krugle

White Paper

Software Development Best Practices: Enterprise Code Portal

An Enterprise Code Portal is an “inside the firewall” software solution that enables enterprise software development organizations to unify and intelligently use existing software source code and related information.

The Enterprise Code Portal helps development organizations leverage development knowledge, optimize utilization of existing code and ensure that the best development resources are deployed against strategically significant software development.

The solution uniquely addresses the cross-project access, visibility and management issues that exist “beyond my project” and that apply to the entire organization.

Background

The Enterprise Code Portal continuously monitors all of an organization’s source code management system (SCM) activity for changes and new submittals. New and modified code is automatically incorporated into the Code Portal’s catalog, along with development artifacts - such as related bug tickets, developer submit requirements and documentation. Together, this information is a critically important element of ensuring future code discoverability understanding, and analysis. As a single point of approved releases and the latest “work in progress”, the Code Portal catalog is the single, complete version of the “truth” for an organization’s published code.

Unlike traditional code library or asset management systems, an Enterprise Code Portal solution operates automatically after setup, and offers more flexibility, power and accuracy in discovering code managed in the catalog.

The Enterprise Code Portal provides access to content, analytics, and other tools for development end users. End user developers use the code portal to:

- Explore code demonstrating internal best practices
- Discover code that can be re-used
- Locate like-minded developers and relevant skill sets within the organization
- Perform impact analysis on code in the maintenance lifecycle

Because the Enterprise Code Portal provides a complete, current view of an organization’s source code, it uniquely enables valuable management information to be collected and monitored across the entire organization. The Enterprise Code Portal directly supports and facilitates:

- Reuse measurement
- Code similarity/duplication monitoring and measurement

- Trending of aggregate code development metrics
- Adoption of CMMI level 5 and similar organizational quality practices
- Compliance with internal or industry coding standards

A Look at the Problem - Where Traditional Software Development Solutions Fall Short

Traditional development solutions have focused exclusively on the productivity of the individual development professional (IDEs, compilers, test tools, etc) and the organization and execution of the software project. This has resulted in a wide variety of requirements planning, defect tracking and source code management tools that are largely incompatible.

These tools help ensure predictability, visibility and manageability at the software development project level. As such, these solutions have proven critically important to making development more productive, but have created barriers that make it difficult to use source code as an organizational asset after its initial production.

Traditional development tools make it almost impossible for those who didn't develop the code to access the proper versions of the code or relevant code artifacts. This problem is compounded in most development organizations by an ever growing number of different and proprietary technologies / systems for managing source code development.

As a result, code exists in multiple silos throughout the organization. This makes it extremely difficult to optimize and share the code (and perhaps more importantly, the knowledge embodied in the code) across projects and development teams. At a more basic level, the cross project optimization opportunity cannot be accurately determined.

Despite a lack of quantitative evidence of the problems created by silo'ed code, development managers and executives are aware of the problems. If your organization is concerned about the following types of issues, there is a good chance that silo'ed code has created a business problem:

- What portion of the software development effort is going toward development that is unique and strategically important to the company?
- How effective is the development organization in terms of sharing common, core code across the development organization? What improvement is possible?
- How extensive is development redundancy and code entropy – what are the real costs in lost effort, product quality and $'s?

- How do you improve collaborative development and knowledge sharing in the organization?

- How do we leverage and govern the use of Open Source Software for the development organization?

Additionally, the following are common indicators of a code silo problem:

- **Large # of development teams** – Whenever 3 or more development teams are working on similar projects or using similar development frameworks or components, it is difficult to synchronize or coordinate development work without an Enterprise Code Portal.

- **Legacy code** – Organizations with significant volumes (more than 100,000 lines) of older code (particularly code that exists in old SCM repositories) commonly experience a series of costly and time consuming problems in code maintenance.

- **Multiple code lines / releases** – Organizations that release many, slightly different versions of the same code, or have short release cycles experience many different code coordination issues which can be addressed with an Enterprise Code Portal.

- **Code duplication** – Whenever code is shared or shipped to customers as high level code, there are numerous problems of coordination that become difficult to detect and remedy without Enterprise Code Portal capabilities.

- **Code divergence** – Whenever common code is accessed and then modified, the organization will bear the avoidable costs of maintaining the same code multiple times. The quality and cost problems associated with similar or duplicate code can be minimized with an Enterprise Code Portal

Without an automated solution, understanding and addressing these problems can be impractical and prohibitively expensive. Code must be manually checked out of existing SCM systems and matched with appropriate project metadata for analysis. Depending on the analysis use case, the right analysis tool must be matched to the data. In many cases a non-indexed, local directory based tool (grep, etc) is used to search through the code. This entire process is time consuming (taking anywhere from hours to days to complete a single analysis) and error-prone – due to many manual steps involved and the mismatch of text file based search/discovery tools with computer source code files.

The 4 Key Issues Solved By an Enterprise Code Portal

Issue 1 **Reliable, Low Cost Solution for Cataloging Code and Development Information**

- Collecting code and metadata from multiple SCM systems, across different project is time consuming and error prone

- Maintaining a current record of code "under ongoing development" requires frequent catalog update efforts
- If code cataloging effort is inconsistent, inaccurate, or out of date , developers will not rely on the catalog

An effective Enterprise Code Portal will automatically monitor all specified file servers and SCM systems for additions, edits and deletions. It maintains a current, complete and consistently organized catalog of all code (public, Open Source, Internal) and code artifacts that are of interest to the organization. This catalog becomes a reliable, trusted source of the code "truth" in the development organization.

Issue 2 **Effective Code Discovery**

- To deliver ongoing value to developers, the code catalog must make locating and discovering relevant code fast, intuitive and accurate
- Different development systems have different metadata structures and code access methodologies, which makes manual retrieval confusing, particularly for those with limited exposure to a given system
- Most search utilities search code as "text" and are unable to use the structure and semantics of code structure to facilitate the search. In cases where developers use common tools such as grep, the searches must be performed from scratch at query time. This results in ineffective, inaccurate searching that can take hours or even days to complete.
- Manual attempts to maintain code libraries are rarely successful. One of the biggest problems is that code in the library is not organized for effective discovery across a broad range of use cases.

An Enterprise Code Portal organizes code and code artifacts to promote discovery. The Code Portal must create and use a discovery "index" that provides instantaneous results. The index and associated discovery language should use (i) the structure of the code, (ii) code project metadata and (iii) development artifacts (bug database references, etc) to make ad-hoc code discovery fast and effective.

Issue 3 **Measurement**

- The single biggest problem with library, cataloging and cross project development efforts is the absence of relevant code and code utilization metrics. Development managers currently have little or no visibility into code sharing, code reuse, duplicated effort, or sharable development fixes.

The Enterprise Code Portal supports basic analysis on the entire body of code in the organization, such as static and trended code volume. Advanced analysis can track the use/spread of specified code segments or code libraries and components. These are the "must have" metrics that enable development management to understand and direct efforts to create cross project efficiency in their software development.

Issue 4 **Managed Access**

- Software code is crucial Intellectual Property that must be protected - in all firms with a large software development organization
- The code must be protected from unauthorized access
- Code that is important to organizational productivity must be identified and promoted
- Code utilization must be monitored and analyzed

The Code Portal must provide developers with appropriate level of visibility into the code catalog, and make it easy to find code that is relevant to the developer, subject to existing priorities, access control and governance principles of the organization. The Code Portal must not create a new layer of access control, but instead leverage existing access and governance capabilities wherever possible. Code Portal access must be logged in granular detail.

Enterprise Code Portal Business Benefits

The capabilities that define an Enterprise Code Portal uniquely enable the development organization to manage (typically minimize) the total volume of code in production and maintenance. The Code Portal helps organizations refocus development spending away from common ordinary code, and toward code that is truly unique and strategically important to the organization. This has significant positive impact in the following areas:

Costs – The Enterprise Code Portal promotes effective code consolidation. Less code under management means less effort to modify and maintain the code over time. An analogy would be Body Mass Index (BMI) measurement for code.

Quality – Code consolidation lowers total code volume, and therefore the aggregate volume of defects as well. It also lowers defect density by driving common code to maturity faster. Quality can also significantly improved by leveraging pattern matched fixes and enhancement across the entire code base.

Time-to-market – A lower volume of code is easier to test and more amenable to automated testing. All test milestones can be achieved with higher confidence, sooner in the development cycle.

Competitive Advantage – A smaller, more stable body of common code means that good developers spend less time educating new developers and/or maintaining code. Good developers are 5-10x more productive than typical developers, thus the impact of reducing menial task involvement by 20% - 30%, on output of innovative new code is profound. By allowing top developers to focus on challenging and interesting problems, you also achieve higher retention, further accelerating innovative output.

Detailed Solution Requirements

Table 1 summarizes the features that are commonly required in an Enterprise Code Portal. The two most common deployment options – developing a custom capability from commonly available components, and using a commercially proven engine such as Krugle Enterprise – are compared.

Due to the requirements of integrating, normalizing and preparing vast amounts of development information for developer access and analysis, there are many important components required in a complete solution. Although components such as grep and lucene are available as building blocks, the tasks of designing, building and maintaining an industrial grade Enterprise Code Portal are formidable.

	Capability	Internally Developed	Krugle Enterprise	Notes
Automatic Cataloging	Unified, scaled code catalog	Custom index / DB e.g. Lucene	Built-in	To store and access all code under management, including all versions
	Direct/API integration with popular SCM systems: Clearcase, SVN, CVS, Perforce, PVCS, Starteam, Team Foundation, etc	Custom scripts for each SCM	Built-in	For rich, high integrity synchronization with all source code information
	File system crawling	Custom script	Built-in	For special case code and code artifacts
	API for development database and artifact integration	Custom script for each DB	Built-in	To enhance code discovery
	Automatic catalog updating	Custom application	Built-in	Keep a complete and current point-in-time library of all code
	Editorial code project configuration	Manually or custom app	Via secure admin	Select code to be included, set ranking order for similar results
	Build system integration	Custom for each build system	Built-in API	Automatic addition / inclusion of new projects or baseline releases
	System for backup, high availability	IT config. & mgmt.	Plug & play appliance	For dependable reporting and ad-hoc dev access
Measurement	Measure / log all system activity	Custom	Included	Understand use patterns and activity in system
	Trended code volume analysis	Custom	Included	Measure and manage total volume of code
	Measure code clones (duplicates), code similarity and clone divergence	Custom	Included	ID areas for code sharing, refactoring, SOA. Detect Open Source use
	Defect propagation analysis	Custom	Included	Find a bug once, fix it in all locations
	Code license, language, churn reports	Custom	Included	Measure evolution of code by key attributes
Discovery	Code snippet search	Custom	Included	Find similar code patterns: defects / performance work
	Search files as text	grep /grok on local code	Included	Discover code and artifacts via keyword(s)
	Search by code attributes, artifacts	Custom application	Included	Use project metadata, or code characteristics
	Developer desktop interface – API	Custom	Included	For integration with other portal and team systems
	Developer desktop interface – browser, IDE	Custom	Included	Exploratory research
Access	Administrative control of catalog, users and use logging	Custom	Included	
	Ability to promote, analyze and guide access choices	Custom	Included	Guide users to popular or important code
	Control access by project, code license, file , project and user	Custom	Included	Ensure appropriate use and code governance
	Developer Level Access control	Custom Integration	ActiveDir. LDAP, API	Authorized access option

Table 1. Enterprise Code Portal Requirements

Summary

Cross-project code access problems represent a large and important opportunity in large scale software development efforts. These problems consume a significant portion of the budget in most software development organizations, and can result in violated SLAs, penalty fees, lost revenues, and in some cases, can result in the development executive's job.

These problems can be addressed through an Enterprise Code Portal. A workable solution requires a reliable comprehensive code catalog, useful analytics, discovery capabilities and managed accessed. For a complete and reliable solution, development organizations can evaluate integrated solutions like Krugle Enterprise.

Once deployed the Code Portal helps organizations improve developer productivity through IP sharing, minimizing code under maintenance, and focusing development resources on strategically important projects.