I’ve just inherited 1,000,000 lines of code… now what?

Michael Rozlog
Product Manager, RAD Studio, Delphi for PHP, and 3rd Rail
Who is Mike?

- You just said that
  - Product Manager – RAD Studio, Delphi for PHP, 3rd Rail
- Responsibilities
  - Helping to build the developer tools that help developers develop world class applications
- Current research areas
  - Software Archeology
  - Situational Applications
  - Pushing the RIEA approach – MS’s Software+Services
- Prior Position at Borland
  - Chief Technical Architect – Borland
  - Creator of Mastering JBuilder – Wiley Books

michael.rozlog@codegear.com
Agenda

• What is software Archeology?
• The process: Software Archeology?
• Demonstration
• Q&A
Current state of IT around the world

- Down economy means less new projects, but more maintenance and updating existing software
- However...
- The need for new or modernizing applications is not decreasing or going away
- U.S. had the lowest C/S graduation in 2007
  - Qualified individuals are not filling the market
- Outsourcing is not always a solution on the table?
  - Not allowed or applicable
  - Bigger issues around this area, latest data shows 50% for non-government organizations
- Knowledge transfer gap
Questions to you…

- How many have inherited code in the last 6 months that you are now responsible for?
- How many have either sent off or received outsourced projects in the last year?
- How many find it hard to retain and keep skilled employees?
- How many have new members on there team in the last 3 months?
What is Software Archeology?
What is Software Archeology?

• A process for approaching unknown software that we have become responsible for...

• An approach to unraveling the complexities of an existing application

• The analysis of past applications to learn and understand why it was done that way
Software Archeology

```plaintext
void AddControls(TWinControl* AWinControl, TStrings* Strings)
{
    for(int i=0; i < AWinControl->ControlCount; i++)
    {
        if (dynamic_cast<TWinControl*>(AWinControl->Controls[i]))
            AddControls(dynamic_cast<TWinControl*>(AWinControl->Controls[i]), Strings);
        else
            AddCtrl(AWinControl->Controls[i], Strings);
    }
}

if (host.charAt(0) == '.')
    return non_proxy_dom_list.removeElement(host);

// check for host name
for (int idx=0; idx<host.length(); idx++) {
    if (!Character.isDigit(host.charAt(idx))  &&
        host.charAt(idx) != '.'  &&  host.charAt(idx) != '/')
        return (non_proxy_host_list.remove(host) != null);
}

// must be an IP-address

procedure TFormMain.DoPageShow(Sender: TObject);
var
    Ctrl: TWinControl;
begin
    PackageInstaller.BeginUpdate;
    try
        Ctrl := CreatePage(Sender as TJvWizardInteriorPage, PackageInstaller.Page);
        while (Ctrl <> nil) and (not Ctrl.Enabled) do
            Ctrl := FindNextControl(Ctrl, True, False, False);
        ActiveControl := Ctrl;
        DoUpdateNavigation(Sender);
    finally
        PackageInstaller.EndUpdate;
    end;
```
if (host.charAt(0) == '.')
    return
// check for host name
for (int idx=0; idx<host.length(); idx++) {
    if (!Character.isDigit(host.charAt(idx)) 
        && host.charAt(idx) != '.' 
        && host.charAt(idx) != '/
        return
    (non_proxy_host_list.remove(host) != null);
// must be an IP-address
SOFTWARE ARCHEOLOGY

if (host.charAt(0) == '.')
    return

// check for host name
for (int idx=0; idx<host.length(); idx++) {
    if (!Character.isDigit(host.charAt(idx))
        && host.charAt(idx) != '.'
        && host.charAt(idx) != '/')
        return (non_proxy_host_list.remove(host) != null);
}

// must be an IP-address
if (host.charAt(0) == '.')
    return
non_proxy_dom_list.removeElement(host);

// check for host name
for (int idx=0; idx<host.length(); idx++) {
    if (!Character.isDigit(host.charAt(idx)) &&
        host.charAt(idx) != '.' && host.charAt(idx) != '/')
        return
        (non_proxy_host_list.remove(host) != null);
}
// must be an IP-address
Why is Software Archeology important?

- Always asked to rebuild the wheel
  - How many times have you been asked to recreate this current implementation, but better?
- Inheritance of code
- Outsourcing
- Graying of IT
Why is Software Archeology important?

• Always asked to rebuild the wheel
• Inheritance of code
  – How many times have you picked up someone else's code or project… to be enhanced?
• Outsourcing
• Graying of IT
Why is Software Archeology important?

- Always asked to rebuild the wheel
- Inheritance of code
- Outsourcing
  - Projects get outsourced, rural sourced, in-sourced, local sourced, and every other kind of sourced...
  - Do you understand the systems that were created?
  - Do the outsourcers understand the system they are inheriting?
- Graying of IT
Why is Software Archeology important?

- Always asked to rebuild the wheel
- Inheritance of code
- Outsourcing
- Graying of IT
  - By 2010... almost 37% of the software industry will consist of persons older then 50!
  - 60% of the federal workforce will be eligible to retire over the next 10 years
    - By 2010... 18.5 percent of the government workforce – about 290,000 employees – will retire
  - Do you understand the systems they have created?
What can be learned?

• High-level architectural understanding
• Weak-points in existing software
  – Maintenance nightmares
  – Performance bottlenecks
• Standards compliance or lack there of…
• The overall health and value of the application
• Architectural approaches or how not to solve a problem
• Finding and Harvesting of Algorithms and patterns
What tools are needed?

• Need to be able to get a picture
• Need to be able to do static analysis
• Need to be able to do Dynamic analysis
• Need to be able to test the system

• Can it be done by hand?
  – Yes… but very time consuming… code reviews, system tests, and primitive system performance models

• Is there a better way?
  – Use a software tool
What is the cost of code reviews

• Very controversial subject
• But for me tools can enhance the approach

• Task
  – Review by hand a small example set of code
    • 5 classes, 139 lines of code, 4 “experts” in 15 minutes
    • Average audit violations detected: 21 time: 0.25 hr

• Comparison
  – The same set of code analyzed by a machine:
    • Audit violations detected: 150 time: 2 sec
  – Humans @ 84/hr – U.S.

• Real “sample” Project
  – 187,600 violations in 100 minutes on a project with 1233 classes and 199,146 lines of code
  – 32,443 violations using the default set
  – Actual customer case studies show ~400 man/hour savings per code review

Remember: remove the subjective and move to the objective
Demonstrations

- Example Showing JBuilder
- Example Showing RAD Studio
Questions & Answers
クイズに答えて賞品を当てよう！

問題
「エンパカデロ、2010年はツール○○○○」。
○○○○に入る4文字のカタカナは？

素敵な賞品が当たる！応募はこちらから
http://forms.embarcadero.com/forms/dcamp_quiz

応募締切は3月12日15時まで！
Thank You

michael.rozlog@embarcadero.com
http://blogs.embarcadero.com/michaelrozlog