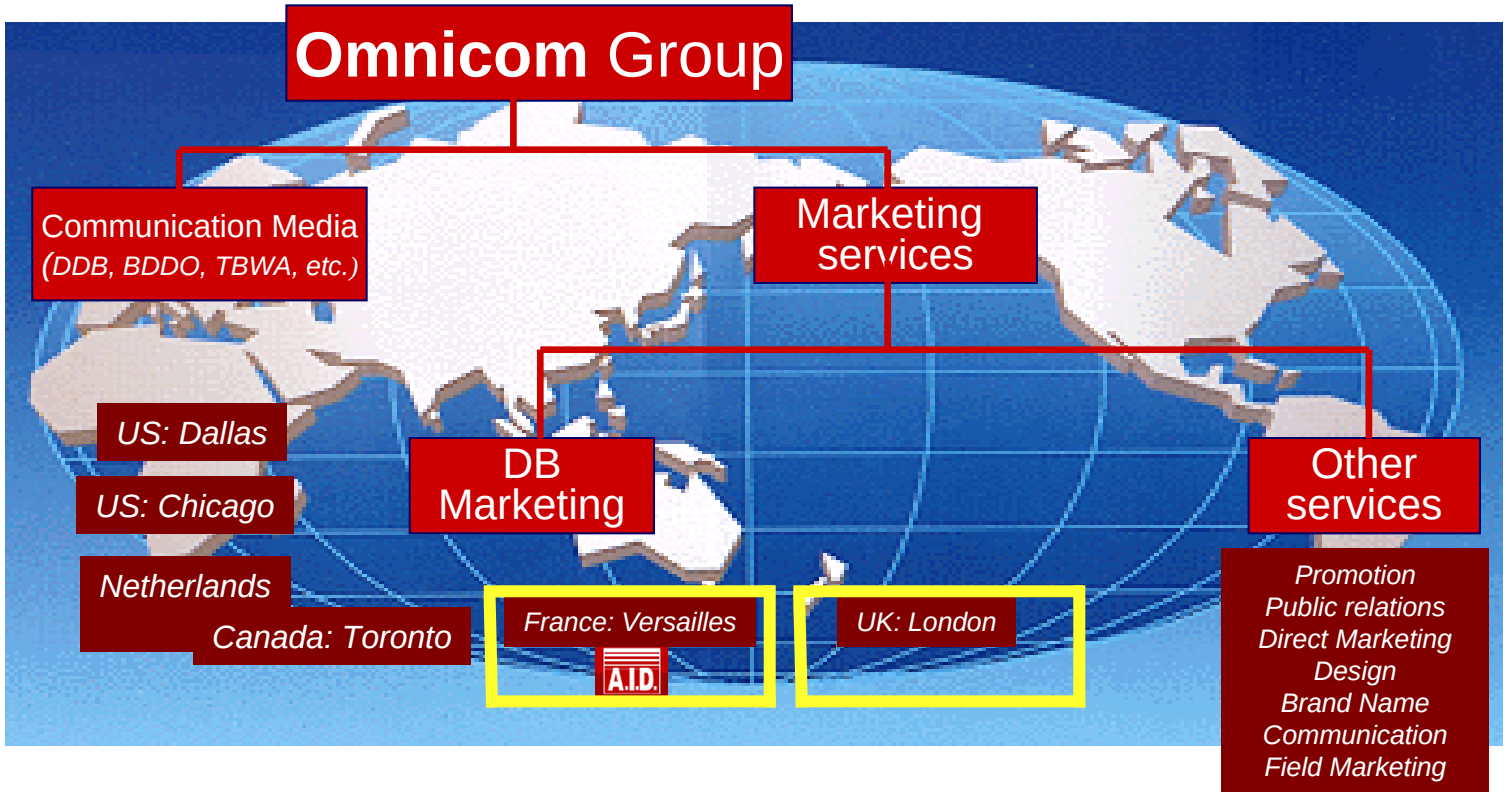
B.D.Q.S.

Why we use it at A.I.D.

GIVING MEANING TO YOUR DATA

A.I.D. – Part of Omnicom group

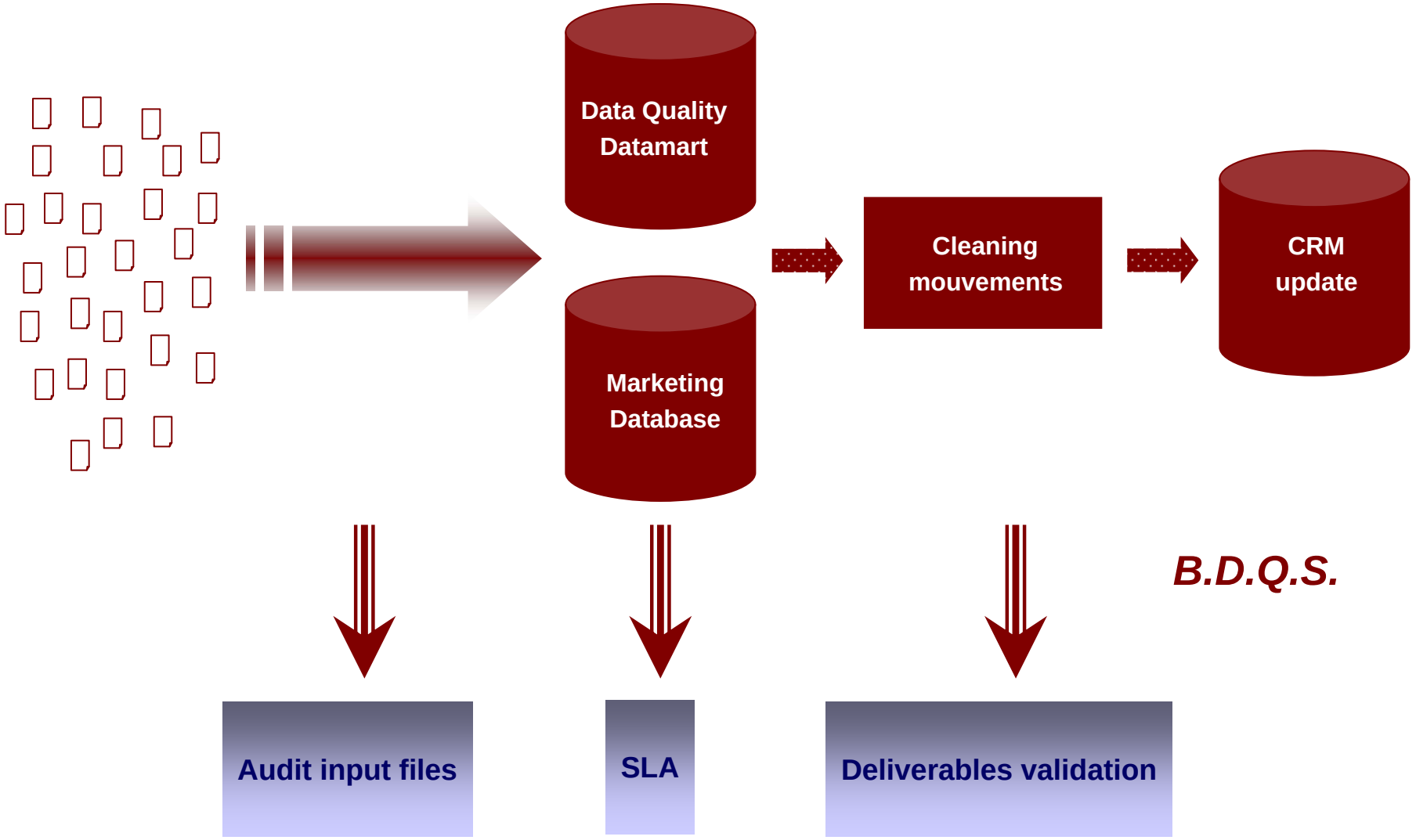
- Omnicom Group is one of the leading international groups in the communications industry.
- A holding company at the forefront in the fields of communications, public relations, marketing and database services, Omnicom is present in over 100 countries.



A range of Services and solutions

Information Systems - CRM	Data Quality	Marketing Performance	Decision Support Marketing
Data Concept	Data Quality Suite		Statistic Drive
Data Develop	Datamart Quality		MarkeTwin
Data Management	Base for DQ System		
Data Storage			
Campaign Management		Market Tracker	Marketing Optimisation
E-messaging & e-mail deployment		Campaign Tracker	

BDQS = To do what ?



BDQS = To do what ?

- **To audit** an input file:
  - A file is received for deduplication, direct marketing operation or integration into a marketing database
  - The contain, the parsing, LOVs, patterns are controlled thanks to BDQS

  ➔ **To list the cleanings to put in place before to use the file**

- **To measure on a regular basis** a system: marketing database, CRM,..hosted by A.I.D:
  - To publish a data quality scorecard
  - To validate our **SLA** engagement using a **tool different from the production tool**
  - To give to the production tracks of improvement

- **To validate** a file before delivery:
  - After a deduplication operation or a CRM cleaning, the files are controlled before delivery
  - Type of control: LOVs, uniqueness, integrity, fake values,…

BDQS = Algorithms available

- List of values, integrity, uniqueness
- Pattern analysis
- Fake values: characters, words, repetitive characters
- Emails check, phone number check
- Duplicates: sophisticated tool based on a statistical method and taken into account data quality indicators
- Statistical analysis implementation in progress and specific algorithms can be implemented by the client

BDQS = Algorithms available
Example of duplicates parameters

```
condition C7
{
        compare_not_null last_name { editdist <= 1 }
        and compare_not_null first_name { letters >= 80 , skel > 50 }
        and compare address { letters >= 80 , skel > 75 }
        and compare_not_null country { equals }
}
```

All the attributes can be combined thanks to indexes as Edit Distance

```
group email {
     conditions {
     C1 (11)
     or C2 (12)
     or C3 (13)
     or C4 (14)
     or C5 (10)
     or C6 (15)
        }
}
```

Compare all the records with the same email

```
group tel {
     conditions {
     C7 (22)
     or C8 (23)
     or C9 (24)
     or C10 (5)
        }
}
```

Compare all the records with the same phone number

BDQS = Algorithms available
How to query the uniqueness

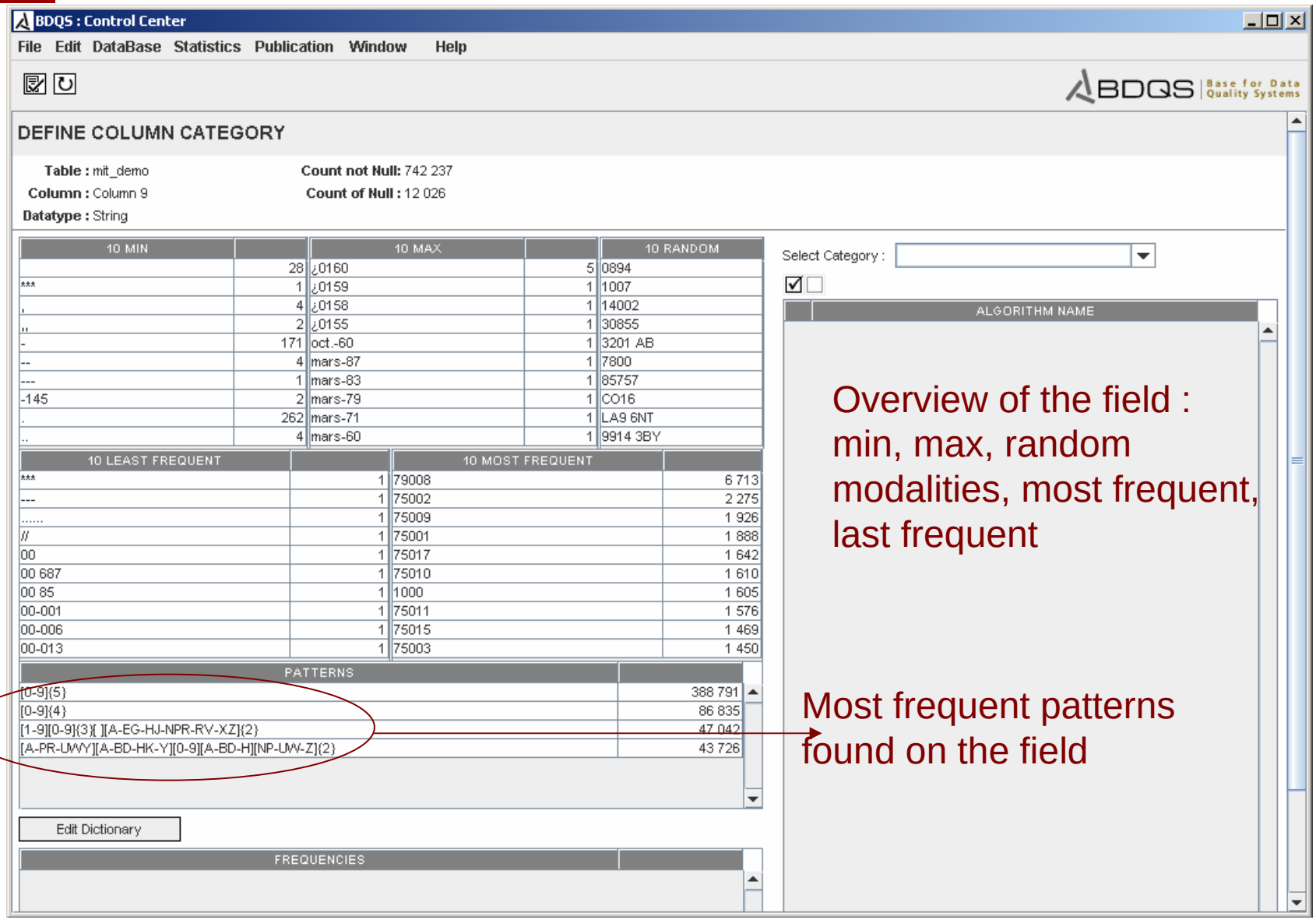
Example :

					Level of confidency
1	Smith	Albert	salbert@hotmail.com		
2	Smitt	A	salbert@hotmail.com		12
3	Smithe	Alberd	salbert@hotmail.com	0139239345	14
4	Smiss	Albert		0139239345	21
5	Smiss	Albert		0139239345	5

2 dedicated queries :

- **Number of masters ?**
  - **Depending on the confidency level accepted**
- **Number of duplicates ?**
  - **To analyze the group of duplicates : origin, …**

BDQS = Algorithms available
Example of pattern detection



Overview of the field : min, max, random modalities, most frequent, last frequent

Most frequent patterns found on the field

BDQS = Access to the data

- **Targeting the entire population, not only samples**

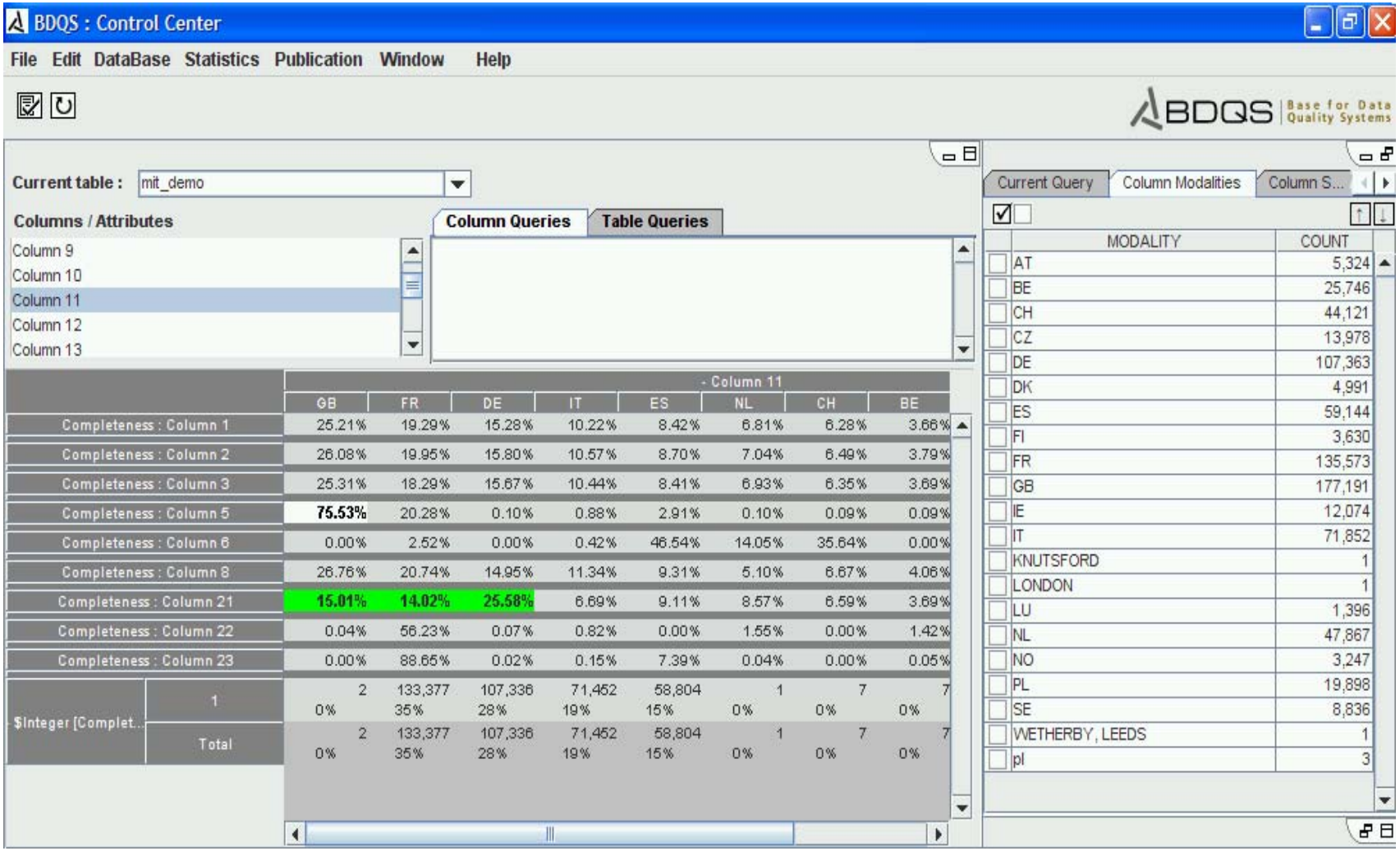
A query tool user-friendly and very fast

Only few seconds for complex queries on millions and more records

To find specific cases

An operational measurement: populations to clean can be directly extracted

BDQS = Access to the data



Questions

- To contact us:

Olivier Coppet

Brigitte Laboisse

A.I.D.

00 33 1 39 23 93 00

http://www.aid.fr