**BESCHWERDEKAMMERN
DES EUROPÄISCHEN
PATENTAMTS**

**BOARDS OF APPEAL OF
THE EUROPEAN PATENT
OFFICE**

**CHAMBRES DE RECOURS
DE L'OFFICE EUROPÉEN
DES BREVETS**

Internal distribution code:

(A) [-] Publication in OJ

(B) [-] To Chairmen and Members

(C) [-] To Chairmen

(D) [X] No distribution

Datasheet for the decision
of 10 January 2019

Case Number:	T 0817/16 - 3.5.07
Application Number:	04784004.6
Publication Number:	1668551
IPC:	G06F17/30
Language of the proceedings:	EN

Title of invention:
Information retrieval based on historical data

Applicant:
Google LLC

Headword:
Document scoring/GOOGLE

Relevant legal provisions:
EPC Art. 56

Keyword:
Inventive step - all requests (no)

EPA Form 3030

This datasheet is not part of the Decision.
It can be changed at any time and without notice.

Decisions cited:

G 0003/08, T 0121/85, T 0107/87, T 0258/97, T 1177/97,
T 0258/03, T 1543/06, T 1784/06, T 1741/08, T 1214/09,
T 1358/09, T 0042/10, T 0306/10, T 2230/10, T 1321/11,
T 1370/11, T 1463/11, T 2035/11, T 2418/12, T 0136/13,
T 0650/13, T 2330/13

Case Number: **T 0817/16 - 3.5.07**

D E C I S I O N
of Technical Board of Appeal 3.5.07
of 10 January 2019

Appellant:	Google LLC
(Applicant)	1600 Amphitheatre Parkway
	Mountain View, CA 94043 (US)
Representative:	Hewett, Jonathan Michael Richard
	Venner Shipley LLP
	200 Aldersgate
	London EC1A 4HD (GB)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 11 November 2015 refusing European patent application No. 04784004.6 pursuant to Article 97(2) EPC**

Composition of the Board:

Chairman	R. Moufang
Members:	R. de Man
	P. San-Bento Furtado

Summary of Facts and Submissions

I.   The applicant (appellant) appealed against the decision
        of the Examining Division refusing European patent
        application No. 04784004.6, which was published as
        international publication WO 2005/033978.

II.   The Examining Division decided that the subject-matter
        of the independent claims of the then main request and
        auxiliary request lacked inventive step within the
        meaning of Articles 52(1) and 56 EPC over a notorious
        computerised information-retrieval system.

III.   With its statement of grounds of appeal, the appellant
        filed a main request and first and second auxiliary
        requests.

IV.   During the appeal proceedings the applicant/appellant
        changed its name from Google Inc. to Google LCC.

V.   In a communication accompanying the summons to oral
        proceedings, the Board expressed the preliminary view
        that none of the requests complied with Article 123(2)
        EPC and that the subject-matter of claim 1 of both the
        main request and the first auxiliary request lacked
        inventive step over a general-purpose computer. It also
        questioned whether the second auxiliary request should
        be admitted into the appeal proceedings under
        Article 12(4) RPBA.

VI.   The appellant replaced its previous requests with a new
        main request and first and second auxiliary requests in
        a letter dated 29 October 2018 (filed first via EPO
        Online Filing and then by fax, the latter submission
        including a corrected first auxiliary request).

VII.   Oral proceedings were held on 10 January 2019 and were
        attended by the appellant. At the end of the oral
        proceedings, the chairman pronounced the Board's
        decision.

VIII.   The appellant requested that the decision under appeal
        be set aside and that a patent be granted on the basis
        of the claims of the main request or, in the
        alternative, the first or second auxiliary request, all
        requests filed with the letter of 29 October 2018.

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

        "A method for scoring a document, comprising:
            identifying a document;
            obtaining one or more types of history data
        associated with the document, the one or more types of
        history data including data relating to changes to a
        content of the document over time,
            wherein obtaining the data relating to changes to
        the content of the document over time includes:
            monitoring signatures of the document to determine
        (i) a frequency at which the content of the document
        changes over time, and (ii) an amount by which the
        content of the document changes over time; and
            generating a score for the document based, at
        least in part, on the one or more types of history data
        associated with the document,
            wherein the generating the score for the document
        includes scoring the document based, at least in part,
        on the frequency at which the content of the document
        changes over time and the amount by which the content
        of the document changes over time."

X.   Claim 1 of the (corrected) first auxiliary request
        reads as follows:

"A computer-implemented method for scoring a document, comprising:

identifying a plurality of documents containing a plurality of terms;

storing, for each document of the plurality of documents, portions of the documents that are determined to be most frequently occurring instead of storing the entire document;

obtaining one or more types of history data associated with each of the plurality of documents, the obtaining performed by monitoring one or more types of history data including data relating to changes to a content of a respective document over time,

wherein the data relating to changes to the content of a respective document over time includes:

a frequency at which the content of the respective stored document portion changes over time, and

an amount by which the content of the respective stored document portion changes over time; and

generating a score for a document based, at least in part, on the one or more types of history data obtained for the document."

XI.   Claim 1 of the second auxiliary request reads as follows:

"A computer-implemented method for scoring a document, comprising:

identifying a plurality of documents containing a plurality of terms;

storing, for each document of the plurality of documents, a signature of the document instead of storing the entire document;

obtaining one or more types of history data associated with each of the plurality of documents, the

obtaining performed by monitoring one or more types of
history data including data relating to changes to a
content of a respective document over time,

wherein monitoring the data relating to changes to
a content of a respective document over time includes:

monitoring signatures of the respective documents
to determine (i) a frequency at which the content of
the respective document changes over time, and (ii) an
amount by which the content of the respective document
changes over time; and

generating a score for a document based, at least
in part, on the one or more types of history data
obtained for the document,

wherein the generating the score for the document
includes scoring the document based, at least in part,
on the frequency at which the content of the document
changes over time and the amount by which the content
of the document changes over time."

XII.   The appellant's arguments, where relevant to the
        decision, are discussed in detail below.

Reasons for the Decision

1.   The appeal complies with the provisions referred to in
        Rule 101 EPC and is therefore admissible.

2.   *The invention*

2.1   The application relates to search engines. Its
        background section explains that, ideally, a search
        engine provides the user with the results most relevant
        to the user's query. Relevant documents are typically
        identified on the basis of a comparison of the search-
        query terms to the words contained in the documents and

other factors such as the existence of links to or from
the documents. The detailed description discloses a
number of techniques for scoring documents, which may
be used to improve the search results returned in
response to a search query.

2.2   The claimed invention is directed to the embodiment
described on page 6, second full paragraph, to page 7,
third full paragraph, of the published application. It
proposes scoring a document on the basis of "history
data" that reflects the frequency at and the amount by
which the content of the document changes over time.
This history data is obtained by "monitoring signatures
of the document".

3.   *Main request – inventive step*

3.1   Unlike claim 1 of the first and second auxiliary
requests, claim 1 of the main request is not worded as
a "computer-implemented" method and therefore arguably
encompasses mental acts as such, which are excluded
from patentability under Article 52(2) and (3) EPC. But
since the appellant at the oral proceedings expressed
its willingness to limit the claim to a "computer-
implemented" method, the Board will, for the purpose of
assessing inventive step, interpret claim 1
accordingly.

3.2   Since the method of claim 1 can be performed on a
general-purpose computer, the Board considers such a
computer to be a suitable starting point for assessing
inventive step. The subject-matter of claim 1 differs
from this prior art in the steps listed in claim 1.

These steps define the algorithm underlying the
computer-implemented method in abstract, functional

terms that do not imply any interaction with specific technical means. In particular, the step "monitoring signatures of the document" calculates and compares signatures for different versions of the document without specifying a technical mechanism by which different versions are detected or retrieved. And "generating a score for the document" and "scoring the document" merely associate the document with a calculated score value. The steps of claim 1 are thus non-technical, apart from their implementation on a computer.

It therefore has to be analysed whether, and to what extent, the steps interact with the technical feature of the claim, i.e. the feature (which the Board reads into the claim for the purpose of assessing inventive step) specifying that the method is "computer-implemented", to produce a technical effect over a general-purpose computer.

3.3   The Board concurs with the Examining Division that assigning a score to a document based on the frequency and the amount of changes to the document is not a technical task, even if performed by a computer. The appellant originally did not dispute this, but at the oral proceedings it suggested that providing good scores improved the search results returned by the search engine and that improved search results resulted in a reduction in the number of search queries, which amounted to a saving of resources.

A similar argument was dealt with in decision T 306/10 of 4 February 2015 in the context of recommendation engines. The board there considered that a reduction in the number of search queries and the corresponding saving of resources did not qualify as a technical

effect of the (improved) recommendations, as they depended on subjective choices made by the user (see reasons 5.2). It referred to decision T 1741/08 of 2 August 2012, reasons 2.1.6, where the argument was made that a chain of effects cannot be used as evidence of a technical effect if one of the links between the effects is not of a technical nature (but, for example, of a psychological nature).

In the present case, the appellant's argument fails for the reason alone that claim 1 is silent on what the generated score is used for. Merely assigning a score to a document is not a technical effect. This is not different if the score is somehow based on the frequency and the amount of changes made to the document.

3.4   The appellant also argued that the method of claim 1 achieved a technical effect by implementing the task of assigning a score to a document based on the frequency and the amount of changes to the document in a particularly resource-efficient manner. Instead of storing the current version of a document in its entirety to allow the amount of changes in the next version of the document to be determined, the method of claim 1 only stored a "signature" and determined the amount of changes by comparing the signatures of the previous and new versions.

3.5   Document signatures are well known in the art but are usually suitable only for determining whether two documents differ, not for measuring the degree in which they differ. In this respect, the application, on page 7, lines 1 to 3, states the following:

"For example, search engine 125 may store 'signatures'
of documents instead of the (entire) documents
themselves to detect changes to document content. In
this case, search engine 125 may store a term vector
for a document (or page) and monitor it for relatively
large changes."

The Board notes that term vectors are well known in the
art. They essentially represent the content of text
documents as vectors of word frequencies. Measuring the
"semantic similarity" between two text documents by
computing the normalised inner product of their term
vectors is a standard technique. Term vectors are thus
indeed suitable for determining the amount of changes
between two documents or two versions of a document.

For the purpose of assessing inventive step, the Board
will therefore - to the appellant's benefit - interpret
"signature" narrowly as "term vector".

3.6   At least for larger documents, it is plausible that the
       term vector of a document takes up less memory space
       than the full document. But the claimed method does not
       achieve any savings of memory space over a general-
       purpose computer - which is the prior art that the
       Board has taken as the starting point for assessing
       inventive step. Indeed, performing the method of
       claim 1 on a general-purpose computer necessarily uses
       more memory resources than not performing the method.
       What performing the method does achieve is a particular
       scoring of documents, but that is not a technical
       effect. It also causes - like any program execution -
       some usage of memory and processor resources, which is
       at least a physical effect, but which is not a
       technical effect for the purpose of inventive step in
       so far as it does not go beyond the inherent effects of

running a program on a computer (see decisions
T 258/03, OJ EPO 2004, 575, reasons 5.4; T 1543/06 of
29 June 2007, reasons 2.7 and 2.8; and T 2230/10 of
3 July 2015, reasons 3.7; see also T 258/97 of
8 February 2002, reasons 6).

3.7   Nevertheless, the jurisprudence of the boards of appeal
        acknowledges the possibility that the design of
        particular non-technical method steps to be implemented
        on a computer has been motivated by technical
        considerations, in particular concerning the internal
        functioning of the computer, resulting in a specific
        technical effect being achieved when the method is run
        on the computer (see decisions T 258/03, reasons 5.8;
        T 1358/09 of 21 November 2014, reasons 5.5; and
        T 2330/13 of 9 May 2018, reasons 5.7.9 and 5.7.10).

        According to opinion G 3/08 (OJ EPO 2011, 10), reasons
        13.5 and 13.5.1, such considerations would have to go
        beyond "merely" finding a computer algorithm to carry
        out some procedure. Mere algorithmic efficiency is
        generally not considered to be a technical effect (see
        decisions T 1784/06 of 21 September 2012, reasons
        3.1.2; T 42/10 of 28 February 2013, reasons 2.11;
        T 1370/11 of 11 March 2016, reasons 10 to 10.5; and
        T 2418/12 of 14 July 2017, reasons 3.3).

3.8   In the present case, the appellant's position is
        essentially that, in the context of a (computer-
        implemented) method of scoring a document on the basis
        of the frequency at and the amount by which the
        document's content changes over time, the decision to
        determine the frequency and the amount of changes
        between two versions of the document by comparing their
        term vectors requires technical considerations, in
        particular relating to memory usage.

If the appellant's point of view is correct, then that
decision cannot be included in the formulation of the
technical problem to be solved. Rather, it contributes
to the solution of the problem of implementing a method
of scoring a document on the basis of the frequency at
and the amount by which the document's content changes
over time in a memory-efficient manner.

3.9   According to a second point of view, determining the
       frequency and the amount of changes between two
       versions of the document by comparing their term
       vectors is merely an algorithmic and thus a non-
       technical solution to the problem of determining the
       frequency and the amount of changes. Although comparing
       document versions in their entirety may be the more
       straightforward solution, the degree of originality of
       a solution is not a criterion for technicality.

       If the decision is indeed non-technical, then it can be
       included in the formulation of the technical problem to
       be solved.

3.10   As a variation on the second point of view, it could
       also be argued that the non-technical purpose of
       claim 1 is not "scoring a document on the basis of the
       frequency at and the amount by which the document's
       content changes over time" but "scoring a document on
       the basis of the frequency at and the amount by which
       the document's term vector changes over time". Indeed,
       a term vector, being a vector of word frequencies, is
       not an inherently technical object.

       It is clear that the argument for the appellant and
       against this variation would be that this formulation
       of the non-technical purpose of claim 1 incorrectly

hides the technicality of the decision to use term vectors in the claimed context.

3.11   Although it cannot be denied that measuring the difference between two text documents by comparing their term vectors is an algorithmic solution, this does not on its own mean that the second point of view is the correct one.

For example, in decision T 650/13 of 2 October 2018, reasons 6, this Board confirmed the holding of the older decision T 107/87 of 26 April 1991 that a data coding rule for identifying and eliminating statistical redundancy contributes to the solution of a technical problem where it is used to reduce the amount of data to be stored or transmitted. This means that if a computer-implemented method includes steps of losslessly compressing and decompressing intermediate results to reduce the amount of memory space required for storing those results, at least those steps will make a technical contribution. Still, the implementation of the coding rule will normally be algorithmic in nature.

In the Board's view, the justification for attributing a technical character to a redundancy-reducing coding rule when used for reducing the amount of data to be stored or transmitted is that such rules can fairly be said to be based on technical considerations: they would have been formulated by an engineer in the field of digital signal processing rather than by a non-technical person such as the "notional mathematician" (Article 52(2)(a) EPC) or the "notional computer programmer" (Article 52(2)(c) EPC).

3.12   More generally, the Board considers that if non-
       technical claim features interact with technical claim
       features to cause a physical effect over the prior art,
       such as an effect on memory usage in a general-purpose
       computer, the physical effect is to be regarded as a
       technical effect for the purpose of assessing inventive
       step if the non-technical features are based on
       technical considerations aimed at controlling that
       physical effect (see e.g. decisions T 2230/10,
       reasons 3.8; and T 2035/11 of 25 July 2014, reasons
       5.2.3).

       A useful test for determining whether such technical
       considerations are present is to ask whether the non-
       technical features would have been formulated by a
       technical person rather than by a non-technical person
       or persons (see e.g. decisions T 1214/09 of
       18 July 2014, reasons 4.8.8; T 1321/11 of
       4 August 2016, reasons 5.3.5; T 1463/11 of
       29 November 2016, reasons 20 and 21; and T 136/13 of
       11 September 2018, reasons 3.6). This is not an enquiry
       into the actual state of technical or non-technical
       knowledge at the effective filing date; the question is
       rather whether the knowledge required for coming up
       with the non-technical features in the particular case
       is of a kind that only a technical person, i.e. a
       person not working exclusively in areas falling under
       Article 52(2) EPC, could possess.

3.13   Compared with techniques for lossless data compression,
       it is less evident that the idea of reducing a text
       document to a term vector to lower memory requirements
       while still being able to determine the amount of
       changes between consecutive versions is technical. The
       concept of determining the semantic similarity between
       documents by means of term vectors belongs to the field

of linguistics, which is a non-technical area falling
under Article 52(2) EPC (see decisions T 121/85 of
14 March 1989, reasons 5.7; T 1177/97 of 9 July 2002,
reasons 3 and 7; and T 2418/12, reasons 3.1). And the
idea to use this concept in a computer program to
reduce the amount of data to be stored is arguably one
that the notional computer programmer would have had -
more data requiring more memory being a concept
inherent to computer programming.

3.14   But in the present case the Board need not make a
judgment as to the technicality of the use of term
vectors in the context of claim 1, as the outcome of
the inventive-step assessment does not depend on it.

Accepting, for the sake of argument, the appellant's
position, the objective technical problem to be solved
is that of implementing, on a computer and in a memory-
efficient manner, a method of scoring a document on the
basis of the frequency at and the amount by which the
document's content changes over time.

Starting from a general-purpose computer and faced with
this problem, the skilled person would have realised
that memory can be used efficiently by storing the
current version of the document in a reduced form which
is still suitable for measuring the difference with
another document or document version. He would
therefore have looked for a suitable reduced form.

At the priority date it was well known that term
vectors, which the application mentions only once (in
the passage cited in point 3.5 above) and without
explaining it, were used for comparing the semantic
content of text documents. At the oral proceedings, the
appellant did not dispute this, but it argued that the

invention used them for a new purpose. However, the
Board judges that the skilled person would have
recognised that term vectors not only were suitable for
comparing text documents but also took up, at least in
the case of larger documents, less memory space than
the entire documents. He would therefore have chosen to
store the term vector of the current document version
and would so have arrived at the subject-matter of
claim 1 without the exercise of inventive skill.

3.15   Hence, the subject-matter of claim 1 lacks inventive
        step (Article 56 EPC).

4.   *First auxiliary request – inventive step*

4.1   Claim 1 of the first auxiliary request differs from
        claim 1 of the main request essentially in that it
        specifies that:

        -    a plurality of documents are monitored (but a score
             is generated for only one, and still on the basis
             of the history data obtained for that document);

        -    for each document of the plurality of documents,
             "portions of the documents that are determined to
             be most frequently occurring instead of [...] the
             entire document" are stored; and

        -    the score is based, at least in part, on the
             frequency at and the amount by which "the content
             of the respective stored document portion changes
             over time".

4.2   The wording of claim 1 suffers from a number of
        imprecisions.

First, the claim states that "portion**s** of the document**s**" are stored for each document. But it also states that the "data relating to changes to the content of a respective stored document over time" includes a frequency at and an amount by which the content "of the respective stored document portion" changes over time. It therefore appears that, for each document, only one portion of that document is stored rather than multiple portions of multiple documents.

Second, the claim leaves undefined what is meant by "portions of the documents" (or perhaps "the portion of the document") that are (is?) "determined to be most frequently occurring". To know whether something is "most frequently occurring", it is necessary to know what kinds of occurrences are being counted. A document portion could be "most frequently occurring" within the document itself or within the plurality of documents or within a document corpus external to the claimed method (e.g. a document corpus representative of the English language).

4.3   These imprecisions cannot be easily resolved by referring to the passage of the description on page 7, lines 3 to 5, on which the amendments are based. This passage reads as follows:

"According to another implementation, search engine 125 may store and monitor a relatively small portion (e.g., a few terms) of the documents that are determined to be important or the most frequently occurring (excluding 'stop words')."

Grammatically, this sentence states that a relatively small "portion of the documents", i.e. a relatively small subset of all documents, is stored. The

parenthesised qualifications "(e.g., a few terms)" and
"(excluding 'stop words')" do shed doubt on this
grammatical reading but do not clarify with any
precision what else could be meant.

4.4   In its letter of 29 October 2018, the appellant
        submitted that the parenthesised qualifications "a few
        terms" and "stop words" did provide clarification and
        that a "portion" did not need to be a contiguous
        section of text. In view of these submissions and the
        above-identified imprecisions, the Board judges that
        claim 1 still encompasses the use of term vectors,
        which essentially list the most frequently occurring
        terms in a document and their frequencies.

4.5   The subject-matter of claim 1 of the first auxiliary
        request is therefore further limited compared with
        claim 1 of the main request as interpreted in point 3
        above only in that a plurality of documents are
        monitored for changes. But if it is obvious to monitor
        a single document for changes, it is also obvious to
        monitor two or more documents for changes.

4.6   Thus, the subject-matter of claim 1 lacks inventive
        step (Article 56 EPC).

5.   *Second auxiliary request - inventive step*

5.1   Claim 1 of the second auxiliary request adds to claim 1
        of the main request essentially that a plurality of
        documents are monitored. As in claim 1 of the first
        auxiliary request, a score is generated for only one
        document, and still on the basis of the history data
        obtained for that document.

Claim 1 further makes explicit that a "signature" of each document is stored.

5.2   In point 3 above, claim 1 of the main request was already interpreted as specifying that a "signature" (or, more narrowly, a term vector) was stored for each document. The subject-matter of claim 1 of the second auxiliary request therefore lacks inventive step within the meaning of Article 56 EPC for the reason given in points 3 and 4.5 above.

6.   *Conclusion*

Since none 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 Chairman:

I. Aperribay   R. Moufang

Decision electronically authenticated