# Title: Freshworks Invention Disclosure Form (IDF) - Smart connections

Instructions : Fill each blank with the requested information or enter NONE as appropriate. Where space on the form is inadequate, you can use links (documents) that points to more details. Please submit the filled IDF by filing a ticket at legal.freshservice.com

(You can replicate this form in confluence, fill it up and export as PDF and attach it to the ticket that you file in legal.freshservice.com, see solutions article <https://legal.freshservice.com/solution/categories/3000052696/folders/3000082583>)

|  |
| --- |
| PART I BASIC INFORMATION |

### 1.a DESCRIPTIVE INVENTION TITLE

|  |
| --- |
| Customisable system for identifying related entities in a multi-tenant system |

### 1.b INVENTOR(S) - list yourself and any colleagues who worked on the idea with you

|  |  |
| --- | --- |
| Name | email address |
| Srivatsa Narasimha | [srivatsa.narasimha@freshworks.com](mailto:srivatsa.narasimha@freshworks.com) |
| Bharathi Balasubramaniam | [bharathi.balasubramaniam@freshworks.com](mailto:bharathi.balasubramaniam@freshworks.com) |
| Suvrat Hiran | [suvrat.hiran@freshworks.com](mailto:suvrat.hiran@freshworks.com) |
| Swaminathan Padmanabhan | [swaminathan.padmanabhan@freshworks.com](mailto:swaminathan.padmanabhan@freshworks.com) |

### 1.c INVENTOR INFO - Are any inventors located outside of India , are any of them not employed Full Time by Freshworks ?

|  |
| --- |
| No |

### 1.d PRODUCTS/TIMELINE - What products will incorporate your idea and when will these products be released ?

Smart connections feature released in Oct 2019 (Freshsales)

### 1.e PRIOR PUBLICATION - Identify any publishing or journals that you referred to during formation of ideas

1. <https://www.academia.edu/27404789/Holistic_Entity_Matching_Across_Knowledge_Graphs>
2. <https://www.academia.edu/21004651/Record_linkage_with_uniqueness_constraints_and_erroneous_values>
3. <http://www.cs.utexas.edu/~ml/papers/marlin-dissertation-06.pdf>
4. [M. Bilenko and R. J. Mooney. Adaptive duplicate detection using learnable string similarity measures. In Proceedings of ACM SIGKDD2003, Washington, DC, 2003.](http://www.cs.utexas.edu/~ml/papers/marlin-kdd-03.pdf)
5. [R. Ananthakrishna, S. Chaudhuri, and V. Ganti. Eliminating fuzzy duplicates in data warehouses. In Proceedings of VLDB-2002, Hong Kong, China, 2002.](https://www.cse.ust.hk/vldb2002/VLDB2002-proceedings/papers/S17P01.pdf)

      6. [K. Goiser and P. Christen. Towards automated record link-age. AusDM ’06: Proceedings of the fth Australasian con- ference on Data mining and analystics , pages 23–31, 2006.](https://pdfs.semanticscholar.org/393c/d1728ac1e5b798b59ac75c67a9f7092bd620.pdf)

|  |
| --- |
| PART II DISCLOSURE OF INVENTION |

### 2.a PURPOSE State the problem / challenge addressed by your idea.

We developed a system to process and store potential related entities in a multitenant database. Entities here include records of people and companies; and relationships could be a potential duplicate entity in a database, a record of a person associated with a company, or a company which is subsidiary of another company. This system processes the entity as it is added to the database and stores the relationship in a manner such that associations and duplicates can be fetched in O(1) time complexity. It also gives flexibility to the users of the system to customise what relationships should be suggested as related. Once relationships between entities are confirmed by the user, they can be merged, marked related, or not related.

### 2.b PRIOR EXISTING TECHNOLOGY - Describe the previous old methods used by others and why those methods fell short. Also describe the technologies that will be integral to your solution that we do not intend to claim as an invention. This is a good place to put background that will facilitate understanding section 2c.

Most systems work in isolation, they store entities, deduplicate records or explicit relation between people and company is maintained. For deduplication alone several methods exist to identify if two records are duplicates. In other scenarios, the system can associate company websites and company email with domain to identify associations. All existing systems lack ability to continuously process any new entity entered in the system to first search for existing duplicates followed by searching for associated entities (relationships) in the entire database. To add to this, duplicates and associations (relationships) should be identified in spite of having different kinds of errors (spelling mistake, phonetic errors, partial match). Our system ensures that when searching for associations (relationships), we restrict the search space and keep it under a few milliseconds. No prior system lets users adjust what should be seen duplicate/association by adjusting simple toggle bar. In most cases you have to specify rules like one edit distance is acceptable or phonetic match should be allowed, etc.

Some unique features -

1. Identification of associated companies based on similarity in company name, website, domain and emails of users associated to those companies. E.g. [abs@freshdesk.com](mailto:abs@freshdesk.com) will be associated with the "Freshworks" company.
2. Providing flexibility to users to adjust what should come as duplicate or associations using simple toggle bar or dynamically learning thresholds based on sample provided to user.
3. Associations are searched and restricted by user provided inputs. For example, we would only allow associations to be found if the records are created by the same user.
4. System learns continuously based on feedback provided by the user.

### 2.c Description of invention (use flow charts or other pics to explain)

