Supporting Software Development as A Knowledge-Intensive and Collaborative Activity

Yunwen Ye
Center for LifeLong Learning and Design
University of Colorado
SRA Key Technology Lab
Software development is collaborative

Software systems are built by many people
  - Power of many hands
  - Division of labor

Many kinds of collaboration exist
  - Consensus building
  - Brainstorming of design ideas
  - Coordination of individual development
## Software development is knowledge intensive

**Knowledge from multiple domains**

- Problem domain
- Programming language
- Development tools
- Operating systems
- Programming libraries
- The system under development

**Programming is essentially a learning process**
From *brawny* collaboration to *brainy* collaboration

Software systems are built by many people

- Power of many hands --> Power of many heads
- Division of labor --> Distribution of knowledge

Supporting *knowledge collaboration* in software development

- Accessing and integrating distributed knowledge
Knowledge vs. knowing capability

Knowledge: stored artifacts in the head

Knowledgeable behavior: effective behavior to perform a task

Knowing capability: the capability of accessing and integrating external knowledge for performing a given task
Overall capability of a project team

NOT the sum of the knowledge of individual developers

But the capability of activating knowledge transfer and integrating what individuals know
Software project as a self-organizing and evolving knowledge ecosystem

Knowledge is distributed over resources artifacts (code, documents) and developers

Knowledge flows with interactions of knowledge resources
- developer -> artifact
- artifact -> artifact
- artifact -> developer
- developer -> developer
Supporting knowledge collaboration

Activate all available knowledge resources (technical and social support) in the knowledge ecosystem
Continuous socio-technical support

Technical support (knowledge-repository)
- At the bid of users
- Difficult to capture tacit and contextual knowledge
- Inexpensive

Social support (community-based)
- Dealing with tacit knowledge
- At the mercy of experts
- Expensive

The seamless integration of the two approaches
The challenges

Collaboration with cognitive tools (the technical dimension)
What information should be added to complement developers’ insufficient knowledge in head

Collaboration with knowledgeable peers (the social dimension)
Who has the knowledge I need?
Who is willing to help me at this moment?
Do I look stupid if I ask this question?
Is the information reliable?

Why should I help him/her?
What happens if I don’t help him/her?
I have my own work to do!
Learning Java API on demand

CodeBroker

STeP_IN (Socio-Technical Platform for in situ Networking)
Rapid growth of Java API
Rapid growth of Java API
## Libraries used in STeP_IN

<table>
<thead>
<tr>
<th>Library Name</th>
<th>Class</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>activation.jar</td>
<td>38</td>
<td>261</td>
</tr>
<tr>
<td>bcel-5.1.jar</td>
<td>373</td>
<td>3093</td>
</tr>
<tr>
<td>commons-collections-3.1.jar</td>
<td>446</td>
<td>4021</td>
</tr>
<tr>
<td>commons-dbcp-1.2.1.jar</td>
<td>44</td>
<td>935</td>
</tr>
<tr>
<td>commons-pool-1.2.jar</td>
<td>25</td>
<td>277</td>
</tr>
<tr>
<td>jun547.jar</td>
<td>2640</td>
<td>18412</td>
</tr>
<tr>
<td>mail.jar</td>
<td>240</td>
<td>1966</td>
</tr>
<tr>
<td>postgresql.jar</td>
<td>82</td>
<td>1216</td>
</tr>
<tr>
<td>resolver.jar</td>
<td>29</td>
<td>298</td>
</tr>
<tr>
<td>StPL75.jar</td>
<td>175</td>
<td>1384</td>
</tr>
<tr>
<td>xercesImpl.jar</td>
<td>784</td>
<td>7463</td>
</tr>
<tr>
<td>xml-ParserAPIs.jar</td>
<td>207</td>
<td>1748</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5083</strong></td>
<td><strong>41074</strong></td>
</tr>
</tbody>
</table>
Creating awareness of unknown API

```java
package test;

public class TestExecOutput {
    /**
     * execute an external command *
     */

    home/yunwen/cb/demo/data/cbindindex/
```
Creating the awareness of unknown API

package test;

public class TestExecOutput {
    /*** execute an external command ***/

    0.3 getCommand Retrieves this RowSet object's command property
    0.3 exec Executes the specified string command in a separat
    0.3 exec Executes the specified command and arguments in a
    0.3 executeBatch Submits a batch of commands to the database f
    0.3 *RCI-display* 05-19 11:40 AM 0.10 (ReusableComponentInfo)--
Mouse triggered information

define

class TestExecOutput {
    /** * execute an external command */
}

package test;

getCommand
exec
executeBatch

getCommand
exec
executeBatch

RCI-display
java.lang.Runtime::public java.lang.Process exec(String command)
A Java method `exec` is shown, which takes a `String` command and returns a `Process` object. The method executes the specified string command in a separate process. Parameters include a specified system command, and exceptions handled include `SecurityException`, `IOException`, `NullPointerException`, and `IllegalArgumentException`. The documentation also includes see also links to related methods.
Finding examples

```
java.lang.Runtime
java.lang.Process exec(java.lang.String)
```

This example is provided by vincent.

```
// execute external commands in a Java program
try {
    String command = "ls";
    Process p = Runtime.getRuntime().exec(command);
    // do something with p
}
catch (IOException e) {
    // do something
}
```

This example is: ☑ Helpful   ☐ Not helpful   evaluate
Reading archived discussion

Discussion Archive for:

java.lang.Runtime
define.ById(Objects,java.lang.String)

Discussion:

[Examples] [Ask Expert]

From: vincent
Subject: how to use exec
Date: 2006/03/02 12:24:25

I want to execute an external command, such as 'ls', but I don't know how to use exec in context. Could someone please to help me. Thanks in advance

[Examples]

From: vincent
Subject: call an external command
Date: 2006/03/02 12:33:26

I need to call an external command from a Java program. Could someone help me using exec. Thx in advance.

From: vincent
Subject: help on exec
Date: 2006/03/02 16:02:08

could someone please tell me how to use exec? thanks.

read output of command
Asking the experts

Requesting collaboration from experts on:

Subject: read the output
Question: How do I read the output of the executed command in the Java program?
Emails to dynamically chosen experts

Message 22:
From dync+17@bud.cs.colorado.edu   Thu Mar 9 05:27:03 2006
X-Original-To: ginnny@bud.cs.colorado.edu
From: dync+17@bud.cs.colorado.edu
To: ginnny@bud.cs.colorado.edu
Subject: [DynC 17] read the output of the command
MIME-Version: 1.0
Content-Type: text/plain; charset=UTF-8
Content-Transfer-Encoding: 7bit
Date: Thu, 9 Mar 2006 05:27:02 -0700 (MST)

harry has requested help on
java.lang.Runtime
Process exec (String)

Your email address will be automatically changed to your STeP_IN UserName when your reply is sent.
Please don't include your signature if you want to remain to be known only by your UserName in STeP_IN.

------Question Contents--------
how to read the output of the command's execution.
thanks.

&
A DynC is a small group of knowledge workers that forms *ad hoc* in support of a particular *user* working on a particular *task*, and dissembles as the task is finished.

- High expertise on a given task
- Established social relationship with the person
The forming process of a DynC
Triggering event for $DynC(A, \alpha)$
From information to information
From information to people (experts)
From people to people
$$DynC(A, \alpha) = \{A, B, C, D, E\}$$
### Initializing Technical Profile (Developer-Info)

![Technical Profile Manager](image)

#### Classpath:
```classpath
D:\Projects\DynC\STeP_IN\lib\StPL71.jar
```

#### Table of Class Definitions and References:

<table>
<thead>
<tr>
<th>Package</th>
<th>Class</th>
<th>Method</th>
<th># of Definitions</th>
<th># of References</th>
<th>Check</th>
</tr>
</thead>
<tbody>
<tr>
<td>java.lang</td>
<td>StringBuffer</td>
<td>java.lang.String</td>
<td>0</td>
<td>60</td>
<td>✓</td>
</tr>
<tr>
<td>java.lang</td>
<td>StringBuffer</td>
<td>java.lang.StringBuffer</td>
<td>0</td>
<td>56</td>
<td>✓</td>
</tr>
<tr>
<td>java.lang</td>
<td>StringBuffer</td>
<td>StringBuffer()</td>
<td>0</td>
<td>55</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>StObject</td>
<td>StObject()</td>
<td>1</td>
<td>40</td>
<td>✓</td>
</tr>
<tr>
<td>java.lang</td>
<td>Object</td>
<td>Object()</td>
<td>0</td>
<td>34</td>
<td>✓</td>
</tr>
<tr>
<td>java.awt</td>
<td>Point</td>
<td>Point(int, int)</td>
<td>0</td>
<td>29</td>
<td>✓</td>
</tr>
<tr>
<td>java.io</td>
<td>PrintStream</td>
<td>void</td>
<td>0</td>
<td>28</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>StImage</td>
<td>java.awt.image.Buffer</td>
<td>1</td>
<td>27</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>StRectangle</td>
<td>StRectangle(int, )</td>
<td>1</td>
<td>22</td>
<td>✓</td>
</tr>
<tr>
<td>java.lang</td>
<td>Object</td>
<td>java.lang.Class</td>
<td>0</td>
<td>22</td>
<td>✓</td>
</tr>
<tr>
<td>java.util</td>
<td>Vector</td>
<td>int size()</td>
<td>0</td>
<td>21</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>StImage</td>
<td>int width()</td>
<td>1</td>
<td>19</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>StImage</td>
<td>int height()</td>
<td>1</td>
<td>18</td>
<td>✓</td>
</tr>
<tr>
<td>java.lang</td>
<td>StringBuffer</td>
<td>java.lang.StringBuffer</td>
<td>0</td>
<td>17</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>SmalltalkException</td>
<td>SmalltalkException</td>
<td>1</td>
<td>17</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>SmalltalkException</td>
<td>SmalltalkException</td>
<td>1</td>
<td>17</td>
<td>✓</td>
</tr>
<tr>
<td>java.util</td>
<td>ArrayList</td>
<td>int size()</td>
<td>0</td>
<td>16</td>
<td>✓</td>
</tr>
<tr>
<td>java.lang</td>
<td>String</td>
<td>boolean</td>
<td>0</td>
<td>15</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>SmalltalkTestExa...</td>
<td>SmalltalkTestExa...</td>
<td>1</td>
<td>15</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>StObject</td>
<td>jp.co.sra.smalltalk.St</td>
<td>1</td>
<td>14</td>
<td>✓</td>
</tr>
<tr>
<td>java.util</td>
<td>Vector</td>
<td>java.lang.Object</td>
<td>0</td>
<td>14</td>
<td>✓</td>
</tr>
<tr>
<td>jp.co.sra.smalltalk</td>
<td>StBlockClosure</td>
<td>StBlockClosure()</td>
<td>1</td>
<td>14</td>
<td>✓</td>
</tr>
<tr>
<td>java.awt.event</td>
<td>WindowAdapter</td>
<td>WindowAdapter()</td>
<td>0</td>
<td>14</td>
<td>✓</td>
</tr>
</tbody>
</table>
 Initializing Social Profile (Human-Human) 

```
<table>
<thead>
<tr>
<th>name</th>
<th>domain</th>
<th># of messages</th>
<th>check</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yunwen.Ye</td>
<td>Colorado.EDU</td>
<td>154</td>
<td>✔</td>
</tr>
<tr>
<td>m</td>
<td>sra.co.jp</td>
<td>39</td>
<td>✔</td>
</tr>
<tr>
<td>g</td>
<td>cs.colorado.edu</td>
<td>8</td>
<td>✔</td>
</tr>
<tr>
<td>m</td>
<td>is.aist-nara.ac.jp</td>
<td>14</td>
<td>✔</td>
</tr>
<tr>
<td>i</td>
<td>ist.osaka-u.ac.jp</td>
<td>13</td>
<td>✔</td>
</tr>
<tr>
<td>o</td>
<td>empirical.jp</td>
<td>15</td>
<td>✔</td>
</tr>
<tr>
<td>y</td>
<td>computer.org</td>
<td>30</td>
<td>✔</td>
</tr>
<tr>
<td>m</td>
<td>ist.osaka-u.ac.jp</td>
<td>36</td>
<td>✔</td>
</tr>
<tr>
<td>b</td>
<td>ics.uci.edu</td>
<td>16</td>
<td>✔</td>
</tr>
<tr>
<td>k</td>
<td>sra.co.jp</td>
<td>87</td>
<td>✔</td>
</tr>
<tr>
<td>a</td>
<td>sra.co.jp</td>
<td>42</td>
<td>✔</td>
</tr>
<tr>
<td>l</td>
<td>uci.edu</td>
<td>1</td>
<td>✔</td>
</tr>
<tr>
<td>k</td>
<td>ics.es.osaka-u.ac.jp</td>
<td>1</td>
<td>✔</td>
</tr>
<tr>
<td>t</td>
<td>colorado.edu</td>
<td>2</td>
<td>✔</td>
</tr>
<tr>
<td>n</td>
<td>is.aist-nara.ac.jp</td>
<td>3</td>
<td>✔</td>
</tr>
<tr>
<td>k</td>
<td>gray.plala.or.jp</td>
<td>3</td>
<td>✔</td>
</tr>
<tr>
<td>m</td>
<td>sra.co.jp</td>
<td>6</td>
<td>✔</td>
</tr>
<tr>
<td>y</td>
<td>kid.rock.u-tokyo.ac.jp</td>
<td>8</td>
<td>✔</td>
</tr>
<tr>
<td>m</td>
<td>is.naist.jp</td>
<td>39</td>
<td>✔</td>
</tr>
<tr>
<td>d</td>
<td>gte.net</td>
<td>3</td>
<td>✔</td>
</tr>
</tbody>
</table>
```

**MBOX:** C:\shared\dyic-mbox

[Image of Step_IN Profile Manager interface]
Search interface

Search for...
normalize token

Options...
Scope: 
Filter: 
Return only methods: ○ I used ○ I never used ○ Both

normalize
org. apache. lucene. analysis. CharTokenizer.
Called on each token character.
normalize it before.
protected char normalize(char).

normalize
org. apache. lucene. analysis. LowerCaseTokenizer.
Collects only characters where
{link Char...} protected char normalize(char).

setOmitNorms
Expert: If set, omit normalization
associa...

normalize
... public void normalize(float weight)...

normalize
org. apache. lucene. search. Weight.
Assigns the query normalization
this.... public void normalize(float weight)...

isTokenChar
... Returns true iff a character...

Done
normalize

protected char normalize(char c)

Called on each token character to normalize it before it is added to the token. The default implementation does nothing. Subclasses may use this to, e.g., lowercase tokens.

throws java.io.IOException

Returns the next token in the stream, or null at EOS.

Specified by:
    next in class TokenStream

Throws:
    java.io.IOException

Example:
...
Example

Examples for: org.apache.lucene.analysis.CharTokenizer
cchar normalize(char)

This example is provided by lu1612.
1 found it helpful.
0 found it not helpful.

```java
if (isTokenChar(c)) { // if it's a token char
    if (length == 0)
        start = offset - 1;
    buff[length++] = normalize(c);
}
```

This example is: Helpful Not helpful evaluate

This page was created by STeP_IN (001) on 2006/10/23 04:49:17
Discussion Archive for: org.apache.lucene.analysis.CharTokenizer
char normalize(char)

DynCs:

**Similarity score computation documentation**
lu1192
2004/09/21 03:52:09

**Re(2): Lucene with Number+Text**
lu229
2002/03/26 16:10:00

any_other_operations_performed_in_this_method?
lu1283
2006/10/04 17:05:20

Discussions:

From: lu1192
Subject: Similarity score computation documentation
Date: 2004/09/21 03:52:09

Hi,

I was looking through the source code and think there may be a discrepancy in org.apache.lucene.search.Similarity's actual implementation.

I believe the problem is only occurring in specific situations.

I'm pretty sure that there is a discrepancy in the implementation of org.apache.lucene.search.TFIDF and the official documentation.

I can see that first sumOfSquares is called and then normalize() is called at the end of search.

[Examples] [Ask Expert]
Requesting collaboration from experts on: org.apache.lucene.analysis.CharTokenizer char normalize(char)

Subject: normalize
Question: Could some one please tell me the algorithm used in this method?
Evaluate a DynC

lu1283's Involved DynC

Active DynC you initiated

<table>
<thead>
<tr>
<th>DynC ID</th>
<th>Subject</th>
<th>Method</th>
<th>Start Date</th>
<th>This DynC was helpful; thank you for your help. This DynC was not helpful; thank you for participation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2530</td>
<td>normalize</td>
<td>CharTokenizer char normalize(char)</td>
<td>2006/10/23 01:55:16</td>
<td></td>
</tr>
</tbody>
</table>

This page was created by STeP_IN (001) on 2006/10/23 02:54:00
Forming a DynC in STeP_IN

Expert identification
Create an ordered list of candidate experts based on their expertise on the specific API

Expert selection
From the list of candidate experts, select those who are likely willing to help the programmer
Expert identification

Level1: Confirmed expertise
Level2: Claimed expertise
Level3: Inferred expertise
Level4: Future expertise
Expert identification

Level 1: Confirmed expertise

Lu953's Involved DynC

Active DynC you initiated
Others' active DynC you are participating
Past DynC you initiated
Past DynC you participated

Past DynC you participated

<table>
<thead>
<tr>
<th>DynC ID</th>
<th>Subject</th>
<th>Method</th>
<th>Start Date</th>
<th>End Date</th>
<th>Helpful/Not helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>834</td>
<td>can lucene be backed to have an update field</td>
<td>MultipleTermPositions int doc()</td>
<td>2004/11/10 06:16:55</td>
<td>2004/11/10 06:16:55</td>
<td>helpful</td>
</tr>
<tr>
<td>975</td>
<td>Similarity score computation documentation</td>
<td>CharVectorizer char normalize(char)</td>
<td>2004/09/21 03:52:09</td>
<td>2004/09/21 03:52:09</td>
<td>helpful</td>
</tr>
</tbody>
</table>
## Expert identification

### Level2: Claimed expertise

### lu953's Technical Profile

#### Profile Summary
- Total methods used: 78
- Methods declared as Expert: 0
- Methods declared as No Knowledge: 0

#### Display methods...
- All methods used

#### Options...
- Scope:
- Filter:

#### Method Table

<table>
<thead>
<tr>
<th>Method</th>
<th>Author</th>
<th>Usage #</th>
<th>Declare</th>
</tr>
</thead>
<tbody>
<tr>
<td>org.apache.lucene.analysis.CharTokenizer</td>
<td>No</td>
<td>27</td>
<td>Expert, Not Declared, No Knowledge</td>
</tr>
<tr>
<td>org.apache.lucene.analysis.Token next()</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>org.apache.lucene.analysis.CharTokenizer</td>
<td>No</td>
<td>2</td>
<td>Expert</td>
</tr>
<tr>
<td>char normalize(char)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>org.apache.lucene.analysis.Token</td>
<td>No</td>
<td>1</td>
<td>Expert, Not Declared, No Knowledge</td>
</tr>
<tr>
<td>int getPositionIncrement()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>org.apache.lucene.analysis.Token</td>
<td>No</td>
<td>3</td>
<td>Expert, Not Declared, No Knowledge</td>
</tr>
<tr>
<td>Token java.lang.String termText()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>org.apache.lucene.analysis.TokenFilter</td>
<td>No</td>
<td>5</td>
<td>Expert, Not Declared, No Knowledge</td>
</tr>
</tbody>
</table>
Expert identification

Level 3: Inferred expertise

**lu1192's Technical Profile**

**Profile Summary**
- Total methods used: 10
- Methods declared as Expert: 0
- Methods declared as No Knowledge: 0

**Display methods...**
- All methods used

**Options...**
- Scope:
- Filter:

```plaintext
<table>
<thead>
<tr>
<th>Method</th>
<th>Author</th>
<th>Usage #</th>
<th>Declare</th>
</tr>
</thead>
<tbody>
<tr>
<td>org.apache.lucene.analysis CharTokenizer</td>
<td>No</td>
<td>4</td>
<td>Expert Not Declared</td>
</tr>
<tr>
<td>char normalize(char)</td>
<td></td>
<td></td>
<td>No Knowledge</td>
</tr>
<tr>
<td>org.apache.lucene.document Document float getBoost()</td>
<td>No</td>
<td>3</td>
<td>Expert Not Declared</td>
</tr>
<tr>
<td>org.apache.lucene.index Term java.lang.String field()</td>
<td>No</td>
<td>1</td>
<td>Expert Not Declared</td>
</tr>
<tr>
<td>org.apache.lucene.search DefaultSimilarity float idf(int, int)</td>
<td>No</td>
<td>1</td>
<td>Expert Not Declared</td>
</tr>
</tbody>
</table>
```

[Previous] 1 [Next]

**Done**
Expert identification

**Level4: Future expertise**

Asked a question on the same API in the past

Did not get help from the DynC and gave the evaluation of “not helpful”
Expert selection

4 kinds of relations

help<A,B>
friend<A,B>
exclude<A,B>
email<A,B>
Expert selection

4 kinds of relations

help\langle A,B\rangle
friend\langle A,B\rangle
exclude\langle A,B\rangle
email\langle A,B\rangle

A(lu953) contributed (sent an email) to the DynC initiated by B(lu1281)
A (lu953) declares that he/she will *always* participate in the DynCs initiated by B (lu1283).
A (lu953) declares that he/she will never participated in the DynCs initiated by B(lu1282).
Expert selection

4 kinds of relation

help<A,B>
friend<A,B>
exclude<A,B>
email<A,B>

The number of emails that A (lu953) sent to B (lu1281) outside of STeP_IN.
Expert selection process

for each X in the list of candidate experts

1. If exclude<X, A> // X does not want to participate in A’s DynC
   Remove X from the list of candidate experts
for each X in remaining list of candidate experts

2. If friend\(<X, A>\)  // X wants to participate in A's DynC

Add X to the DynC
for each X in remaining list of candidate experts

3. If $\text{help}(A, X) - \text{help}(X, A) > 0$

   // A has helped X more than the other way around

   Add X to the DynC
Expert selection process

for each X in remaining *list of candidate experts*

4. A recently helped X

   Add X to the DynC
Expert selection process

for each X in remaining list of candidate experts

5. Add those Xs that have been helped more than they have helped others.
for each X in remaining list of candidate experts

6. Choose Xs according to the number of email<X, A>
Socially aware communication

To sustain knowledge collaboration

Freedom of no-participation to avoid forced collaboration
Give high priority to the motivation of experts
The success of one collaboration should not come at the cost of future collaboration

Asymmetrical information disclosure
No-participation is socially acceptable
Socially aware communication

To say “NO” to certain DynCs in a socially acceptable way

Notification email of DynC formation
Socially aware communication

To say “NO” to certain DynCs in a socially acceptable way

Leave this DynC

Notification email of DynC formation
Socially aware communication

To say “NO” to certain DynCs in a socially acceptable way

Notification email of DynC formation

No participation in all DynCs on this method
Socially aware communication

To say “NO” to certain DynCs in a socially acceptable way
Socially aware communication

Asymmetry of information disclosure
DynC members are not made public
Members who contributed are acknowledged publicly
Software development is knowledge intensive and collaborative

Providing technical and social support for the easy acquisition of external knowledge when needed

http://stepin.cs.colorado.edu/STeP-IN/