Internal distribution code:

(A) [-] Publication in OJ
(B) [-] To Chairmen and Members
(C) [-] To Chairmen
(D) [X] No distribution

Datasheet for the decision of 23 September 2024

Case Number: T 2401/22 - 3.5.07

Application Number: 18204942.9

Publication Number: 3547161

IPC: G06F17/27, G06Q50/00, G06N3/02

Language of the proceedings: EN

Title of invention:
Position-dependent word salience estimation

Applicant:
SAP SE

Headword:
Word salience estimation/SAP

Relevant legal provisions:
EPC Art. 56
RPBA 2020 Art. 13(2)

Keyword:
Inventive step - main request (no) - auxiliary request (no)

Decisions cited:
G 0001/19, T 1351/04, T 2330/13, T 0598/14, T 0755/18, T 0702/20, T 1903/20, T 1952/21

Europäisches Patentamt
European Patent Office
Office européen des brevets

Beschwerdekammern

Boards of Appeal

Chambres de recours

Boards of Appeal of the
European Patent Office
Richard-Reitzner-Allee 8
85540 Haar
GERMANY
Tel. +49 (0)89 2399-0

Case Number: T 2401/22 - 3.5.07

D E C I S I O N
of Technical Board of Appeal 3.5.07
of 23 September 2024

Appellant: (Applicant)	SAP SE Dietmar-Hopp-Allee 16 69190 Walldorf (DE)
Representative:	DeltaPatents B.V. Fellenoord 370 5611 ZL Eindhoven (NL)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 27 June 2022 refusing European patent application No. 18204942.9 pursuant to Article 97(2) EPC**

Composition of the Board:

Chair	J. Geschwind
Members:	P. San-Bento Furtado
	M. Jaedicke

Summary of Facts and Submissions

I. The appeal lies from the decision of the examining division to refuse European patent application No. 18204942.9. The decision was given in writing after cancellation of the oral proceedings.

II. In the decision under appeal, the examining division decided that the subject-matter of claim 1 of the main request and first auxiliary request lacked inventive step over a notoriously known general purpose computer used for indexing and retrieving documents such as that disclosed in prior-art document D3:

D3: US 2013/0041921 A1, published on 14 February 2013.

The examining division further decided that the claimed subject-matter related substantially to non-technical linguistic and mathematical models for generating word salience values of words in text data.

III. With the statement of grounds of appeal, the appellant resubmitted the main request and first auxiliary request considered in the decision under appeal and filed a new second auxiliary request. It requested that the decision under appeal be set aside and that a patent be granted on the basis of the main request or the first or second auxiliary request.

IV. In a communication accompanying a summons to oral proceedings, the board cited also the following documents:

D1: Chawla et al.: "Investigating how well contextual features are captured by bi-directional recurrent neural network models", Proceedings of the 14th

International Conference on Natural Language Processing, pages 273 to 282, Kolkata, India, December 2017;

D2: G. Zhou et al.: "Cross-lingual sentiment classification with stacked autoencoders", Knowledge Information Systems 2016, 47, pages 27 to 44;

D4: D. Bahdanau et al.: "Neural Machine Translation by Jointly Learning to Align and Translate", Cornell University, https://arxiv.org/abs/1409.0473, 19 May 2016.

Documents D1 and D2 were listed in the search report. Document D4, which was cited in document D1, was introduced into the proceedings by the board.

The board expressed its preliminary opinion that claim 1 of the main request added subject-matter beyond the content of the application as filed. The subject-matter of claim 1 of the main request and first auxiliary request was not inventive over a prior-art method for retrieval of electronic documents being indexed and associated with word salience values determined for words in the text data, which the application acknowledged as prior art. The board was not inclined to admit the second auxiliary request into the proceedings.

V. With a letter of reply dated 19 August 2024 the appellant filed new third and fourth auxiliary requests.

VI. Oral proceedings were held as scheduled. The board informed the appellant that it admitted the third and fourth auxiliary requests into the proceedings. The appellant then withdrew the mean request and first and second auxiliary requests and maintained the third and

fourth auxiliary requests as its main request and first auxiliary request. At the end of the oral proceedings, the Chair announced the board's decision.

VII. The appellant's final request was that the contested decision be set aside and that a patent be granted on the basis of the main request or the first auxiliary request corresponding to the third and fourth auxiliary requests filed with the letter of 19 August 2024.

VIII. Claim 1 of the main request reads as follows:

"A computer-implemented method (300) for retrieval of electronic documents being indexed and associated with word salience determined for words in text data, the method being executed by one or more processors (410) and comprising:

receiving (302, 304), by the one or more processors, two or more electronic documents, each electronic document comprising text data, a second electronic document comprising a link to a first electronic document;

processing (306), by the one or more processors, word representations of words of the first electronic document using a first encoder (202) to provide first output and a context vector (210);

processing (308), by the one or more processors, text data of the second electronic document and the context vector using a first decoder (206) to provide second output;

determining (310), by an attention mechanism (204) executed by the one or more processors, a plurality of weights for each word in the text data of the first electronic document based on the first output, and the second output;

providing (312), by the one or more processors, a word salience value for each word, a word salience

value comprising a sum of weights of a respective word; and

using the word salience values in a system (100; 400) to perform retrieval of electronic documents being indexed and associated with word salience values of words included in electronic documents."

IX. Claim 1 of the first auxiliary request differs from claim 1 of the main request in that the text "to provide first output and a context vector (210)" has been replaced with

"to provide first output and an encoder hidden state (h_i^{enc}) for each word of the first electronic document"

and the text "text data of the second electronic document and the context vector using a first decoder (206) to provide second output" has been replaced with

"text data of the second electronic document using a first decoder (206) to provide second output, each decoder step using a different encoder context vector (c_j) based on each encoder hidden state (h_i^{enc}) and a previous decoder state (h_{j-1}^{dec})".

Reasons for the Decision

Invention

1. The invention concerns the determination of the "word salience" of words in electronic documents for the retrieval or summarisation of electronic documents (see original description, page 1, lines 4 to 6; original claim 8).

1.1 Word salience is "the relative importance of a word within an electronic document" (page 2, lines 11 to 19). It can be determined based on term frequency-inverse document frequency (TFIDF), which increases proportionally to the number of times a word appears in the document, and is offset by the frequency of the word in the document. According to the description, the traditional TFIDF fails to account for the relative importance of words in different sentences of the document (page 2, lines 11 to 26).

1.2 The invention proposes position-dependent word "salience score" determination which takes into account "secondary data" having an association with the text data in electronic documents (e.g. articles). For example, the secondary data can be social media data or "linking tweets" with uniform resource locator (URL) links to articles. These are expected to reflect reader interest in the articles (page 6, line 31 to page 7, line 29).

1.3 A method according to the invention, as described on page 5, lines 19 to 28, receives electronic documents, each including text data. The electronic documents comprise a first electronic document (e.g. an article) and a second electronic document including a link to the first electronic document (e.g. a linking tweet with URLs to the article). The method uses a "first encoder" (e.g. a recurrent neural network, RNN), a "first decoder" (e.g. a bidirectional grand recurrent unit, GRU) and an attention mechanism. The first electronic document is processed using the "first encoder" to provide "first output" and a context vector. The second electronic document and the context vector are processed using the first decoder to provide "second output". The hidden representation at the last

step of the encoder is considered as the context vector (page 9, lines 16 to 31). The attention mechanism determines a plurality of weights for each word in the text data of the first electronic document based on the first output and the second output. These weights are added to provide a word salience value for each word (page 5, lines 15 to 28).

1.4 The word salience values can be used in an information retrieval system in which electronic documents (e.g. articles) can be indexed and associated with word salience values included in the electronic documents. A word can be indexed to multiple electronic documents. In response to a query including a word, the information retrieval system can return ranked results representative of the electronic documents based on the word salience values of the word (page 14, lines 5 to 20).

Admission of claim requests into the appeal proceedings

2. The main request and first auxiliary request differ from the requests considered in the decision under appeal in that the feature specifying the summarisation of electronic documents has been deleted. This amendment addresses the objection of added subject-matter raised for the first time in the board's communication under Article 15(1) RPBA. These circumstances constitute an exceptional circumstance within the meaning of Article 13(2) RPBA. Hence, the board admits the main and first auxiliary requests into the appeal proceedings.

Main request

3. *Inventive step*

3.1 Claim 1 specifies a computer-implemented method for "retrieval of electronic documents being indexed and associated with word salience determined for words in text data". The method of claim 1 uses an encoder-decoder architecture with an attention mechanism (known in the art, see e.g. document D4, abstract and page 4, third paragraph) to calculate word salience values of words of a first document as explained under point 1.3 above. As a last step of the claimed method, the word salience values are used for retrieving "electronic documents being indexed and associated with word salience values".

3.2 As explained in the background section of the description, on page 1, line 9 to page 2, line 26, at the date of priority of the present application, information retrieval systems were known which used one or more processes to calculate word salience of words and use them in order to identify, retrieve and/or summarise electronic documents based on the words present in the documents and their salience values.

3.3 Such prior-art systems were known to index the electronic documents based on word salience values in order to retrieve electronic documents. In particular, the board notes that neither the description nor claim 1 describe how the word salience values are used in a system to perform retrieval, or how the documents are indexed in the system. The application thus assumes that the retrieval and indexing are performed using standard techniques.

Such a prior-art method for retrieval of electronic documents being indexed and associated with word salience values determined for words in the text data (or simply "prior-art method for retrieval of indexed documents") is an appropriate starting point for assessing inventive step.

3.4 The subject-matter of claim 1 differs from the prior-art method for retrieval of indexed documents in that

(a) a second electronic document comprising a link to a first electronic document is also used in calculating the salience values;

(b) the word salience value for each word comprises a sum of weights of a respective word;

and in that the following steps are performed:

(c) processing word representations of words of the first electronic document using a first encoder to provide first output and a context vector;

(d) processing text data of the second electronic document and the context vector using a first decoder to provide second output;

(e) determining, by an attention mechanism, a plurality of weights for each word in the text data of the first electronic document based on the first output, and the second output.

3.5 The appellant argued that the invention of claim 1 achieved high quality word saliency, and thus led to better indexing than in the prior art. Word saliency had a direct impact on how the computer performed a search. It was used to steer the searching towards or away from certain documents. Searching in large databases was an inherently technical task and data that determined the search outcome was technical data in view of decision T 1351/04. It was inconsistent to consider that searching and its related infrastructure,

such as the indexing which made the searching possible, was technical, yet which documents were produced by the search was irrelevant.

According to the appellant, this dichotomy — searching is technical, but search results are irrelevant — was untenable. Searching existed for a reason. Modern large scale storage systems stored hundreds of thousands of documents, if not millions. Such a storage was technically useless without search. A user who needed a document would have to manually go through those documents one by one until they had found the relevant one. Without an appropriate search function, a storage system beyond a couple of hundred documents would become dysfunctional. They ceased to provide the function that they were created for, namely providing the needed document.

Furthermore, the assumptions that the invention translated linguistic considerations into a mathematical model were incorrect. This did the invention no justice, as in fact, it was rather the opposite. The model did not learn which documents were responsive to which queries from a linguist. The neural models were fully general, and learned from the data itself. Nor did the model learn from a mathematician, as might be the case if a fixed formula were used. Instead, a novel machine learning architecture was defined that allowed the machine to learn the desired mapping itself. The machinery to learn this mapping was not known, or at least not disclosed in the cited documents.

The appellant further argued that the invention provided a technical mechanism to narrow down search results. The invention used word salience values, which were determined not by arbitrary user preferences but

through the training of a neural network. This training process ensured that the word salience values reflected the relative importance of terms within the context of the documents themselves, rather than subjective or user-specific criteria. As such, the word salience values served as a robust, technical solution to the problem of information overload in large-scale document retrieval systems. Furthermore, the invention’s method for calculating and applying these salience values was inherently technical as it involved data processing techniques, including the use of encoder decoder architectures and attention mechanisms, to derive salience values that enhanced the accuracy and relevance of search results.

At the oral proceedings the appellant also argued that the idea of using other documents that refer to the document through links, as in the well-known PageRank approach, was technical. The problem solved by the invention was: how to translate the interest expressed in secondary documents into numerical data that could be used in a normal information retrieval process. The solution used technical data derived from training which was different from linguistic concepts. It was not known how to take secondary documents into account and it was not obvious to devise a way to extract the relevant information from a large set of documents referencing another set of documents.

3.6 The board does not find the appellant's arguments convincing. The case law clearly establishes that even though information retrieval in a computer system may include technical tasks, not everything related to searching is technical. Aspects related to user preferences, linguistics and semantics are in principle not technical. Obtaining search results which better

meet the user's interests or more closely match the semantics of the search terms is not a technical effect (T 598/14, Reasons 2.4). As explained in decision T 598/14, Reasons 2.5, while functional data, such as an index structure, which "controls the computer by directing it to a certain memory location" is considered technical in accordance with decision T 1351/04, not all data used in an information retrieval system is considered functional data contributing to a technical effect. Index terms which merely correspond to keywords are as such not technical.

3.7 In the present case, the indexing mechanism itself is not new. The way the electronic documents are indexed and associated with word salience values and the way the word salience values are used to retrieve electronic documents are known from the prior-art method for retrieval of indexed documents (see also points 3.2 to 3.4 above). The link from the second to the first document of feature (a) is well known in the art. In the context of the invention, it is not used for the purpose of accessing data in the computer, but to derive the user interest in a document. The distinguishing features are thus not functional data, i.e. data making a technical contribution, within the meaning of decision T 1351/04.

3.8 The result of the distinguishing features is that different salience values are obtained. Word salience values reflect both the relevant information in the electronic documents and the readers' interest in them (see e.g. page 2, lines 11 to 19; page 7, lines 21 to 29; page 10, lines 21 to 26). A salience value is thus a non-technical mathematical parameter representing users' interests in the content of a document.

In the claimed method, a second electronic document, e.g. a linking tweet, is used to obtain information about the user interest in the electronic documents (see e.g. page 7, lines 24 to 27). This is the result of a non-technical consideration. Furthermore, the use of e.g. web pages referencing a web page for evaluating the user interest in the web page is well known.

Taking the above into account, feature (a) merely uses notoriously known technical means for a non-technical purpose and feature (b) does not make a technical contribution.

3.9 Features (c) to (e) specify features of a computer program for calculating the salience values of words. Computer programs are as such not patentable (Article 52(2)(c) and (4) EPC). The board does not recognise any technical considerations in the manner a first encoder, a first decoder and an attention mechanism are used to calculate the non-technical salience values of words (see also T 702/20, Reasons 7 to 21; T 598/14, Reasons 2.3 and 2.4; T 1903/20, Reasons 3.3). Features (c) to (e) merely produce non-technical salience values, no technical effect being apparent from the claim.

3.10 The board is not persuaded by the appellant's argument that the claimed method was patentable because it was based on a machine learning architecture that allowed the machine to learn the desired mapping itself.

In the board's opinion, the claim does not specify the features necessary to achieve such a learning effect. In any case, even if the board recognised that the claimed method caused "the machine to learn the desired mapping itself", the board would not acknowledge the learning effect to cause a technical effect. In

decision T 755/18 the board concluded that if neither the output of a machine nor the output's accuracy are technical, an improvement of the machine achieved automatically through supervised learning to generate a more accurate output is not in itself a technical effect (Reasons 3.2). In the board's opinion, the same holds true for unsupervised machine learning of a machine that produces a non-technical output. In decision T 1952/21, the examined claim specified a machine learning system based on neural networks. The board considered that the claim did not restrict the invention to a technical context. The board did not recognise reinforcement learning as a further technical use, even if the advantages in reinforcement learning brought forward by the appellant were to be acknowledged (Reasons 18 to 27.4 and 33.2 to 34).

In line with the cited case law, in the present case, the use of the results of the method in the machine itself in order to learn how to achieve "enhanced ... accuracy and relevance of search results", as argued by the appellant, is not a further technical use within the meaning of G 1/19. Even if features (a) to (e) of the claimed method were considered to contribute to a machine capable of learning itself how to produce better salience values, this would constitute a non-technical algorithmic change of a computer program to produce salience values which better represented the user interests, which is not a technical effect.

3.11 Therefore, the subject-mater of claim 1 lacks inventive step (Article 56 EPC).

First auxiliary request

4. Claim 1 of the first auxiliary request differs from claim 1 of the main request in that features (c) and (d) have been replaced with the following features:

(f) processing word representations of words of the first electronic document using a first encoder to provide first output and an encoder hidden state (h_i^{enc}) for each word of the first electronic document;

(g) processing text data of the second electronic document using a first decoder to provide second output, each decoder step using a different encoder context vector (c_j) based on each encoder hidden state (h_i^{enc}) and a previous decoder state (h_{j-1}^{dec}).

5. *Inventive step - claim 1*

5.1 Since the features added by the amendments are not disclosed in the prior art, the distinguishing features of claim 1 of the first auxiliary request over the prior art are features (a), (b), (e), (f) and (g).

5.2 The appellant argued that the context vector could be quite long to encode the different relevancies in the first document, so that it could contribute to predicting the decoder targets in the decoder phase of the system, as was mentioned in paragraph [0033] of the application as published. A longer context vector had an impact throughout the system, requiring more resources to process.

The inventors had been confronted with the technical challenge of using long context vectors to make the encoder-decoder paradigm work best, in spite of the significant technical burden that this would entail. The solution comprised breaking up the long context

vector into multiple short context vectors. Rather than using the same context vector for each prediction target of the decoder, the system was allowed to build a series of context vectors for each decoding target. The series of context vectors could be based on all existing hidden encoder states, rather than only on the last one. This case was similar to that in decision T 2330/13, in which the board had recognised a technical contribution of the implementation of a non-technical method.

5.3 In reply to the board's argument that the result of the claimed method with features (f) and (g) was different from a hypothetical method according to claim 1 of the first auxiliary request but without these features, the appellant argued that even if the approach of claim 1 did not yield exactly the same result as a single, comprehensive context vector, this was irrelevant because the goal of the invention was not to replicate existing methods but to provide a novel and technically advantageous approach to document retrieval. The novel application of multiple shorter context vectors allowed useful word saliency computation at lower cost. Whether or not the exact same results were obtained was secondary to the fact that the claimed method introduced a new way of processing information, which could offer as good or better outcomes, even if potentially slightly different. Using multiple shorter context vectors provided practical advantages, such as reduced computational complexity, better scalability, and improved performance in processing large datasets.

5.4 In the board's opinion, the fact that the result is not the same reflects a non-technical trade-off regarding the desired search results. Furthermore, the claim, which describes the invention only in abstract terms,

does not specify all the features necessary for achieving the alleged technical effects. The board is thus not convinced that features (f) and (g) solve the technical problem of reducing the use of resources (see also T 702/20, Reasons 14.1) or to be adapted to a more efficient implementation, as in the case of decision T 2330/13. Consequently, distinguishing features (f) and (g) are non-technical and do not enter into the assessment of inventive step.

6. Therefore, the subject-matter of claim 1 of the first auxiliary request lacks inventive step (Article 56 EPC).

Conclusion

7. Since neither of the requests on file is allowable, the appeal is to be dismissed.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chair:

B. Brückner J. Geschwind

Decision electronically authenticated