method might in a particular case be to identify the alleged "inventive concept" of the claim, as explained in *Pozzoli*. In other cases, it would be appropriate to focus on whether differences between the cited prior art and the inventive concept contribute to the technical character of the invention as a whole (taking care not to consider those differences in isolation).

20 While being clear that the *Aerotel* approach should no longer be followed, the Supreme Court in *EPAI* stated in paragraph 64 that there is nothing wrong with the first step in the *Aerotel* approach of properly construing the claim. Taking all this into account, and in agreement with the approach in FLIR[7], it seems that the approach now has the following steps:

i) Properly construe the claim;
ii) Determine whether the claim involves the use of any hardware;

---

[5] EPO Enlarged Board of Appeal G 0001/19 *(Pedestrian simulation)*
[6] *Pozzoli SPA v BDMO SA* [2007] EWCA Civ 588
[7] BL O/0673/26

iii) Perform the intermediate step: determine whether a feature contributes to the technical character of the invention as a whole and filter out features which do not contribute to, or interact with, the technical character of the invention; and

iv) Consider the novelty and inventive step of the filtered claim based on matter cited as forming part of the "state of the art".

21 The applicant agreed that I should follow the EPAI approach, but provided no further submissions on the detailed steps I should take.

The invention

22 The description of the application states that the disclosure relates to natural language processing (NLP). The application states that NLP is a branch of Artificial Intelligence (AI) concerned with giving computers the ability to understand text and spoken words in much the same way as human beings can (paragraph 3 of the application as published). The claimed invention defines a method for updating (called "retrofitting") the vectors representing words (called "word embeddings") to provide closer synonymous vectors. These closer synonymous vectors can be later provided for the training of other machine learning models, such as a Question-Answer model assisting a chatbot, a Query-Understanding model for Retail, Health, and Education sectors, a Semantic analysis model for processing tweets or feedback from users, and a Name Entity Recognition model (paragraph 93).

23 The application sets the context of the invention as relating to manipulating vector word embeddings, so that they are more semantically aware. It is stated that the vector word embeddings in GloVe (a well-known model for distributed word representation) are insufficiently semantically aware. As an example, it is stated that the word "rose" has multiple meanings – a colour, a flower, an object to express love, amongst others. English is said to have many words that are ambiguous – words that have multiple meanings. However, it is stated that a singular representation for each meaning of a word is detrimental to machine learning model performance. It is said to be important that the linguistic characteristics of the word should be preserved when fine-tuning a vector word embedding for any NLP task and that prior art methods of fine tuning are restricted to antonyms and synonyms.

24 The application as amended has a single independent claim, claim 1. Claim 1 is long and includes many steps. For ease of reference and understanding, I have split the claim into identifiable steps and emphasised some words. Claim 1 reads:

1. A system (110) for facilitating user meta data discovery and retrofitting framework, the system (110) comprising:
a) a processor (202);
b) a memory (204) coupled to the processor (202), wherein the memory (204) comprises processor-executable instructions, which on execution, causes the processor (202) to:
c) receive a set of word embeddings, wherein the word embeddings comprise vector representations of a word;

d) compute a similarity score and a dissimilarity score between every pair of word embeddings in the received set of word embeddings;
determine a margin value for every pair of word embeddings based on the similarity score and the dissimilarity score computed for the respective pair of word embeddings;
e) process the set of word embeddings through a K-nugget discovery pipeline comprising a plurality of stages including:
e1) a lexical stage configured to extract basic lexicons from the word embeddings, wherein the lexical stage utilizes a FastText model to capture morphological variants and morphological prefixes of words using cosine similarity scores and applies a Lexical Delta Boosting Method to generate a weighted score by adding a transformation embedding to a root word and projecting the transformation embedding to near vocabulary words, the weighted score being based on a weighted similarity score and a rate of similarity deviation from a root-word similarity score;
e2) a syntactic stage configured to extract syntactic relations including synonyms, antonyms, hypernyms, hyponyms, entity substitutes, and entity complements of words and to retrofit the word embeddings using global domain knowledge;
e3) a semantic stage configured to perform intent understanding using a stacked embedding approach by comparing similarity scores based on a hierarchy of concepts;
e4) a transactional stage configured to extract long-range context information and short-range context information from the word embeddings; and
e5) a language-agnostic stage configured to extract entities based on dialects across geographical locations;
f) discover user meta data lexically for a character-level understanding of words for every pair of word embeddings based on the respectively determined margin value by a dynamic retrofitted embedding approach,
wherein the user meta data is discovered by a combination of:
f1) a stacked embedding approach to solve perplexity of words with respect to surrounding context based on the similarity score and the dissimilarity score;
f2) a traditional knowledge approach by extracting a long-range context information of words and a short-range context information of words; and
f3) a language-agnostic knowledge approach by identifying dialects of pronunciation of a word;
g) compute, for each pair of word embeddings, a use-case relevance score using a classifier and an entropy indicative of mixed senses, and calculate a dynamic lambda hyperparameter by using a monotonic decreasing function of the relevance/entropy, the dynamic lambda regulating a regularization term in a retrofitting cost function comprising synonymy loss and antonymy loss;
h) retrofit the word embeddings using global domain knowledge with the retrofitting cost function regulated by the dynamic lambda;
i) perform hierarchical block iteration using cosine similarity to generate granular-level or category-level retrofitted embeddings that are knowledge-aware, domain-specific, and domain-agnostic;

j) apply a Self-Organizing Tree Algorithm to the retrofitted embeddings to form a hierarchical organization used to determine closer synonymous vectors based on the user meta data discovered for each pair of word embeddings; and
k) provide the closer synonymous vectors for training machine learning models such as a Question-Answer model assisting a chatbot, a Query-Understanding model for Retail, Health, and Education sectors, a Semantic analysis model for processing tweets or feedback from users, and a Name Entity Recognition model.

Argument and Analysis

Step (i): Claim construction

25 Claim 1 reads "A system for facilitating user meta data discovery and retrofitting vector word embeddings". …" which is widely accepted to be construed as "a system *suitable for* facilitating user meta data discovery and retrofitting vector word embeddings…". Unsurprisingly, I had no submissions to the contrary. The 'word embeddings' are the input to the claimed system. The word embeddings are vector representations of words that the processor modifies, in numerous ways, in order to create the closer synonymous vectors. The closer synonymous vectors are the output of the system and they are also construed as being *suitable for* use in training machine learning models such as those mentioned in claim 1.

26 At the hearing, the applicant's submissions largely concerned claim 1 and how it should be construed.

27 The 'knowledge nugget (K-nugget) discovery pipeline' (e) has five stages (e1-5) that extract different aspects of contextual information of the word embeddings. When queried, the applicant stated that 'global domain knowledge' within the syntactic stage was said to be understandable to the skilled reader and would be taken to mean an external reference repository. I accept that and will proceed on that basis.

28 The applicant explained that 'user meta data' (f) provides additional data for the meaning and context of the words in use, which is discovered using a variety of approaches: the stacked embedding approach to solving perplexity (f1) allows the context of ambiguous words to be discovered and accounted for, which together with the traditional knowledge approach (f2) and the language-agnostic approach (f3), give more data of the meaning and context of the words in use. I will proceed on that basis.

29 In relation to the computation of the use-case relevance score (g), it was submitted that 'entropy' relates to the ambiguity or the number of meanings or tenses that a word may have and regulates when the vector retrofitting is to occur. It was submitted that the entropy was used to prevent the overfitting or underfitting of the updated word embeddings and therefore acts as a control step to the modification of the word embeddings, in (h), which in turn feeds into the hierarchical block iteration (i), which generates the retrofitted word embeddings. I accept this and will proceed on that basis.

Step (ii): Any hardware test

30 It is common ground between the examiner and the applicant, that the invention requires hardware since claim 1 defines a system that requires a processor and memory. This test is satisfied.

Step (iii) Intermediate step

31 The next step is to consider which features contribute to the technical character of the invention as a whole. Non-technical features which do not contribute to, or interact with, the technical character of the invention are non-technical "as such" and are therefore filtered out before assessing novelty and inventive step

Argument

32 At the hearing, Drs Khan and Pearson pointed out that there were technical advantages in both the manner in which the closer synonymous vectors were developed and in the use of the machine learning models that are trained using the closer synonymous vectors. Summarising what the applicant considered to be the result of the claimed invention, Dr Khan said that the invention provided a better trained machine learning model input to a computer, which provides for more efficient and more meaningful search results in the field of artificial intelligence and language processing by a computer. The advantages were said to be (i) fewer computer resources used in the generation of the closer synonymous vectors, (ii) fewer computer resources required when the closer synonymous vectors were used downstream, and (iii) the use of the closer synonymous vectors provides more meaningful results.

33 At the hearing and in the correspondence on file, there is a consistent line of submissions from the applicant on specific technical features contributing to the technical character of the invention, which can be summarised into the three arguments below.

34 Firstly, it is submitted that the claimed invention controls the transformation and organisation of embedding vectors to prevent unnecessary embedding updates and act as a computational control mechanism. Specifically, the applicant argued that the present invention requires a specific computer implemented framework for discovering user meta data and dynamically retrofitting word embeddings. The invention was said to require "*a series of specific computational operations …(which) collectively define a concrete machine-implemented processing architecture that controls the transformation and organization of embedding vectors, rather than merely determining the meaning of words."* In particular, it was submitted that the use of the dynamic lambda equated to an adaptive control function which directly controls the extent of embedding updates applied during retrofitting operations.

35 Secondly, it was submitted that the hierarchical block iteration (i) and self-organising tree algorithm (j) provide more efficient computation of large collections of embedding vectors, and that technical advantages arise from the computational organisation of processing operations. The hierarchical block iteration, which partitions embedding space processing into multiple levels, reduces the number of comparison operations required during retrofitting. By organizing embedding

contributions hierarchically, the claimed framework improves scalability and computational efficiency when processing large collections of embedding vectors. Additionally, the Self-Organising Tree Algorithm generates a specific machine-implemented hierarchical data structure that efficiently organises and compares embedding vectors, to further increase the computational efficiency.

36 Thirdly, it was submitted that the specific model and method for the lexical stage (e1) reduce errors in similarity evaluation resulting in more reliable transformation and organisation of embedding vectors.

37 Drs Khan and Pearson argue that the entirety of the invention as claimed should be considered to be technical on the basis of any or all of the three technical advantages and three specific technical features in the above paragraphs, and therefore the whole claim should progress through the intermediate step. Although it is by no means determinative, I note that no EPO or UK case law has been provided by the applicant to support their submissions or specific arguments on the facts of the application.

38 The examiner has a different view. The examiner determines that the system of the claim involves the use of hardware and then applies the intermediate step by assessing whether each feature of the claim contributes to the technical character of the invention as a whole. The examiner filters out the features that they considered did not make such a contribution. After filtering, they conclude that the remaining technical features are only “a processor and associated memory capable of computing word vectors”. The examiner considers that the filtered features of the claim do not contribute to, or interact with, technical character as they are either algorithmic or linguistic.

39 In coming to this conclusion, the examiner referred to EPO Guidelines for patentability, in particular G-II 3.3 which states that *“ A mathematical method may contribute to an invention's technical character, i.e. contribute to producing a technical effect that serves a technical purpose, by being applied in a field of technology and/or by being adapted to a specific technical implementation (T 2330/13)*” and G-II 3.3.1 which states that *“Artificial intelligence and machine learning are based on computational models and algorithms such as artificial neural networks, genetic algorithms, support vector machines, k-means, kernel regression and discriminant analysis. Such computational models and algorithms are per se of an abstract mathematical nature, irrespective of whether they can be "trained" using training data.*” The examiner concluded that closer synonymous vectors, and their future use to train question-answering, query-understanding, semantic-analysis or named-entity-recognition models, did not establish a specific external technical purpose or a specific technical implementation.

40 The examiner stated that the system of the invention is implemented by a computer program, which runs on conventional computer hardware and that is not specifically adapted to a particular hardware implementation. The program’s design is not motivated by technical considerations relating to the internal functioning of the computer system or network, and so there is no “further technical effect” beyond the normal interactions between the program and the computer.

41 The examiner considers that any efficiency gains are due to the operation of the program or the organisation of the output data (the synonymous vectors), rather than due to an improvement in the operation of the computer as such. Further, the examiner considers that any reductions in updates, comparisons and search space are due to algorithmic optimisations that do not convey a further technical effect.

Analysis

42 It is clear that the claimed invention involves mathematical and computational methods. The EPO Guidelines referred to by the examiner are therefore highly relevant and provide useful guidance when applying the intermediate step, although I note that they are not binding upon me. Only the technical features of the claimed invention, and the non-technical features which interact with the technical subject matter of the invention as a whole, will remain unfiltered and proceed to the inventive step analysis.

43 With reference to the following first paragraph from the EPO Guidelines G-II, 3.3.1 (my emphasis added):

"Artificial intelligence and machine learning are based on computational models and algorithms such as artificial neural networks, genetic algorithms, support vector machines, k-means, kernel regression and discriminant analysis. Such computational models and algorithms are per se of an abstract mathematical nature, irrespective of whether they can be "trained" using training data. However, their use does not by itself render inventions related to artificial intelligence or machine learning non-patentable, and the guidance provided in G-II, 3.3 generally applies. This means that, if a claim of an invention related to artificial intelligence or machine learning is directed either to a method involving the use of technical means (e.g. a computer) or to a device, its subject-matter has technical character as a whole and is thus not excluded from patentability under Art. 52(2) or (3). *In such cases, the computational models and algorithms themselves contribute to the technical character of the invention if they contribute to a technical solution to a technical problem, for example by being applied in a field of technology and/or by being adapted to a specific technical implementation."*

44 However, there is no technical problem or specific technical purpose disclosed in the application as filed. The problem solved relates to efficiently incorporating additional linguistic characteristics of words into word embedding vectors. Therefore, consideration of the problem solved by the invention does not assist in identifying technical features which are not filtered out by the intermediate step.

45 Referring to EPO Guideline G-II, 3.6 "*In order to have technical character, and so not be excluded from patentability, a computer program must produce a "further technical effect" when run on a computer. A "further technical effect" is a technical effect going beyond the "normal" physical interactions between the program (software) and the computer (hardware) on which it is run. The normal physical effects of running a program, e.g. the circulation of electrical currents in the computer, are not in themselves sufficient to confer technical character on a computer program (T 1173/97 and G 3/08)"*

46 As stated by the examiner, there is nothing in the application that relates to a specific technical implementation operable to exploit particular technical properties of the system on which it is implemented to provide a further technical effect. Where the hardware of the computer system used to execute the claimed functionality is discussed in the description, at paragraphs 51-53 & 86-88 for example, there is no suggestion that the program has been specifically designed to work with the particular technical properties of the hardware. Also, the computational models and algorithms are not described as being adapted to suit any specific technical implementation. Therefore, consideration of specific implementation does not assist in identifying technical features which are not filtered out by the intermediate step.

47 The applicant submits that the use of the dynamic lambda equates to a dynamic control function, which controls the extent of the updates to the word embeddings. Control of a specific technical process can be a technical purpose, and EPO Guidelines G-II, 3.3 include examples related to controlling a specific technical process, e.g. controlling an X-ray apparatus or a steel cooling process. However, in this application, the process being controlled is the extent of word vector retrofitting by algorithms, which I do not consider to be technical. Therefore, consideration of process control does not assist in identifying technical features which are not filtered out by the intermediate step.

48 The applicant submits that the invention provides technical advantage through greater computational efficiency, but the EPO Guidelines G-II, 3.3 (Computational Efficiency) are also clear in this regard: "*If the mathematical method does not serve a technical purpose and the claimed technical implementation does not go beyond a generic technical implementation, the mathematical method does not contribute to the invention's technical character. In such cases, it is not sufficient that the mathematical method is algorithmically more efficient than prior-art mathematical methods to establish a technical effect"*. I find that the more efficient processes claimed do not serve a technical purpose because the specified algorithms have not been applied in field of technology or adapted to a specific technical implementation. Therefore, algorithmic efficiency does not assist in identifying technical features which are not filtered out by the intermediate step.

49 I conclude that the computational and algorithmic steps claimed do not contribute to the technical character of the invention as a whole and therefore should be filtered out during the intermediate step. Consequently, steps (d), (g), (h), (i) and (j) will be removed and not proceed to the inventive step analysis.

50 The claim includes five steps (e1-5) relating to the lexical, syntactic, semantic, transactional and language-agnostic analysis of the word embeddings, and also three more steps (f1-3) analysing the context, use and pronunciation of words. I agree with the examiner that linguistic analysis does not establish technical character, and derive some support for this conclusion from the EPO Guidelines G-II, 3.3.1, which state that "…classifying text documents solely according to their textual content is not considered to be per se a technical purpose but a linguistic one (**T 1358/09**)."

51 Although this guideline relates specifically to classifying text documents, rather than relating to any or all inventions involving some linguistic analysis, it seems

reasonable to conclude that the understanding of textual content is considered not to be a technical purpose.

52 The applicant specifically submits that the specific model and method used for the lexical stage (e1) reduce errors in similarity evaluation. While this may well be the case, the improvement in similarity evaluation of word embeddings is also considered to be a linguistic purpose and not a technical purpose.

53 I find that the claimed steps for linguistically analysing and extracting further contextual information about the word embeddings, or pairs of word embeddings, do not contribute to the technical character of the invention, and so further claim features (e1-5) and (f1-3) are also filtered out during the intermediate step.

54 Further, the modification of the inputted vector representations of a word to create "closer synonymous" vectors, is also a linguistic difference which should be filtered. After filtering, the technical part of the output is considered to be merely "modified vectors" without specifying their closer synonymous aim.

55 In applying the intermediate step to filter out the computational and algorithmic steps (d), (g), (h), (i) and (j) and the linguistic analysis steps (e1-5) and (f1-3), I am left with a filtered claim 1, which reads:

A system (110) suitable for facilitating user meta data discovery and retrofitting framework, the system (110) comprising:
a processor (202);
a memory (204) coupled to the processor (202), wherein the memory (204) comprises processor-executable instructions, which on execution, causes the processor (202) to:
receive a set of word embeddings, where the word embeddings comprise vector representations of a word;
process the set of word embeddings; and
provide modified vectors suitable for training machine learning models.

Step (iv): Inventive step

56 The examiner applied the intermediate step to claim 1 and found that the remaining technical features were only "a processor and associated memory capable of computing word vectors". After filtering the claim down to this, the examiner maintained their objection to a lack of inventive step in view of *Sapan Shah et al*[8], an academic publication which had been cited against the European equivalent of the present application.

[8] *A retrofitting model for incorporating semantic relations into word embeddings" by Sapan Shah et al., Proceedings of the 28th International Conference on Computational Linguistics, pages 1292-1298, Barcelona, Spain (Online), December 8-13, 2020.*

57 The applicant made no submissions on the inventiveness of the claim as filtered by the examiner; their position was that all features of the claim should proceed intact through the intermediate step to be considered in the inventive step analysis.

58 It is common ground between the examiner and the applicant that the skilled person is a designer working in the field of, and having common general knowledge of, natural language processing.

59 The prior art cited by the examiner, *Sapan Shah et al*, discloses the computer-implemented modification of vector word embeddings, in which all types of word relations present in a knowledge resource (i.e. synonymy, antonymy, hypernymy, holonymy and so on) are used to obtain constraints in the form of mathematical relation inequalities. These inequalities are then used to map original word vectors to a vector space which respects these constraints[9].

60 This prior art process would necessarily use a processor, memory and processor-executable instructions. Additionally, the prior art process also has vector word embeddings and information about word context as an input, processes the input data, and outputs modified word vectors which more accurately reflect the full context of the word.

61 I consider that all of the features of filtered claim 1 above are disclosed in the prior art. As I find no differences between the filtered claim and the "state of the art" as required by step 3 of *Pozzoli*, there is no need to determine whether any differences constitute steps which could have been obvious to the person skilled in the art, as required in step 4. I therefore find that the invention as defined in claim 1 lacks an inventive step.

62 I received no specific submissions on the dependent claims. After careful consideration, I am of the opinion that the dependent claims specify further details of linguistic analysis which would also be filtered when applying the intermediate step. Consequently, I find that all of the dependent claims also lack an inventive step.

Decision

63 I have decided that the invention defined in the independent claim lacks an inventive step. I have considered the contents of the application and do not believe that any saving amendment is possible. I therefore refuse the application under section 18(3).

Appeal

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

Emma Porter

Patent Group Head, acting on behalf of the comptroller

---

[9] Section 1, Introduction, especially last paragraph of page 1292 to end of second paragraph of page 1293.