# Evolutionary Rapid Development using open source framework for Geospatial Data Processing[†]

**Indranil Misra**[*]
**S.Manthira Moorthi**[**]
**Rajdeep K Gambhir**[***]
**R. Ramakrishnan**[****]

## Abstract

***Purpose:*** *In the digital age of 21st century, open source movement is considered to be a new paradigm for software use and development. Evolutionary Rapid Development (ERD) is the concept of integrating software systems based on the reuse of components and by the use of software and architectural templates. The objective of the study is to use ERD fundamentals for exploratory analysis and development of software features, infrastructure and components with the adoption of leading edge open source geospatial technologies for doing geospatial data processing tasks.*

***Methodology:*** *In the study many open source geospatial tools are evaluated, adopted and integrated using ERD in GIS framework.*

***Findings:*** *The paper provides an architectural framework for geospatial data processing using evolutionary rapid development technology (ERD).*

***Practical Implications:*** *This paper highlights the challenges faced by the geospatial software developers to glue systems together and their resolution by using ERD as a technical model.*

**Keywords:** *Evolutionary Rapid Development (ERD); Geographic Information Systems (GIS); Open Source Softwares; Geo Processing; Classification*

**Paper Type:** *Technical*

## Introduction

Geospatial technology space in today's scenario is able to integrate hardware, software and data seamlessly for storing, processing, analyzing and displaying all forms of geographically referenced information. Here software journey for geospatial data processing tool development is considered to be a challenging as well as an interesting job for software developers. The geospatial software development in the present era doesn't mean *"start from scratch"*. The

---

[†] *Modified version of the paper presented at National Seminar on Open Source Software Systems: Challenges and Opportunities held at University of Kashmir, June 20-22, 2011.*

[*] *Indranil Mishra. Data Products Software Group.Signal and Image Processing Area. Space Applications Centre (ISRO). Ahmedabad 380015, India indranil@sac.isro.gov.in*

[**] *S.Manthira Moorthi. Data Products Software Group. Signal and Image Processing Area.Space Applications Centre (ISRO). Ahmedabad 380015, India smmoorthi@sac.isro.gov.in*

[***] *Rajdeep K. Gambhir. Data Products Software Group. Signal and Image Processing Area. Space Applications Centre (ISRO). Ahmedabad 380015, India rajdeep@sac.isro.gov.in*

[****] *R.Ramakrishnan. Data Products Software Group.Signal and Image Processing Area. Space Applications Centre (ISRO). Ahmedabad 380015, India rama@sac.isro.gov.in*

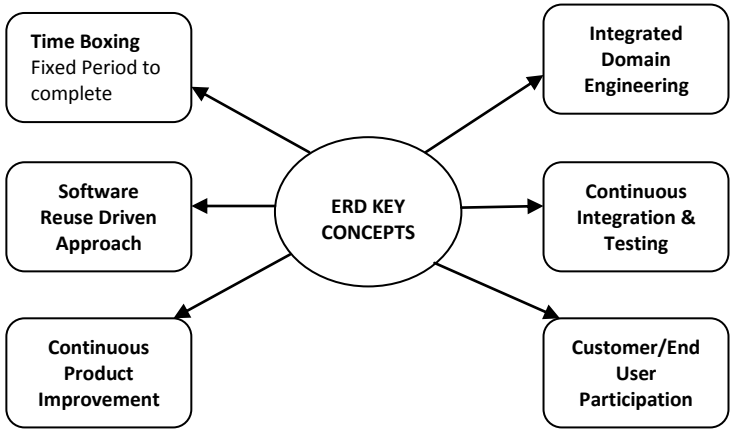
scope is to build new software systems for geospatial data analysis by integrating and customizing various open source geospatial softwares using *Evolutionary Rapid Development* model. The software productivity consortium as a technology development agent of the U.S Defense Advanced Research Projects Agency (DARPA) developed *Evolutionary Rapid Development (ERD)* process for managing the development of complex software systems using software reuse in an environment of rapidly evolving components and architecture.

This paper illustrates an *ERD* approach for building a desktop geospatial toolbox by amalgamating available open source tools which can perform the required image processing and geospatial operation tasks in an effective way. In this paper, details are given on ERD principal concepts and its adoption in the geospatial software domain to glue software systems together for the use by geospatial analysts.

**Principle of ERD**

Information system development is undergoing a revolution. Software developers no longer focus on handcrafting an application; rather they assemble robust, fully functional software systems from readily available and inexpensive components. Here system developer task has begun to transform the system into an optimal solution in accordance with an application framework. In this *"Information Age",* systems must be designed to grow and adapt to future events that are not always certain. The ERD process establishes a set of key characteristics that together address these factors **(Fig.1) (De- Santis, Blyskal, Moini, & Tappan, 1997).**

**Fig. 1: ERD Concepts**



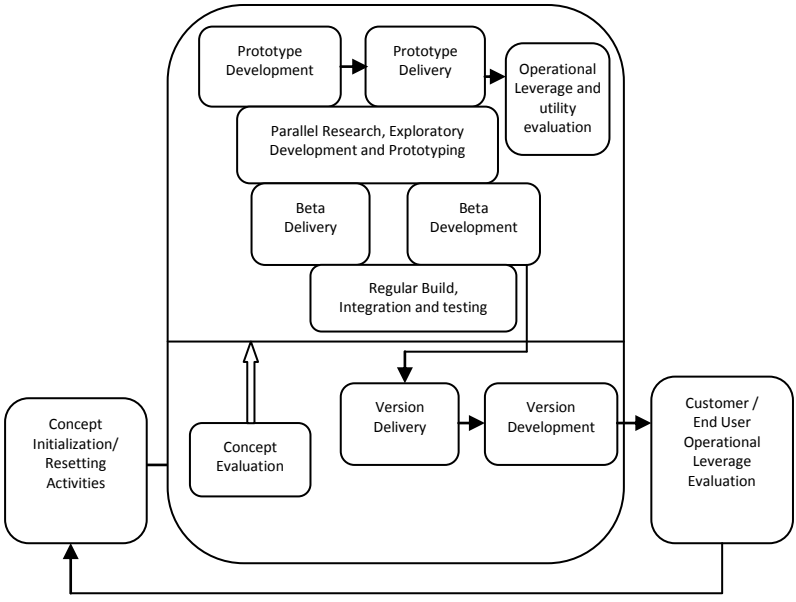
The ERD process description is shown in **Fig. 2**. It mainly consists of three main parts. The parts are as follows:

**1. *Concept Initialization and Resetting Activities*:** It establishes the higher level objectives and constraints for the software system with the customer consultation.

**2. *Evolutionary Development Time Box Activities*:** It refines and implements the software system through a series of prototypes.

**3. *End User Operational Leverage Evaluation Activities*:** It assesses the operational system to determine its value and to identify additional needs.

**Fig. 2: ERD Process Description**



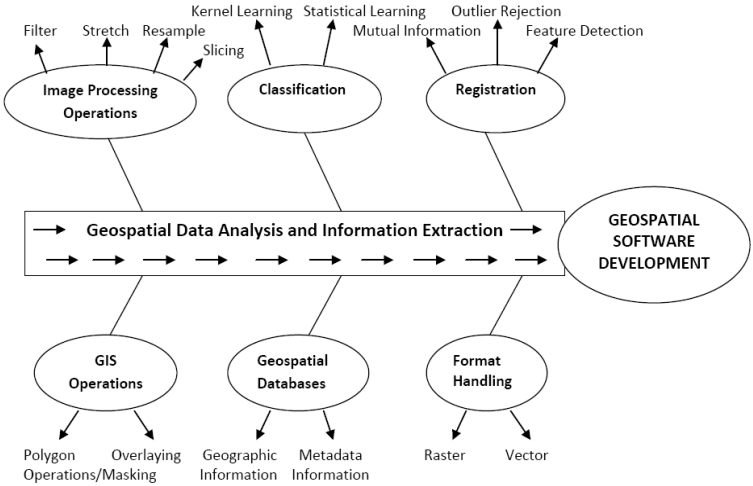

## I. Fish Bone Analysis for Geospatial Problem

The fish bone diagram **(Fig. 3)** shows the causes of a certain event/problem and is in industry practice. The common use of fish bone diagram is to build an effective product design. Here fish bone diagram is used to simulate the geospatial software development problem. The *"causes"* are the functionalities required for geospatial information extraction.

**Fig 3** shows the major functionalities required for geospatial software development. It is discussed in later sections that no single software can do all these jobs perfectly for geospatial analyst. We aimed for parallel research and exploratory development using ERD process model for

building such complex geospatial framework. Many open source geospatial tools can resolve the causes of the fish bone diagram in **Fig 3** and attempt to build, customize and integrate them to resolve the geospatial problem.

**Fig. 3: Fish Bone Diagram**




## II. Free and Open Source Software as an Optimal Solution

Open source software has permeated all the areas of software computing and geospatial analysis is not an exception **(Steiniger & Bocher , 2008)**. The processes/functionalities required for building geospatial software fall over a large range i.e from *geospatial data preprocessing* to *information extraction*. Many open source geospatial tools are available to solve the causes of fish bone diagram in **Fig. 3.** But it has been seen that no single open source software supports all the software evaluation parameters to solve the problems of geospatial software development. Also we need to see the reusability of the source code for building the customized application. **Table I** shows the open source tools available against the functionalities/problems of fish bone diagram.

**Table 1: Open Source Tools and their functionalities**

| Functionalities/Problems | Tools Available | | |
|---|---|---|---|
| Format Handling | GDAL -Geospatial Data Abstraction Library | | |
| Registration | OSSIM | ELASTIX | ILWIS |
| Classification | ILWIS | SAGA | |
| Image Processing Operation | ILWIS | OSSIM | SAGA |
| GIS Operations | ILWIS | OSSIM | |
| Geospatial Databases | PostgresSQL | | |

**1. GDAL: An Abstraction Library**

GDAL stands for *Geospatial Data Abstraction Library*. It is an open source utility for handling raster/vector formats. GDAL is written in ANSI C and C++. It can be compiled with all modern C/C++ compilers. GDAL run on all modern operating systems such as Windows XP, Linux and Unix. GDAL supports all standard raster formats. GDAL also provides set of pre built utilties. These are individual executables run from the command prompt. One of the command line utilty is gdal_translate which can be used to translate one format into another and vice-versa.

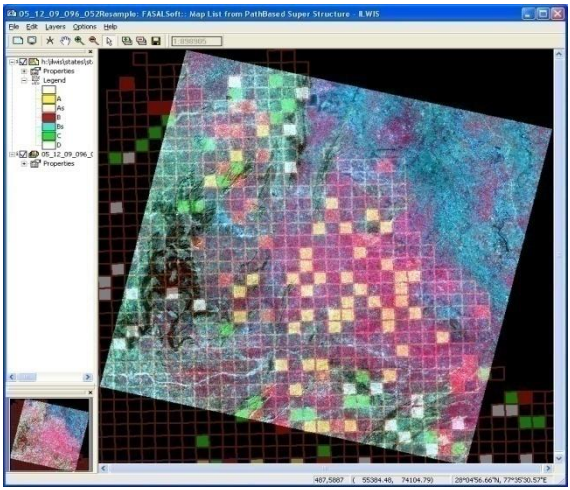
**2. ILWIS: An Open Source Geospatial Toolbox**

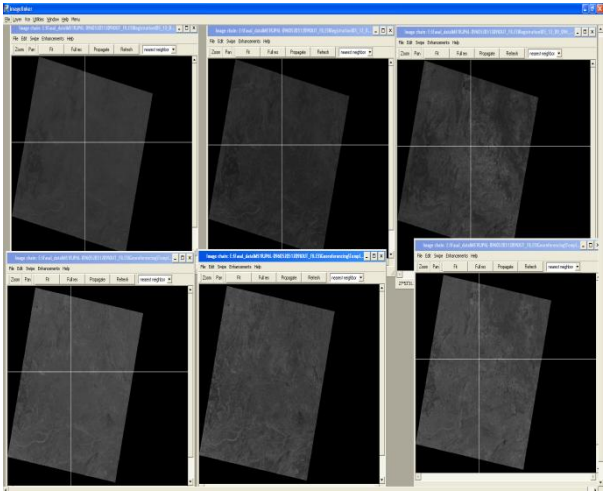
The *Integrated Land and Water Information System (ILWIS)* is PC-based GIS & Remote Sensing software, developed by ITC up to release 3.3 in 2005. ILWIS comprises a complete package of image processing, spatial analysis and digital mapping (as of July 1st 2007), ILWIS software is freely available ('as-is' and free of charge) as open source software (binaries and source code) under the 52° North initiative (GPL license). This software version is called *ILWIS Open* **(Lemmens & Schouwenburg, 2007)**. ILWIS is almost a complete tool for image processing and GIS operations but only for a data type of 8 bit unsigned char images is supported for image classification. The toolbox only provide georeferencing using tie points and does not support automatic image registration framework which is a pre-requisite for today's geospatial software **(Misra, Moorthi, Darji & Ramakrishnan, 2010)**.

**3. OSSIM: Awesome Image Processing**

*Open Source Software Image Map (OSSIM)* is a high performance engine for remote sensing, image processing, geographical information systems and photogrammetry. It has been actively developed since 1996 **(Misra, Moorthi, Darji & Ramakrishnan, 2010)**. Designed as a series of high performance software libraries, it is written in C++ employing the latest techniques in object-oriented software design. It has a powerful automatic satellite image registration plugin as an extension to OSSIM library. Its API documentation is exemplary. It is relatively easy to build an image processing application or a plugin using OSSIM libraries which does not demand much of GIS operations and attribute based calculations of rasters, vectors and table data. Though vector and shape file support is present, its GIS capabilities are limited than other open source toolbox.

**Fig. 4**: **Resourcesat-1 LISS-3 data processing in ILWIS and OSSIM toolbox**





**4. Elastix: Powerful Registration Toolkit**

*Elastix* is developed, as a publicly available computer program for intensity-based medical image registration. It is an open source software, based on the well-known *Insight Segmentation and Registration Toolkit (ITK)*. The software consists of a collection of algorithms that can be used to solve image registration problems in geospatial domain. The modular design of Elastix allows the user to quickly configure, test, and compare different registration methods for a specific application. The command-line interface enables automated processing of large numbers of

datasets, by means of scripting **(Klein, Staring, Murphy, Viergever, & Pluim, 2010)**.

**5. SAGA GIS**

*SAGA GIS (System for Automated Geoscientific Analyses)* is a free and open source geographic information system used for editing spatial data. It was originally developed by a small team at the Department of Physical Geography, University of Gottingen, Germany, and is now being maintained and extended by an international developer community. SAGA has a fast growing set of geoscientifc methods, bundled in exchangeable Module Libraries. The standard modules contain variety of image classification methods. It is available in Windows and Linux.

**6. PostgreSQL: An open source Database Management System**

PostgreSQL is a powerful, open source object-relational database system. It has more than 15 years of active development and a proven architecture that has earned it a strong reputation for reliability, data integrity, and correctness. It runs on all major operating systems, including Linux and Windows. It is fully **ACID** (atomicity, consistency, isolation, durability) compliant, has full support for foreign keys, joins, views, triggers and stored procedures (in multiple languages). *PostGIS* is a spatial database extension for PostgreSQL. PostGIS adds support for geographic objects to the PostgreSQL object-relational database. In effect, PostGIS *"spatially enables"* the PostgreSQL server, allowing it to be used as a backend spatial database for geographic information systems **(The PostgreSQL Global Development Group, 2010)**.

## III. ERD with Available Open Source Geospatial Tools

The need of the hour is to use the concepts of ERD discussed to integrate and build customized geospatial software according to the customer requirement. In ERD, concept initializing step is an important step where the number of processes, the flow and the data interfaces get defined in consultation with the customer. Further, the concept may undergo changes which is called resetting the concepts as the development happens in a iterative mode. While we explore functionalities of the packages, the initial definitions change according to the functionalities and interfaces available.

At present multiple softwares in combination are employed to solve the problem at hand. But the data interface between packages had to be ensured by providing relevant metadata. Usually the metadata values are common but the keywords syntax and semantics are different for different packages. The data interface not only at entry point for any process workflow but also at different stages has to be ensured.

**1. Evaluation of Open Source Packages**

There are significant opportunities to support the process of comparing, reasoning about, integrating, and modifying *free and open source softwares (FOSS)*. Apart from technical functionality requirements, software aspects such as programming languages mix, easy source compilation, third party library dependencies, GUI library for cross platform compatibility, data models, process chain design, and customizability are definite parameters to consider open source softwares for developers with specific skill set. For the present software exercise we preferred single or few library dependencies for each geospatial processing task in the windows operating system platform **(Barik, Samaddar & Gupta, 2009)**

**2. Rapid Prototyping Versus Long Term Tool Building**

In rapid prototyping, the development simply wants to amalgamate functionalities in a non cohesive manner. This means accepting tools instead of just functionalities and accommodates them in tools management architecture. Rapid prototyping approach can also be the approach taken by not so experienced software developers or by an effort not well resourced and there is a pressing need for prototyping of potential new functionality. The hands-on participation of developers and quick rollout time of these applications enable an early demonstration of feasibility by avoiding the traditional process of software development. If the functionality becomes popular, it can then become part of an existing or new software system that meets all quality requirements.

**3. Combining Geospatial Tools**

The open source geospatial softwares encountered have their own advantages and disadvantages. But still they provide a good chance for a pragmatic programmer to start customizing open source software for doing the geospatial tasks as per the custom requirements. *FOSS ILWIS* is a good choice for all GIS tasks with raster and vector data layers, visualization and presentation but *OSSIM library* scores better on geospatial tasks especially on remote sensing images with its rich plugin tools on image registration and mosaicing. *ILWIS'* supervised classification module is not satisfying the custom requirement in this study because there's no functionality to accept the training sample data in the form of *ESRI* shape file before image classification. *ILWIS* does not have the capability to classify images other than 8 bit images. The need for specific image classification tools, pushed the team to develop a new tool for the purpose **(Moorthi, Misra, Ramakrishnan & Srivastava, 2011)**.

## Conclusion

*ERD* is a powerful software development process model. It requires a software development team to develop functionalities by aggregating functionalities from free and open source software (FOSS) using evolutionary rapid prototyping development. This kind of functional aggregation from more than one tool and manage it in a specific software environment depend upon the organizational culture and practice. Users for softwares of this kind need to be educated properly and convinced that it is really a solution that they are looking for without any compromise. The functionality set, design and user interface of open source software mainly depend upon the objectives set in the concept initialization and resetting phase. A desktop geospatial tool box comprising of FOSS tools *ILWIS, OSSIM, SAGA, ELASTIX, GDAL* and *POSTGRESQL* was developed, complete in functionalities to satisfy a geospatial analyst need for their case studies. *ERD* using open source softwares is best suited for organisations where software needs are met with in house developments. We find ERD has potential to be a formal approach in open source software developments and hence this concept is employed in arriving to solve the geospatial data processing tasks.

**Acknowledgment**

The authors thankfully acknowledge Dr.R.R.Navalgund, Director, Space Applications Centre, for giving us opportunity to explore open source for geospatial data analysis and Dr. P.K Srivastava, Deputy Director, SIPA, Space Applications Centre, for his encouragement and support.

## References

Steiniger, S., & Bocher. E. (2008). An Overview on Current Free and Open Source Desktop GIS Developments. *International Journal of Geographical Information Science.* 23 (10), 1345-1370. doi:10.1080/1365881080802634956

De- Santis, Richard. M., Blyskal, John., Moini, Assad., & Tappan, Mark. (1997). Evolutionary Rapid Development. *Technical Report*. Herndon, Virginia: Software Productivity Consortium. Retrieved December 13, 2010 from http://www.dtic.mil/cgiin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA327979

Moorthi, S.Manthira., Misra, Indranil., Ramakrishnan, R., & Srivastava, P.K. (2011).Discovery of software systems for composition and their customizability in free and open source geospatial software environments – A Case Study. *Map World Forum 2011, Hyderabad, India.* Retrieved January 01, 2011 from http://www.geospatialworldforum.org/2011/proceeding/Dforum.htm

Barik, R.K., Samaddar, A.B. & Gupta, R.D. (2009). Investigations into the Efficacy of open source GIS Software. *GIS Development,* Retrieved December 23, 2010 from http://www.gisdevelopment.net/technology/emerging/mwf09_Barik.htm

Misra, Indranil., Moorthi,S.Manthira., Darji, Nikunj. P., & Ramakrishnan, R. (2010). Open Source software solution for land and crop studies by Geospatial Analysis of Remote Sensing Data. *XXX INCA International Congress Dehradun, India.* 10-12 November, 2010.

Klein S,. Staring M,. Murphy K,. Viergever MA,. & Pluim, J. (2010). elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*. 29 (1), 196-205. doi: 10.1109/TMI.2009.2035616

Lemmens, Rob., & Schouwenburg, Martin. (2007). ILWIS and 52°North: From closed source to open source and interoperable image services. *FOSS4G, Victoria, Canada*. 24-27 September 12, 2007.

The PostgreSQL Global Development Group (2010). PostgreSQL 9.0.4 Documentation User Manual. Retrieved February 12, 2011 from http://www.postgresql.org/docs/manuals/