UMBC

**IS 722: Systems & Information Integration**

**Project Deliverable III: Integrating information from multiple repositories**

**DESCRIPTION**

Now that we have created an integration layer on top of the participating databases, we can proceed to integrate information from these databases. Currently there are two layers: the global layer (with all the metadata information and canonical representation) and the local layer with the participating databases (local DBs).

Global users submit global queries to the global layer of the system. A global query is decomposed into a set of subqueries, one per each participating local database. Each subquery is submitted to the appropriate local database, where it is executed and the results are shown to the user. Some details for this final phase of the project follow:

* Input queries are expressed in the canonical format. They are called global queries. This means that one needs to TYPE a sql query using the canonical names of fields in the integration/metadata layer. However, the global query needs to be typed by the global user as input to the system for further processing; it cannot run by itself (it will produce syntax errors, since there is no actual database with the table names and fields which show up in the global query).
* When a global query is received, it must be split into subqueries, one for each local database. String manipulation functions in Oracle PL/SQL may be used for this. Do not write hard-wired code that works only for specific global queries. Make code as generic as possible. Remember, this code is to *automatically generate code*.
* Once the subqueries have been generated (each one as a string) one can use Dynamic SQL to execute the subqueries. Execute each subquery and show the result for each subquery separately.
* Put together the results of the subqueries by using a UNION operation. Make sure you have union compatible results to demonstrate this case
* OPTIONAL: Consolidate the results of the local subqueries using a JOIN to produce the final answer to the global query

**DELIVERABLES**

**This is the demo of the entire project. You have about 20 minutes to show how your product works:**

1. **[10 pts] Presentation.** Create a short slide presentation with the architecture of your system (boxes connected together) and the ER diagram. Make sure to mention any items that are worthwhile that may differentiate your product from the competition. All members of the team must participate in the presentation.
2. **Demo.** This is the major part. It shows all you have learned from the entire course. You will run the system in real-time and you will show the entire class the flow of certain scenarios (global queries) as they are passing through the system. Prepare a script with the global queries and the appropriate calls. This script should start with DELETE commands, and then CREATE TABLE commands, etc. The script starts on a clean database, populates it, and runs the code. Make sure to show:
   1. **[5 pts] Global query.** The global query as a string
   2. **[55 pts]Subqueries.** Each subquery as it has been automatically generated by your system
   3. **[10 pts] Partial Results.** The temporary tables that contain the results for each submitted subquery to each local database
   4. **[20 pts] Overall result**. The consolidated result of multiple subqueries