Q1. How would you solve this? Might be that the solution is straightforward. Please write it down even if it seems stupid or simple.

I would make a request through the api in order to obtain the metadata, like for example in Oracle using the metadata API, i would make a request based on the table schema and name. The metadata should be also fetched in a XML file for easier parsing. Also the decision of how to categorize the files should also be made on the server.

Q2. What are the considerations in terms of performance? What are the solutions you are thinking for speeding this computation up.

For speeding up the computations I would consider using indexes storing the most searched columns from views or tables, obtaining this most searched data from the in-memory search engine on the server.

Q3. What if you have 1 billion distinct queries? How would that affect the overall time of computing this ? How can you speed it up ?

I think something similar to Q2 with a lot of indexes and containerization in such a way that the search engine doesn’t need to go through a lot of data to obtain the necessary information