GLCS: A Generic Location Coding System for e-Governance Applications

Susanta Kumar Mohapatra[1]*, Lalatendu Dash[1] and Jagannath Prasad Bakshi[1]

ABSTRACT

Besides human beings, geography or geographic locations act as the most common determinant behind any human oriented initiative. e-Governance being one of the most important human administrative initiative of the present era, naturally relies upon and extensively makes use of the different geographic entities and locations to represent and store the relevant data. This naturally calls for evolving a suitable geographic location codes for use in the e-Governance oriented applications. But with the proliferation of large number of initiatives at the International, National and the local levels, making use of the data pertaining to the same geographic locations, it has become absolutely imperative to devise a suitable location coding system for use by the various applications so as to ensure sharing, exchange and integration of data. Such a location coding system, to be applicable and acceptable universally, should be based on a completely invariant foundation of permanent nature but having sufficient flexibility to not only handle the geographic and administrative variance but also allow for continuous growth and evolvement. This paper provides the outline of such a system which could be adopted by various the e-Governance applications.

Keywords: Location Code, e-Governance, RGI Code, Economic Census Code, PIN Code, Service Oriented Architecture (SOA), Code Server, UDDI, SOAP, WSDL, BPEL4WS, SQL XML, Application Server, Metadata, Standards & Best Practices

1. Introduction

e-Governance initiatives in India has taken a giant stride with the launching of the National e-Governance Plan (NeGP). The plan seeks to create the right governance and institutional mechanism towards ensuring citizen-centricity by adopting suitable standards and best practices. Under NeGP, a number of mission mode projects are being implemented throughout the country. Moreover, a number of independent initiatives are also underway in various states. In this context, it is quite pertinent to address the standards and methods to be adopted in order to exchange and share the data among the various e-Governance applications. The majority of data used in such applications are linked to various location units, corresponding to geographic or administrative domains. This arises from the fact of location being an important entity for data representation. This being the case, coding of locations are deemed necessary for efficient storage and retrieval of data. But presently such applications are adopting separate and non-standard location coding systems to represent and store their data. In this context, it is deemed appropriate

[1] National Informatics Centre, Orissa State Unit, Bhubaneswar, Orissa, India
* *Corresponding Author:* (Email: skm@ori.nic.in, Telephone: 91-674-2501339)

to take stock of the existing location codes, their short comings (if any) and their relevance in the current emerging scenario of proliferation of e-Governance applications and the need for location code standardization for ensuring data sharing, exchange and application interoperability.

2. Existing Location Coding Systems: A Critical Study

There are a number of initiatives in India which have devised location coding systems. This study takes a critical look at some of the important systems such as (i) RGI Code (Census of India), (ii) Economic Census Code (Ministry of Statistics & Programme Implementation, Govt. of India) and (iii) PIN Code (Dept. of Post, Govt. of India).

***2.1 RGI Code*:** The decennial Census of India is the largest single administrative exercise of its kind in the World. Even the preparations for it make news because of the size and impact of the operation.

In 1991, the data was made available as several database (.DBF) files. These used a 16 digit location code (eg. 1400303001010002) for urban data and 18 digit code for rural (eg. 140030003000300005) data at the State level. The first two digits represented the district. The next four digits represented the taluk/ tahsil in case of urban areas. In case of rural areas, six digits were used to represent the development blocks. The next six digits stood for the town/village panchayat. The last four digits were for representing the wards.

In the 2001 Census, the coding is slightly different from 1991. The State/Union Territory is represented by two digits each in the code. In the State level data, the first two digits denote the district. The next four represent taluk, tahsil, P. S., development block, circle or mandal as is relevant to each State. The towns as well as the villages are represented separately through Permanent Location Code Numbers (PLCN) consisting of eight digits for villages and four for wards. Beginning with the first village of the first district to the last village in the last district, there is a continuous running number code for each village. The eight digits provided to represent the village will help addition of new villages in future without disturbing the overall scheme.

Table 1: Location Code Structure used in 2001 Census

State/ UT	Two digits (within the country)
District	Two digits (within the State/UT)
Subdistrict	Four digits (within the district)
Village	Eight digits (within the State/ UT)
Town	Eight digits (within the district)
Ward	Four digits (within the town)

The general pattern followed in coding the geographical units was a serpentine one, beginning from the northwest corner and completing at the farthest southeast corner within the defined higher level of the geographical hierarchy. In so far as the state/union territory is concerned Jammu & Kashmir has code number 01 and code number 35 has been allotted to the union territory of Andaman & Nicobar Islands. Generally, the administrative jurisdiction below a district is a sub-district, for example, a tehsil or a taluk. However, in certain states and union territories exceptions had to be made since the nomenclature for sub-districts is not uniform throughout the country. For example, in Orissa Police station is the sub-district, in West Bengal it is the Community Development Block and so on. In Assam, Mizoram and Orissa there are a few sub-districts which cut across the district jurisdictions: In such situations the sub-districts have been provided a separate series of code numbers viz. 9000, 9001 and so on.

One of the major initiatives taken in the Census 2001 was the allotment of Permanent Location Code Number to each and every village within the State and not within a tahsil as in the earlier censuses. PLCN

was thus assigned as one continuous number from the first village in the first district to the last village in the last district. PLCN is an eight digit unique location code number with the first six digits representing the code number of the village and the last two digits depicting two zeros '00'. These zeros are reserved as buffer to be used for coding any new village(s) that may come up between two villages in future. For example, if a new village comes up between two villages with PLCNs 01254600 and 01254700, the new village will be allotted PLCN 1254601 and so on.

The location code number for a town is also an eight digit number starting with the digit 4 situated at the extreme left acting as the unique identifier. The next two digits depict the code number of the district in which the town falls followed by two digits representing the town serial number in the district. There are three zeros at the end as buffer, mainly to meet the requirement for bringing the number of digits to eight to match the number of digits in the PLCN for the villages. Thus a town location code number 40305000 represents the town serial number 5 of the district number 03 in a State. The out-growths of towns have not been given any independent location code numbers. These have been given notional ward codes after the last ward code of the respective towns.

2.2 Economic Census Code: Central Statistical Organisation undertook countrywide Economic Census, fifth in the series, during the year 2005-06 in all the States/UTs in collaboration with State/UT Directorates of Economics & Statistics (DESs). The Census covers all the entrepreneurial activities of the country (except crop production and plantation). The details of the location coding for rural and urban areas are given in Table 2 and Table 3.

Table 2: Location Code for Rural Areas

1-2	State Code
3-4	District Code
5-8	Tehsil/Taluk/PS/Dev. Block/Circle/ Panchayat Code
9-16	Village code
17-20	Enumeration Block No.

Table 3: Location Code for Urban Areas

1-2	State Code
3-4	District Code
5-6	Town Code
7-8	Frame Code (Year of updation of UFS)
9-13	Ward No.
14-18	Investigator Unit (IV) Number (within town)
19-20	(UFS) block number

2.3 PIN Code: A Postal Index Number or PIN or Pincode is the post office numbering or post code system used by the Indian Postal Service. PIN stands for Postal Index Number and the code is 6 digits long. The PIN was introduced on August 15, 1972. The first digit for regions (8 regions in India), the next digit is for sub-regions, next for the district, and the last three for the post office.

2.4 Analysis of the Existing Coding Systems

Table 4: Analysis of the Existing Coding Systems

Parameter	RGI Code	EC Code	PIN Code
Coverage	Large but not complete	Similar to RGI	Very small
Coding Type	Intelligent, excluding	Intelligent	Intelligent

	PLCN		
Coding Pattern	Composite	Composite	Composite
Simplicity	No	No	Yes, due to small set of locations
Maintainability	Less, barring PLCN	Less	Less
Extensibility (New Locations)	Not covered	Not covered	Not covered
Generic (can handle different project needs)	No	No	No
Online availability	No	No	No
Data Exchange	Not provided	Not provided	Not provided
Interoperability	Not provided	Not provided	Not provided

As seen from the RGI coding system, there is a trend to go away from composite intelligent coding system to a permanent non-intelligent, non-composite coding system as already initiated for the villages in the form of PLCN. This has been done keeping in mind the maintainability as any code change or addition of new key column in a composite coding system generates a cascade effect down through foreign key relationship. Further more, intelligent keys denote the organizational or application domain specific meaning associated with them, and with any change in the organization/ business structure, the codes also needs to be changed. This explains the trend towards adoption of "Permanent Location Codes" which are non-intelligent and more flexible by nature. Thus the latest trend is from composite intelligent coding system to a permanent, non-intelligent, non-composite coding system.

As the primary purpose of the various codes is to provide an efficient and interlinked data processing mechanism for the computer software/RDBMS, immediate association of the meaning with the codes is not important as it could be obtained from associated link tables. As far as the need to provide meaningful information to the public, computer software could do the same in a user friendly and context sensitive way. Thus the trend initiated by RGI needs to carried on to it's logical end. In addition, the other important parameters such as maintainability, extensibility, genericity, data exchange, interoperability, online availability etc. needs to be addressed. Keeping in mind the above mentioned needs, it becomes imperative to devise a suitable location coding system which does away with the limitations of the existing systems. Such a system would pave the way for an adaptive data infrastructure based on an enterprise data architecture model. This paper illustrates the attempt made in this regard called GLCS, a generic location coding system.

3. GLCS

3.1 *Problem Statement and Stakeholder Descriptions*

Problem Statement:

Table 5: Problem Statement

For	EGovernance application developers, Govt. departments & Standardization bodies
Who	Are the creators or users of location codes
The product	Is an application software
The Problem	Lack of a standardized accessible up-to-date location codes
Affects	The large community of Govt. & Pvt eGov application developers & the concerned Govt. departments entrusted with the responsibility of maintenance of the up-to-date location codes
Impact	Redundant efforts, Duplication of location data, data inconsistency, disputes due to multiple non-matching data sets and lack of standardization and non-robust & bulky applications; lack of data exchange & integration between applications
Solution would be	Centralised web-enabled location code server maintained by concerned departments and accessible and usable by the host of application development community on-line over internet/intranet/VPN or offline

Stakeholders & Users Description:
The identification of various stake holders, their roles & responsibility along with an unambiguous well-defined user needs do play an important role while designing & developing any user-friendly system. Thus a small sketch of the various stake holders & the users of this system is given below;

Table 6: Stakeholders & Users:

Name	Roles	Responsibilities
Government	Provider and User of Data & Application	Supports Standardization & hosting
Application Developers/Academicians/Public	Consumer of the service	Enrich by pertinent feedback
NIC	Design & development	Maintenance

3.2 Objective

- To serve as a suitable Location Code Standard
- The proposed GLCS, is based on an invariant framework addressing the variations which might arise in different geographic, political or administrative domains from time to time.

3.3 Architecture

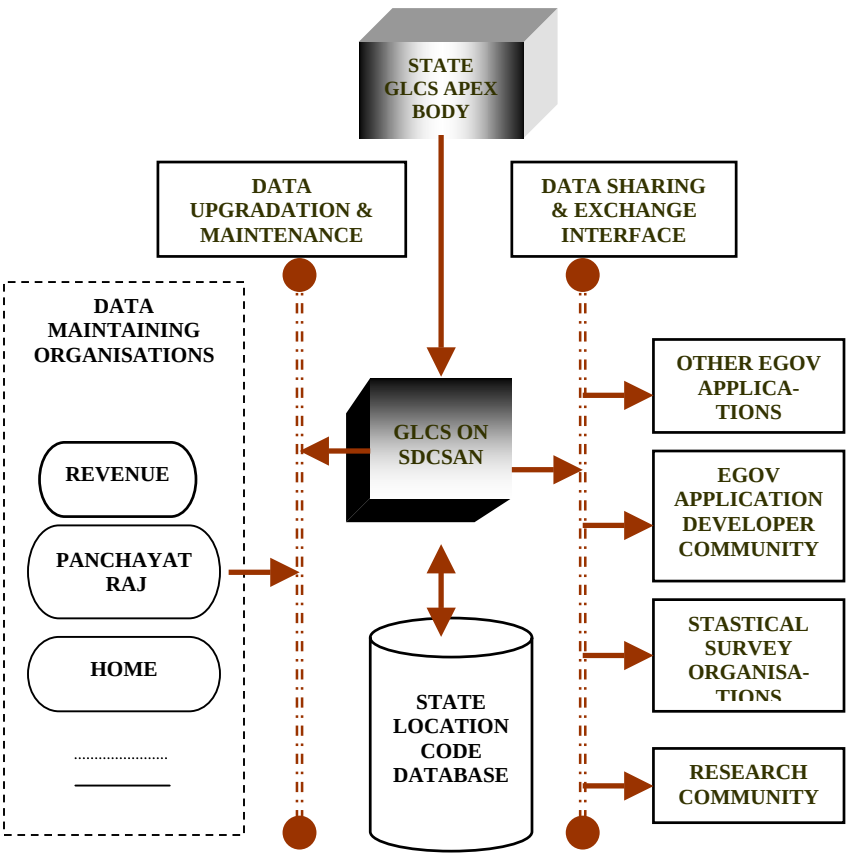

Figure 1 : GLCS Architecture

*3.4 **System Design***

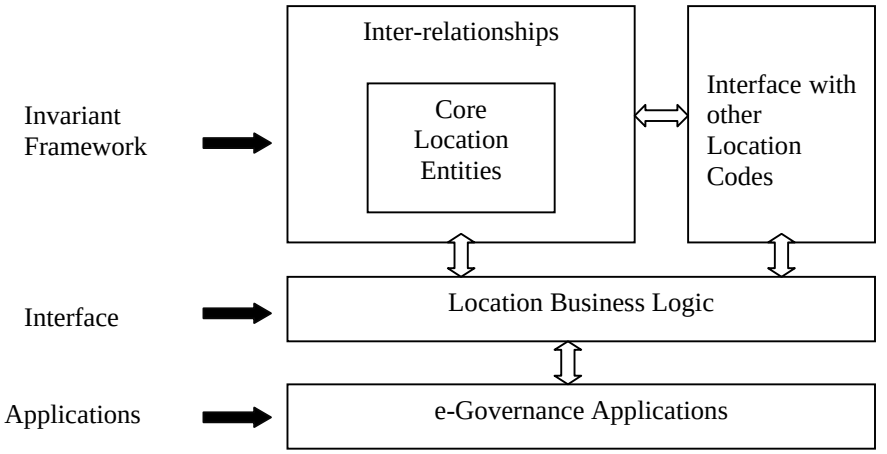

Figure 2: GLCS System Design

The proposed generic location system, henceforth referred to as GLCS, is based on an invariant design framework completely independent of any spacio-temporal variations which might arise in different geographic, political or administrative domains from time to time.

The GLCS Framework consists of:

- Core Location Entities
- Entity Inter-Relationships
- Mapping with other Location Codes

Accessibility to the framework by e-Gov Applications provided through Location Specific Business Logic implemented through web services.

The various data items to be kept under the three invariant components are;

- Core location entities
  - Location Language Codes
  - Location Units
  - Location Codes
  - Location Domains
  - Locational Domain Migration Paths

- Locational Inter relationships
  - Location code Location unit Link
  - Location code Location code Link

- Interface with other location codes
  - Metadata of mapping table
  - External Coding Agency Code
  - Code Mapping Table (for linking RGI, ECI etc codes)

3.5 Interoperability with Other Systems

GLCS makes use of Web services as one of the major technology tool to achieve various data related service providing capabilities. A Web service is a specific kind of service that is identified by a URI, whose service description and transport utilize open Internet standards. Interactions between Web services typically occur as SOAP calls carrying XML data content. Interface descriptions of the Web services are expressed using Web Services Definition Language (WSDL). The Universal Description, Discovery, and Integration (UDDI) standard defines a protocol for directory services that contain Web service descriptions. UDDI enables Web service clients to locate candidate services and discover their details. Service clients and service providers utilize these standards to perform SOA's basic operations shown in the figure. Service aggregators may use the Business Process Execution Language for Web Services (BPEL4WS) to create new Web services by defining corresponding compositions of the interfaces and internal processes of existing services. GLCS activates interoperability with other e-Governance applications through web services, metadata and directory service. Data sharing is enabled through data mapping to other standard location codes. Online data access by other applications is to be provided by distributed database architecture based in SQL XML.

3.6 Advantages

There are host of attributes of the GLCS which are given as follows;

- Flexibility: Facilitates easy incorporation of changes
- Simplicity: Non-composite standardized coding, less no. of tables
- Extensibility: New location entities could be added/ deleted without structural change
- Maintainable: No structural information in code, thus independent and maintenable
- Comprehensive: Has provision for accommodating all conceivable locations
- Customizable: As it can be mapped to any other new location scheme
- Application Friendly: Through SOA using WSDL, UDDI & BPEL4WS
- Suitability in Adaptive Enterprise Data Architecture: Enterprise network using application & data server

4. Implementation

The implementation strategy for GLCS must start with the formation of an apex committee at the international/national/state levels. In addition to work out the broad policies and guidelines, the apex committee must work out all the relevant locational units relevant to various e-Governance oriented applications. They must also entrust the creation, maintenance of the location codes and finalization of location business rules pertaining to the different locational units to specified organization/ department. Those organizations/departments will work as the sole custodian for the assigned location item and corresponding location codes. An initial prototype of GLCS has already been developed and put to test run by National Informatics Centre, Orisssa State Unit.

Learning from our experience so far, some of the pertinent issues for successful implementation of GLCS are listed as follows;

- Political will and need of a administrative champion: In Orissa, this has been one of the most important bottleneck for most of the e-Governance initiatives
- Sense of ownership by Government: Ownership is found to rest with the consultants or lower administrative authorities by default
- Constraint of funds: Lack of funds with the state govt. and inability to obtain from centre
- Lack of skilled and capable private organizations for supporting e-governance initiatives
- Illiteracy, poverty, and digital divide: Less than 1% of total population are computer literate

- Process re-engineering and change management: Mostly ignored due to lack of knowledge, difficulty in undertaking administrative reforms and legal issues

Presently the GLCS provides services such as; (i) Location Code Information, (ii) Downloading Location Code Tables, (iii) Downloading Location Code Query Outputs, (iv) Updation of Location Code Data, and (v) Getting the list of Services Provided by GLCS

- Availing GLCS tables as web based distributed data on industry standard RDBMS (ORACLE, MS SQL SERVER)
- Implementing GLCS on a Service Oriented Architecture (SOA) for ensuring availability, flexibility & manageability.
- Starting the implementation strategy for GLCS with the formation of an apex committee at the national & state levels.
- Working out all the location units relevant to various e-governance oriented applications by the apex committee.
- Entrusting the creation, maintenance of the location codes and finalization of location business rules pertaining to the different location units to the specified organization/department by the apex committee.

5. Concluding Remarks

GLCS is quite relevant to the currently emerging e-Governance scenario, where standards and best practices are playing a crucial role. Department of Information technology, Government of India has embarked on a major plan in this regard. But location coding, though quite important, has not yet been extensively dealt with. Numerous e-Governance applications should no longer take resource to isolated non-standard location coding system so as to avoid potential future pitfalls in terms of lack of access and sharing of location based data. Initiatives such as GLCS should be offered the due support by all concerned so as to make this initiative a success.

References

1. RGI Coding System, Registrar General of India 2001 *Census Handbook*
2. Concept Paper on Permanent Location Masters for Data Exchange in e-Governance projects, NIC Karnataka State Unit. Available at http://egovstandards.gov.in. Accessed July 15, 2006.
3. http://mospi.nic.in/economic_census.htm Accessed June 1, 2007.
4. http://www.indianpost.gov.in/ Netscape/Pincode.htm Accessed July 15, 2006.

About Authors

Susanta Kumar Mohapatra has done his B.Tech in Mechanical Engineering from I.I.T, Kharagpur in 1985 and M.Tech in Computer & Information Sciences from Hyderabad Central University in 1987. Since 1988 he has been working at Natioanl Informatics Centre, Orissa State Unit at Bhubaneswar. He is currently working as a Technical Director and is in charge of a number of initiatives such as the Land records computerisation project and the e-governance initiatives such as e-Sahayata and Janavani in Orissa. He may be contacted at skm@ori.nic.in and his phone No is 91-674-2501339.

Lalatendu Dash has done his M.Sc & Diploma in Computer Science in 1987 & 1989 from Utkal University & College of Engg. & Technology respectively. He has done his MCA from IGNOU in 2000. Since 1989 he has been working at Natioanl Informatics Centre, Orissa State Unit at Bhubaneswar. He is looking after the ICT initiaves in Natural Calamity & Public Health sectors. He is also involved in the eSahayata Poject.

Jagannath Prasad Bakshi has done his Msc in Physics from Utkal University and subsequently done MBA from the same University. Joined NIC in 1997 and since 2001 he has been working at National Informatics Centre, Orissa State Unit at Bhubaneswar. He is currently working as a System Analyst and is involved in projects such as the Land records computerisation project and the e-Governance initiatives such as e-Sahayata and Janavani in Orissa.