Designing of Exotic Germplasm Database Using Client/Server Technology

RC Agrawal[1] and Manoj Kumar Burman[2]
[1]*National Bureau of Plant Genetic Resources, New Delhi-110 012*
[2] *Tezpur University, Department of Computer Science, Napaam, 784 028 (Assam)*

An efficient management of large volume of data related to Plant Genetic Resources (PGR) has become very essential for researchers, agricultural scientists and policy makers for research and reference purposes. An attempt is made to address the problem of managing data related to germplasm material imported by National Bureau of Plant Genetic Resources. A relational information system is designed and developed based on client-server approach. The detailed study also includes the taxonomic details and the addresses of national and international institutes. Some issues and steps related to the database designing of information management in the area of import of exotic material have been discussed. The detailed Entity-Relationship (ER) diagram and Data-Flow Diagrams (DFD) are also included which can help in the detailed designing of the database.

Key Words: Data Flow Diagram (DFD), Database Management System, Diagram, Entity Relationship (ER), Germplasm Exchange, Plant Genetic Resources

National Bureau of Plant Genetic Resources (NBPGR, hereafter referred to as Bureau) is the only nodal agency for import/export of germplasm material for agricultural research in India. It maintains a large number of crop species which include the "local" species (that grow naturally and collected from within the country) called "*Indigenous Collections*" (IC) and the "foreign" species (species that have been brought in/imported from various parts of the world) called "*Exotic Collections*" (EC). A wide variety of crop species have been imported and distributed by the Bureau for research purposes against requests by researchers and scientists in the area of agricultural sciences. The Bureau has been engaged in this activity since 1946 and nearly 0.5 million of samples have been received and distributed under import.

Any scientist, researcher or organization requiring to import some material has to make a formal request to the Exchange Division of the Bureau which in turn after scrutiny, issues an Import Permit (IP) and forwards the same to the source institute/country with a copy for information to the indentor. Imported exotic material may be received first-hand only at the Bureau where it undergoes for quarantine procedures to screen for interceptions of any pathogen. Any infected material is either cleaned or immediately destroyed. Only clean(ed) samples are returned to the Exchange Division where this material is accessioned before being distributed to the indentor. Accessioning is a process of giving a unique sample number called "EC number" to each and every sample to be used for any future reference. Each recipient including the original indentor needs to provide a feedback on the utility of material received by them.

Till date, information about all the exotic collections were maintained manually, the volume of which has reached around 200 registers. One of the major hurdles faced by the scientists and research management people is the management of this large amount of EC data. The EC data includes a variety of information such as the crop and species name, crop variety, type of material (such as seed, fruit, stem, root, culture etc.), alternate identity, other characteristics (such as plant quarantine, grow conditions etc.) and place of collections for material imported from various countries. The EC data may, therefore, be classified into six inter-related groups, namely, the *taxonomy* data (which includes details of all the crop species, their classifications and characteristics etc.), the *germplasm material* data (which includes the taxonomic details, quantity and other characteristics of the material imported from abroad), the *indentor* data (which includes details of the scientist, researcher or organization placing a request for import, namely, national addresses) the *source* data (which includes details of the source institute, organization from where the material is to be imported, namely, international addresses), the *import* data (which includes information about the material actually received and quarantined) and the *feedback* data (which includes information about feedback on material received and utilized by the indentor).

Maintaining and accessing above information in record books poses several data problems. Hence, what is required is a system to manage the voluminous exotic data in an efficient way and to provide quick retrieval, is an imperative. Apart from the need to aid the Bureau in proper and efficient management, maintenance of

information such as whether a species is a local or foreign species, it has become essential for documentation of such information in view of the recent global developments such as *General Agreement on Trade and Tariffs (GATT).*

In this paper we attempted to explain some issues and steps related to the database designing of information management regarding import of exotic material. Keeping in view the above mentioned facts, the following objectives were defined.

(i) ***Primary objectives:*** (a) to study and analyze the current manual system (requirement analysis), (b) to design and normalize various data tables to a reasonable degree, (c) to create and maintain a database of plant genetic material imported by the NBPGR and (d) to develop and implement a multi-user, user-friendly web enabled information management system.

(ii) ***Other objectives:*** (a) to facilitate the Bureau in bringing it at par with documentation capacities of other international organizations of repute engaged in similar research activities, (b) to assist NBPGR in maintaining a strong documentation base to tackle problems related to IPR (Intellectual Property Rights), Patents etc in the changing global scenario, and (c) to assist researchers, agricultural scientists and policy makers to access online the information about germplasm for research and reference purposes.

The detailed Entity-Relationship (ER) diagram and Data-Flow Diagrams (DFD) are also included which can help in the detailed designing of the database.

Problem Statement and Requirement Analysis

Any indentor, who wants to import some germplasm, has to make a formal request to the Bureau which after scrutiny issues an Import Permit for the purpose. The steps currently being followed manually and problems anticipated by the client in the process were identified and are listed below:

(1) Indentor places a request by filling a pre-printed form available at Germplasm Exchange (GEX) Division of the Bureau for import, giving the details of material required. The indentors request for material may be very vague or specific e.g. some samples of tropical fruits or say 5-6 samples of Tomato (*Lycopersicon esculentum* L.) exhibiting resistance to a particular disease. However, at a time same indentor may request for importing multiple materials from a single source. The client wishes to see the probable source(s) of requested material, which in the current manual system is difficult. Each request is registered and a Diary No. is assigned to the indentor for any further reference.

(2) GEX Division scrutinizes the request. Here, the client wishes to verify a history of material(s) imported by same indentor in recent past (up to 5 years) which in the current manual system is difficult.

(3) For valid requests, an Import Permit is generated and issued by the Bureau. An acknowledgement along with the Import Permit *number* (for reference) and a copy to this effect is sent to the indentor for his/her information. Separate permit is issued against each request for import from one country.

(4) The Import Permit is forwarded to the source country/ institute for sending the material. The supplier institute may send the requested material along with necessary documentation. The client needs to issue acknowledgement to the indentors and reminders to non-responding source(s) and keep a track of pending requests. The imported material can be received first-hand only by the Bureau and not directly by the indentor. The client requires crop-wise, country-wise and year-wise reports of imported material.

(5) Material (Consignment) received from abroad is given a *PQ number* and forwarded by GEX Division to the Plant Quarantine (PQ) Division to screen for any possible infection (due to virus, bacteria, disease etc.). Only clean(ed) samples out of the consignment are returned by the PQ Division to GEX Division. Here the Bureau needs to maintain a record of material sent to and cleared by PQ Division concerning quarantine. Status of cases pending with PQ Division is also needed. Thus the correspondence between PQ and GEX is mainly on the basis of PQ number.

(6) Each sample received by GEX Division after quarantine is accessioned *i.e.*, given a unique identification number called Exotic Collection (EC) number. From here onwards this *EC number* is the reference for all future operations of these samples.

(7) After accessioning, the material is passed on to the original indentor. **Multiple material distributions** are to be handled. Here, the client also needs to maintain a record of material distributed crop-wise, country-wise, and year-wise and category-wise which in the current manual system is a very cumbersome process.

(8) Each recipient of the material is expected to provide a feedback on the material utilized by him/her. The client needs to send reminders to non-responding parties in the matter. All feedback responses are to be compiled and made available for reference by other researchers. This is expected to be an EC number or crop name-based activity.

(9) Periodic reports and a versatile query system are also imperative needs of the client.

Advantages of the Information System over the Manual System

There are many drawbacks in the existing system. The present system was completely manual and all the records were kept in paper files that stand at a risk of getting damaged through gradual deterioration. The system, like other manual systems has some limitations, which can be enumerated as follows:

(i) Since the volume of data is very large, it is very difficult and inefficient to maintain them, (ii) The process of generating report is very time consuming, (iii) Retrieving information is time consuming, (iv) Time to time updating of records is very difficult, (v) Involvement of lots of manual works, (vi) Manipulation and repudiation possible and (vii) Lack of structured storage of data and efficient summarization.

The proposed system shall overcome the drawbacks of the existing system (manual system) in the following ways:

(i) Increase efficiency and reliability by suitable electronic medium for data storage, (ii) Increased information security by using suitable security model for information storage and processing, (iii) Increased effectiveness by providing efficient and convenient information retrieval methods and by providing value added information output, (iv) Increased efficiency and convenience by computerizing processing, (v) Backup can be done easily, (vi) Updating of records is an easy task, (vii) Generating of Reports is quite simple, (viii) Recovery of data can be done after any system crash, (ix) Information can be available online, (x) Administration of database can be done online and (xi) Increased productivity expectancy with the changed work environment.

The Platform/tools Required

The system was developed with following hardware and software tools:

(a) Hardware

One of the available P-IV (@1.5 GHz, 256MB RAM, 40GB HDD) desktop machines was used as Server.

Server : Intel Pentium III or above, Min. 256 MB RAM and 40GB HDD

Client : PII or above desktop with minimum 32 MB RAM

(b) Software

Following is a reference list of software used to develop and implement the software:

Operating System : Linux 7.3 (Server)

Client-Side Interface : Browser (I.E., Netscape etc.)

Database Server *(Back-end)* : ORACLE 9i for Linux

Web Server : ORACLE HTTP Server (Middle Tier)

System Architecture

It is based on a 3-tier Client Server model (Fig. 1).

(a) Database Server

At the lowest layer PGR-IMS employs the Database Server. The Server receives requests from clients for various operations. The server is responsible for concurrency control, recovery, transaction and storage management. The Database Server is responsible for query processing and all the operations on PGR-IMS database. After initialization, the database server waits for a query from its client. The results of the query

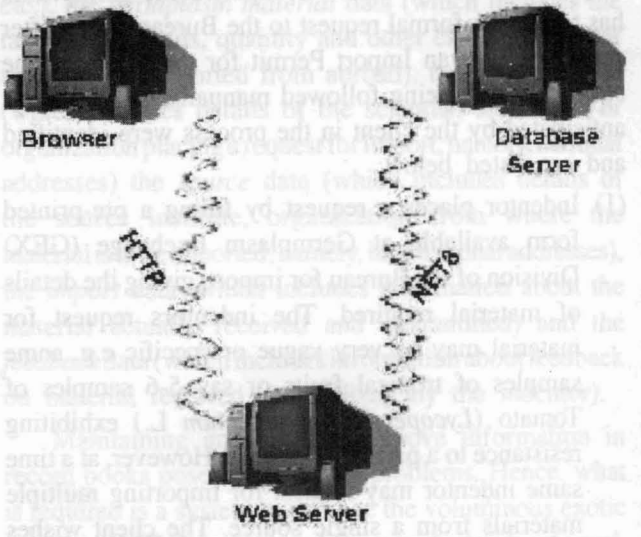

Fig. 1. Architectural design

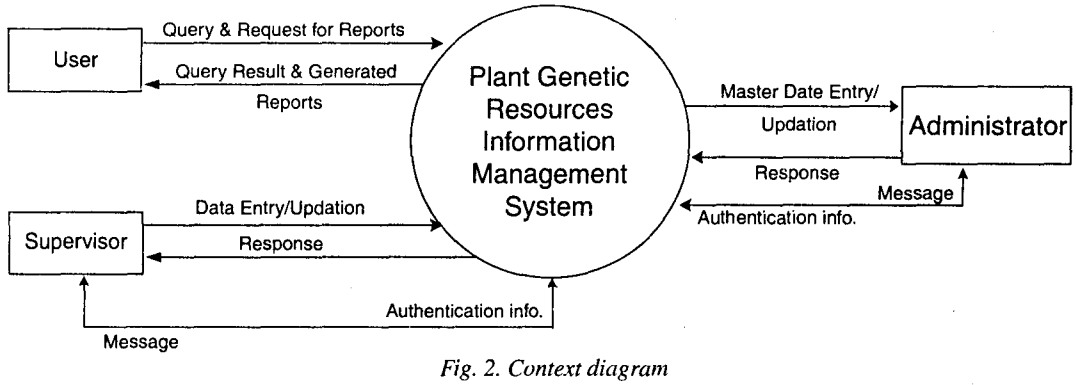

Fig. 2. Context diagram

that are obtained by processing it are sent back to the client.

(b) The Web Server

This is the middle tier between the browser and the Database Server. The user interacts with the Form based GUI through browser and the request that is sent to the Database Server is served by the Web Server. The Web Server waits for the result of the request from the Database Server and sends it back to the browser in suitable form when it is received.

(c) Browser

The browser is the client side interface for interacting with the database server.

Conclusions

The aim of this paper is to design and develop a Management Information System for National Bureau of Plant Genetic Resources to reduce the paper record filing works and for providing the safety for valuable data. The system is able to achieve quick and efficient processing of data, provision for multi-user data update/ entry, generation of quick and well-formatted reports, proper storage of data etc. Thus it enhances the performance in comparision to the existing system. A context diagram (Fig. 2) was drawn to show the process node (process 0) that generalizes the function of the entire system in relationship to external entities. Initially the existing system was thoroughly studied and observations recorded for use in the proposed system. Tools like DFD (Fig. 3) and ER diagram (Fig. 4) were use for requirement analysis. Following that was the design phase in which we designed the database schema for system.

The system is developed using Oracle 9i as RDBMS. The system should be developed in such a way that it becomes user friendly. If desired more enhancements can be made in the system in future to make it more fanciful, convenient and secure. An internet based SQLPLUS named *miSQL*PLUS* has been developed for Bureau which helps the administrators for maintain it online.

Apart from the above insights, we feel that there are other insights that could be gained from this project in the field of software application development. The importance of proper system design and the power of adhering to the correct software engineering development procedures have been made prominent in our mind.

References

Bobrowski S (2000) Oracle 8i for Linux. Tata McGraw Hill. Bradley D Brown, (2002) Oracle 9i Web Development, Oracle Press.

Clements N and G Williams (2000) Java Programming Student Guide. Oracle Corporation Student Guide.

Deitel and Deitel (2000) Java 2 How to Program. Pearson Education.

Groff JR and PN Weinberg (2000) SQL Complete Reference. Tata McGraw Hill.

Hobbs (1998) Database Programming with JDBC. Techmedia.

J L Weber (2000) Using Java2 Platform PHI.

Loney K and G Koch (2002) Oracle 9i The Complete Reference. Oracle Press.

Pekowsky L (2000) Java Server Pages. Pearson Education.

Powell T and F Schneider (2002) Java Script - The Complete Reference. Tata McGraw Hill.

Ramez Elmasri and Navathe B Shamkant (1994) Fundamentals of Database Systems- *Functional Dependencies and Normalization for Relational Databases*, Addison-Wesley Publishing Company.

Schildt H (2002) Complete Reference Java 2.Tata McGraw Hill.

Senn James A, (1989) Analysis and Design of Information Systems, Tata McGraw Hill.

SQL Language Reference Manual (1994) ORACLE RDBMS Version 8.0.

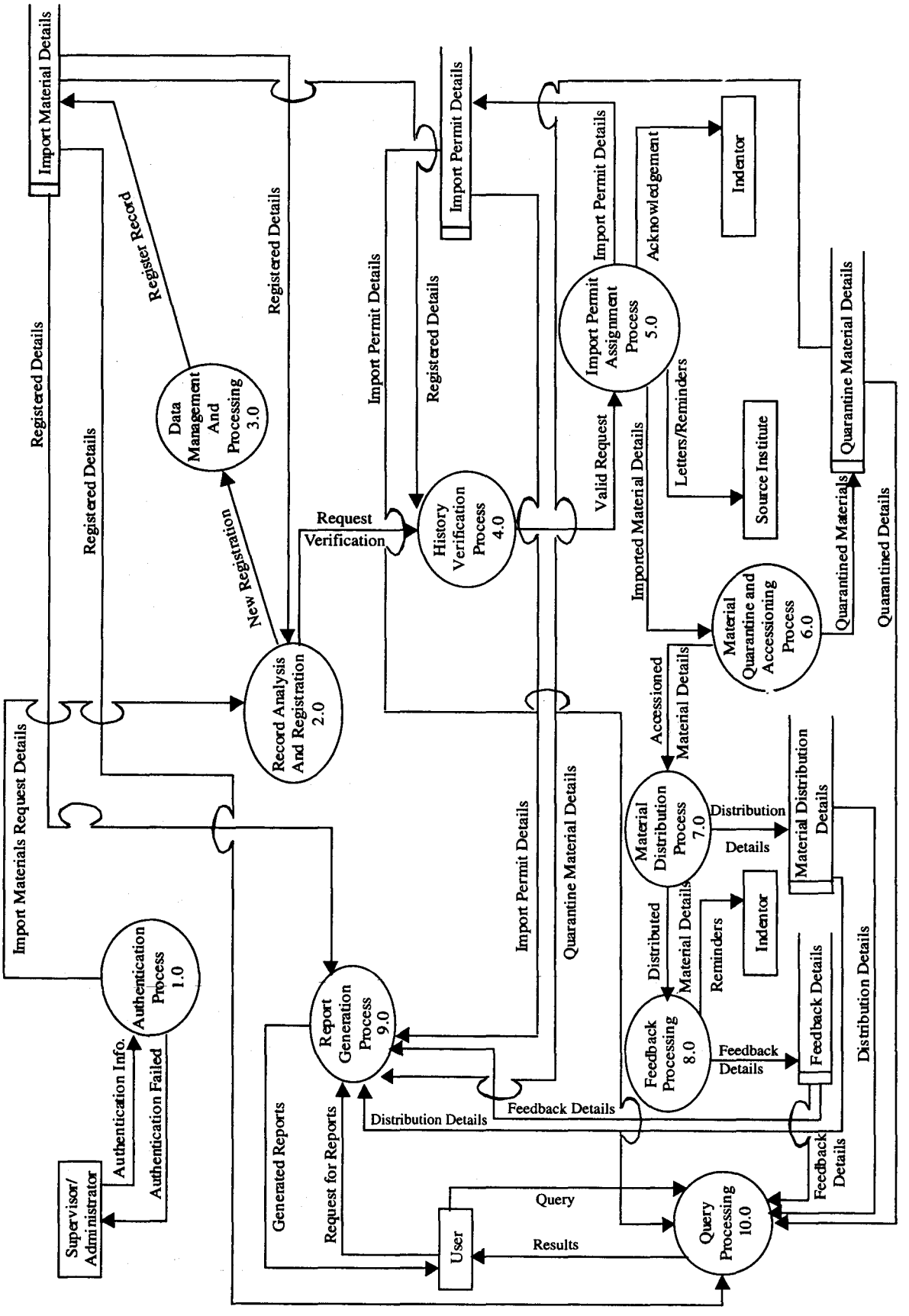

Fig. 3. Data flow diagram (Ist level)

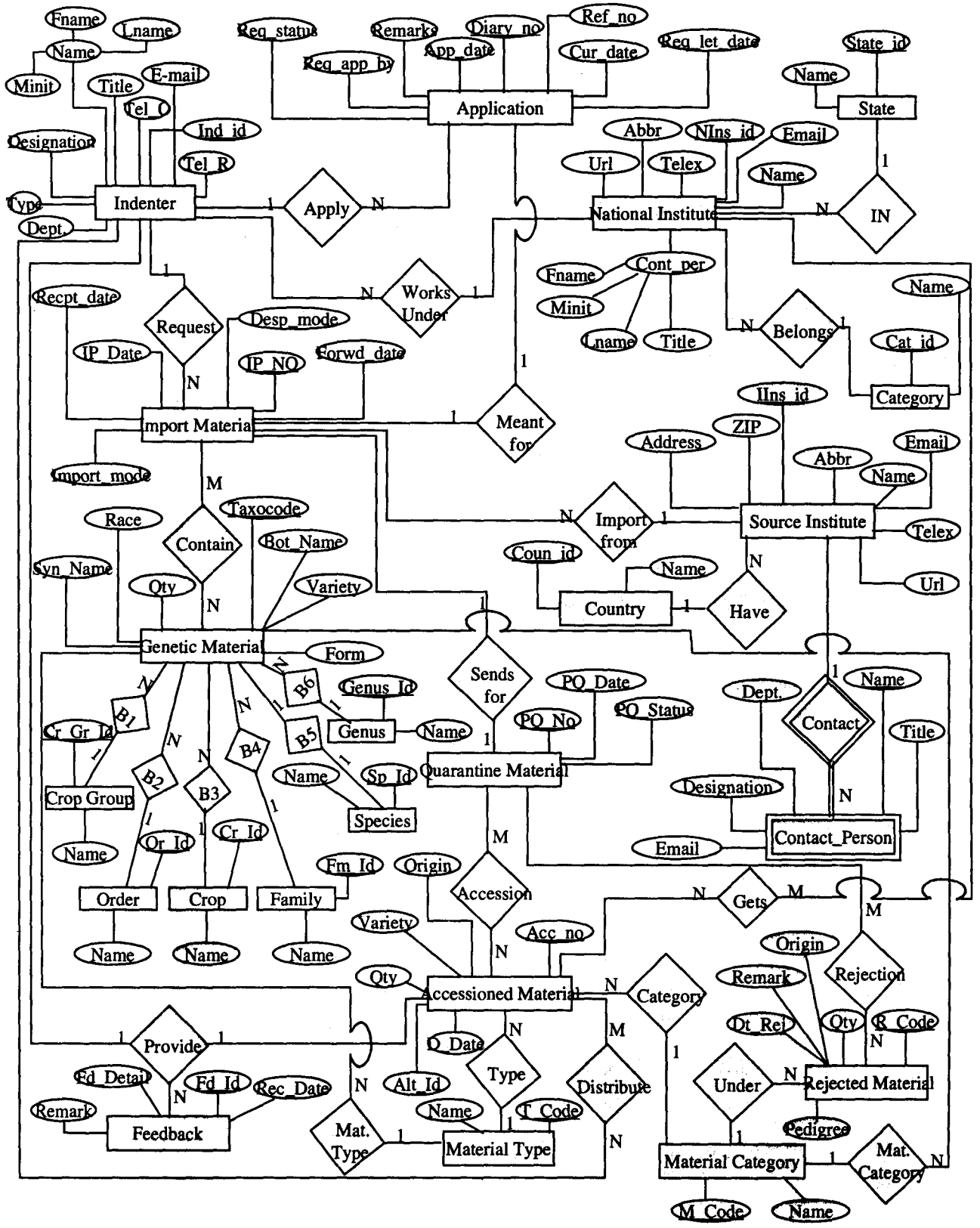

Fig. 4. Entity relationship model