

# PATENT COOPERATION TREATY

From the  
INTERNATIONAL SEARCHING AUTHORITY

# PCT

WRITTEN OPINION OF THE  
INTERNATIONAL SEARCHING AUTHORITY  
(PCT Rule 43*bis*.1)

To:

see form PCT/ISA/220

Date of mailing  
(day/month/year) see form PCT/ISA/210 (second sheet)

Applicant's or agent's file reference  
see form PCT/ISA/220

**FOR FURTHER ACTION**  
See paragraph 2 below

International application No.  
PCT/EP2018/071829

International filing date (day/month/year)  
10.08.2018

Priority date (day/month/year)  
11.08.2017

International Patent Classification (IPC) or both national classification and IPC  
INV. G06F17/30

Applicant  
INFOSUM LIMITED

1. This opinion contains indications relating to the following items:

- Box No. I Basis of the opinion
- Box No. II Priority
- Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- Box No. IV Lack of unity of invention
- Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- Box No. VI Certain documents cited
- Box No. VII Certain defects in the international application
- Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will usually be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

Name and mailing address of the ISA:



European Patent Office  
P.B. 5818 Patentlaan 2  
NL-2280 HV Rijswijk - Pays Bas  
Tel. +31 70 340 - 2040  
Fax: +31 70 340 - 3016

Date of completion of  
this opinion

see form  
PCT/ISA/210

Authorized Officer

Moran, Matthew

Telephone No. +31 70 340-0



---

**Box No. I Basis of the opinion**

---

1. With regard to the **language**, this opinion has been established on the basis of:
  - the international application in the language in which it was filed.
  - a translation of the international application into , which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1 (b)).
2.  This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43bis.1(a))
3.  With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing:
  - a.  forming part of the international application as filed:
    - in the form of an Annex C/ST.25 text file.
    - on paper or in the form of an image file.
  - b.  furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
  - c.  furnished subsequent to the international filing date for the purposes of international search only:
    - in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
    - on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
4.  In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
5. Additional comments:

---

**Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

---

1. Statement

Novelty (N)	Yes: Claims	<u>1-18</u>
	No: Claims	
Inventive step (IS)	Yes: Claims	
	No: Claims	<u>1-18</u>
Industrial applicability (IA)	Yes: Claims	<u>1-18</u>
	No: Claims	

2. Citations and explanations

see separate sheet

---

**Box No. VIII Certain observations on the international application**

---

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

1 **Re Item V**

**Reasoned statement with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1.1 Reference is made to the following documents:

D1 US 2015/242407 A1 (FROHOCK RON [US] ET AL) 27 August 2015 (2015-08-27)

D2 US 2017/075898 A1 (DESHPANDE PRASAD M [IN] ET AL) 16 March 2017 (2017-03-16)

**Independent Claims**

1.2 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claim 1 does not involve an inventive step in the sense of Article 33(3) PCT.

1.3 Notwithstanding the issues raised below in section §2 with respect to clarity, D1 is regarded as being the prior art closest to the subject-matter of claim 1, and discloses (the references in parentheses apply to D1):

A method of joining a first dataset with a second dataset, the first dataset configured to store a set of data entries each identified by a respective key of a first type and the second dataset configured to store a second set of data entries identified by a respective key of a second type

*(§19, §22, claims 4 + 5 - joining a first and second dataset using related attributes in the respective datasets as a join key wherein the attributes have defined types),*

the method comprising:

- selecting an intermediate mapping entity ~~from a set of possible intermediate mapping entities~~, each mapping entity storing association between keys of the first type and keys of the second type; *(directly implied by §28 - a key transform translates values of attributes of one type into values of attributes of another type. This is understood to imply that an association between the types is known, and stored, in advance of the translation)*

- providing the selected intermediate mapping entity for use in joining the first data set with the second data set;  
(§28)
- ~~wherein the step of selecting the intermediate mapping entity is based on the intersection weight between the first and second data sets via each of the intermediate mapping entities,~~
  - wherein the intersection weight is the proportion of overlapping data entries between the first and second datasets  
(§19, §25, §43, §49, §51 and §52: *A Relationship Confidence Metric (RCM) is used to measure the overlap between datasets. The RCM is calculated for a relationship datasets which indicates a probability of the strength of the relationship between the values of the attributes in the respective datasets, using various advisors which check for similarities between the values and types of attributes and which determine if the attribute values represent entity IDs. The results of said advisors are weighted and included in the RCM.*).

1.4 The subject-matter of claim 1 therefore differs from this known D1 in that:

- selecting an intermediate mapping entity from a set of possible intermediate mapping entities, each mapping entity storing association between keys of the first type and keys of the second type (*note: the underlined feature is that which differs from the disclosure of D1*);
- wherein the step of selecting the intermediate mapping entity is based on the intersection weight between the first and second data sets via each of the intermediate mapping entities

1.5 The effect of the above differences, taken in combination, is considered to be the detection of a mapping entity which reduces the number of missing translations between potentially overlapping datasets.

1.6 Therefore, the corresponding problem may be considered to be how to modify D1 to reduce the number of missing translations between potentially overlapping datasets.

1.7 The solution proposed in claim 1 of the present application cannot be considered to involve an inventive step (Article 33(3) PCT) for the following reasons.

- 1.8 The skilled person would be aware, from the teaching of D1, of the concept of use an intermediate mapping entity between datasets to improve the chance of detecting overlaps, or intersections, of entities between the two datasets. The skilled person would further be aware, from their common general knowledge in the technical field of the invention, that intermediate mapping entities can be automatically obtained to facilitate the detection of overlapping entries between datasets, as exemplified by D2 (§2-6).
- 1.9 When faced with the above problem, therefore, the skilled person would know that a plurality of intermediate entity mapping entities can exist to assist in the detection of overlap between datasets. Hence, the skilled person would consider it a trivial task, without any inventive activity, to apply each of a set of said intermediate mapping entities to the pair of datasets, in a trial-and-error fashion, to determine which of said intermediate mapping entities leads to the calculation of an intersection weight, considered to be the best for said pair of datasets.
- 1.10 Therefore the solution to the above identified problem, provided by the differing features of the present claim 1, is not considered to contribute an inventive step with respect to Article 6 PCT.
- 1.11 The subject-matter of independent claim 17 corresponds to that of the method of independent claim 1, albeit in apparatus form. Thus, the reasoning as set out above applies, *mutatis mutandis*, to the subject-matter of said independent claim 17 which therefore also does not represent an inventive step under Article 33(3) PCT.

### **Dependent Claims**

- 1.12 Dependent claims 2-16 and 18 do not appear to contain any additional features which, in combination with the features of any claim to which they refer, meet the requirements of the PCT in respect of inventive step, the reasons being as follows.
- 1.13 Claim 2. D1 discloses the use of a mapping entity which is understood to represent a *path* between datasets. The location and definition of a plurality of paths does not appear to contribute any further technical effect with respect to Article 33(3) PCT.

- 1.14 Claims 3, 4. Choosing a set of parameters and a corresponding heuristic to determine the best path would be a conventional choice for the skilled person when choosing between available paths representing entity mappings. Furthermore, in the absence of any technical detail on the features of said claims, the subject matter of said claims does not appear to contribute any further technical effect and, hence, is not considered to be inventive.
- 1.15 Claims 5 and 6. The concept of calculating an overlap intersection weight between two intermediate entity mappings is not at all clear as intermediate mapping entities are understood to be "store association[s] between keys of the first type and keys of the second type (*see description, page 2, lines 27-29 and also present claim 16*). It is not possible to take a clear understanding of the above claims, even in light of the description, and hence, it can not be considered to contribute an inventive step.
- 1.16 Claim 7 and 18. Storing data, e.g. intersection weights, in association with data to which it relates, e.g. intermediate mapping entities, is considered to be a trivial task and hence, does not contribute an inventive step.
- 1.17 Claims 8 and 9. Analogously to claim 5, it is unclear how an intersection weight can represent an overlap between a data set and an intermediate mapping entity. Even if it were to be credibly demonstrated that the weight represents an overlap between a data set and the output of an intermediate mapping entity, the direction of the mapping is a trivial design choice which does not contribute an inventive step.
- 1.18 Claims 10 and 11. The choice of how to calculate the proportion represents an arbitrary design choice and no apparent further technical contribution is apparent.
- 1.19 Claim 12. The subject matter of this claim is not considered inventive for the same reasons as those provided above in sections §1.8 and §1.9.
- 1.20 Claim 13. Using a combination of intermediate entity mappings appears to be a variation of the use of intermediate mapping entities which would be obvious to the skilled person without any inventive step.
- 1.21 Claim 14. It is not clear what further technical effect is achieved by the subject matter of the claim as it encompasses that any one of the multiple intermediate mapping entities, independent of calculated intersection weights, is used to join the two datasets in each of the two steps and then the results of the two steps are joined again. In the absence of any clear further technical effect, the subject matter of said claim is not considered inventive.

- 1.22 Claim 15. The subjective and technically-unclear parameters of *quality* and *speed* can not contribute a further technical effect in the calculation of the subjective decision of what a *best plan* is. Hence, claim 15 is not considered to be inventive.
- 1.23 Claim 16. Notwithstanding several clarity issues, the subject matter of claim 16 is understood to be the identification of keys in a second data set that overlap with translated keys from a first dataset. This is considered to be indicated by D1 (§28) and known from the automatic identification of mappings between datasets, as exemplified by D2 (§2 and §6). Hence claim 16 is not considered to be inventive.

## 2 **Re Item VIII**

### **Certain observations on the international application**

- 2.1 The application does not meet the requirements of Article 6 PCT, because claims 1, 4, 5, 6, 8, 9, 15 and 16 are not clear.
- 2.2 Claim 1, line 1, recites a "method of joining a first dataset with a second dataset". However, there appears to be no feature in claim 1 which carries out an action of joining two datasets. Instead, an intermediate mapping entity *for use* in joining datasets is selected and provided. Therefore the claim does not clearly achieve its stated aim which renders the scope of the claim unclear with regard to Article 6 PCT.
- 2.3 Claim 1. At lines 5 and 6, it is unclear whether the "respective key of a second type" identifies a set of data entries or each entry in a set. This renders the scope of the claim unclear with regard to Article 6 PCT.
- 2.4 Claims 1, 5-6, 8-9 and 16. The term *intermediate mapping entity* appears to have varying inconsistent meanings in said claims which renders the scope of said claims unclear with regard to Article 6 PCT. For claims 1 ad 16, in light of the description, the term is understood to mean an entity which "store[s] association[s] between keys of the first type and keys of the second type (*see description, page 2, lines 27-29 and also present claim 16*). However, in claims 5-6 and 8-9, the term appears to be treated as equivalent to a data set or to an element of a data set. This inconsistency renders the subject matter of said claims unclear with respect to Article 6 PCT.



- 2.5 Claim 1. The step beginning "wherein the step of selecting the intermediate mapping entry is based ..." is not clear, as it is presently worded, rendering the scope of claim 1 unclear with regard to Article 6 PCT. In light of the description, the step is construed to mean that all intermediate mapping entities in the set, identified at line 7, are used with the pair of datasets, in a trial-and-error fashion, to determine which of the intermediate mapping entities provides what is considered to be the best overlap between the two datasets.
- 2.6 Claims 4 and 15. The following terms are non-technical and hence the scope of the features to which they refer is unclear with regard to Article 6 PCT: single best path (claim 4), *quality* and *speed* (claim 15).