

Exemplar: EXEcutable exaMPLes ARchive

Mark Grechanik, Chen Fu, Qing Xie
Accenture Technology Labs

Collin McMillan, Denys Poshyvanyk
College of William & Mary

Chad Cumby
Accenture Technology Labs



A Search Engine for Finding Highly-Relevant Applications

- Online repositories contain many millions of lines of code, but reusing them is a very difficult problem.
- The **high-level descriptions** of applications do not usually match its **low-level implementation** details.

```
public static long mystery(long a, long b)
{
    if (b==0)
        return a;
    else
        return mystery(b, a % b);
}
```



Compute the **Greatest Common Denominator**

- We present Exemplar, a source code search engine to bridge this mismatch by integrating API help documentation into the search process.

Example Programming Task

Write an application to record musical instrument data to a file in the MIDI file format.



What programmers do

The programmer may check other applications for API calls from third-party packages used to read data from a MIDI device and then print to a file.

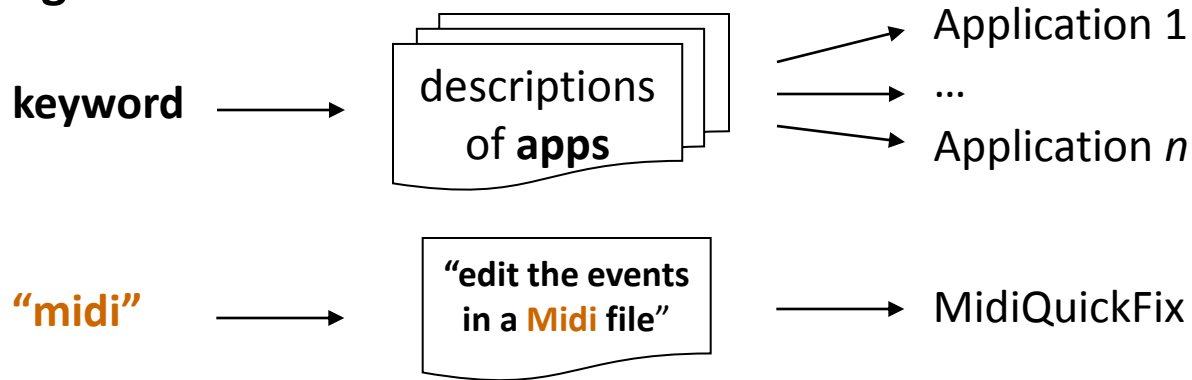
```
List allInfos = new ArrayList();
List providers = getMidiDeviceProviders();
...
MidiDevice.Info[] infosArray
    = (MidiDevice.Info[]) allInfos.toArray(new MidiDevice.Info[0]);

for(int i = 0; i < infosArray.size(); i++) {
    fileOutput.print(infosArray[i]);
    ...
}
```

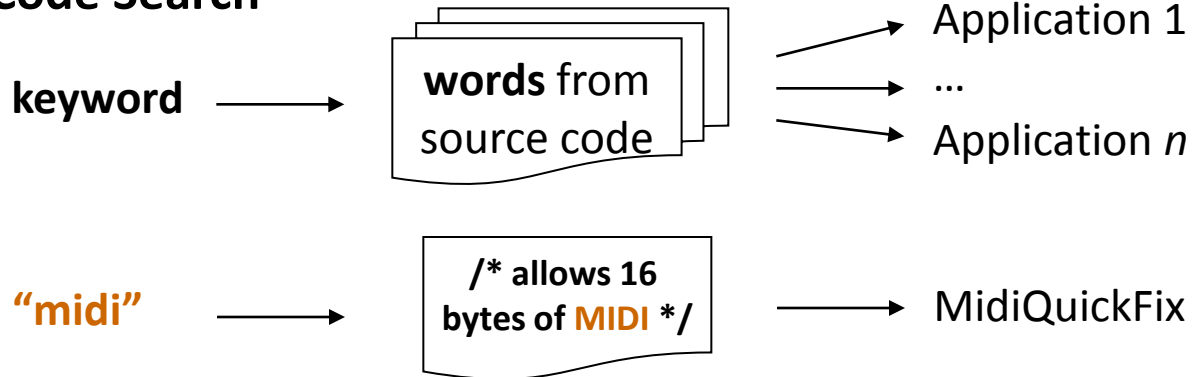
8,000 Java projects that we extracted from SourceForge make over 11,000,000 calls to the official Java API.

Many search engines rely on words from applications.

SourceForge

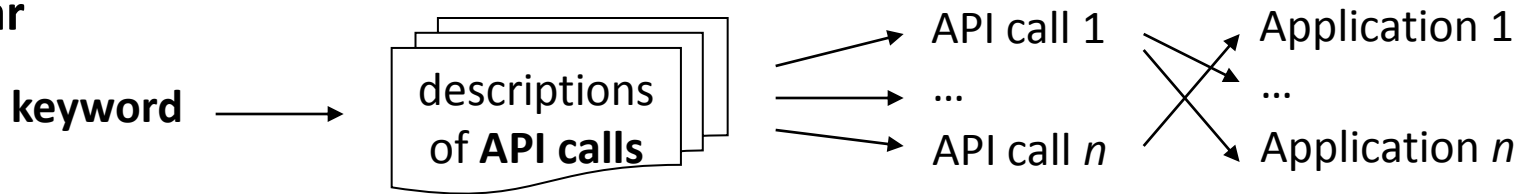


Google Code Search

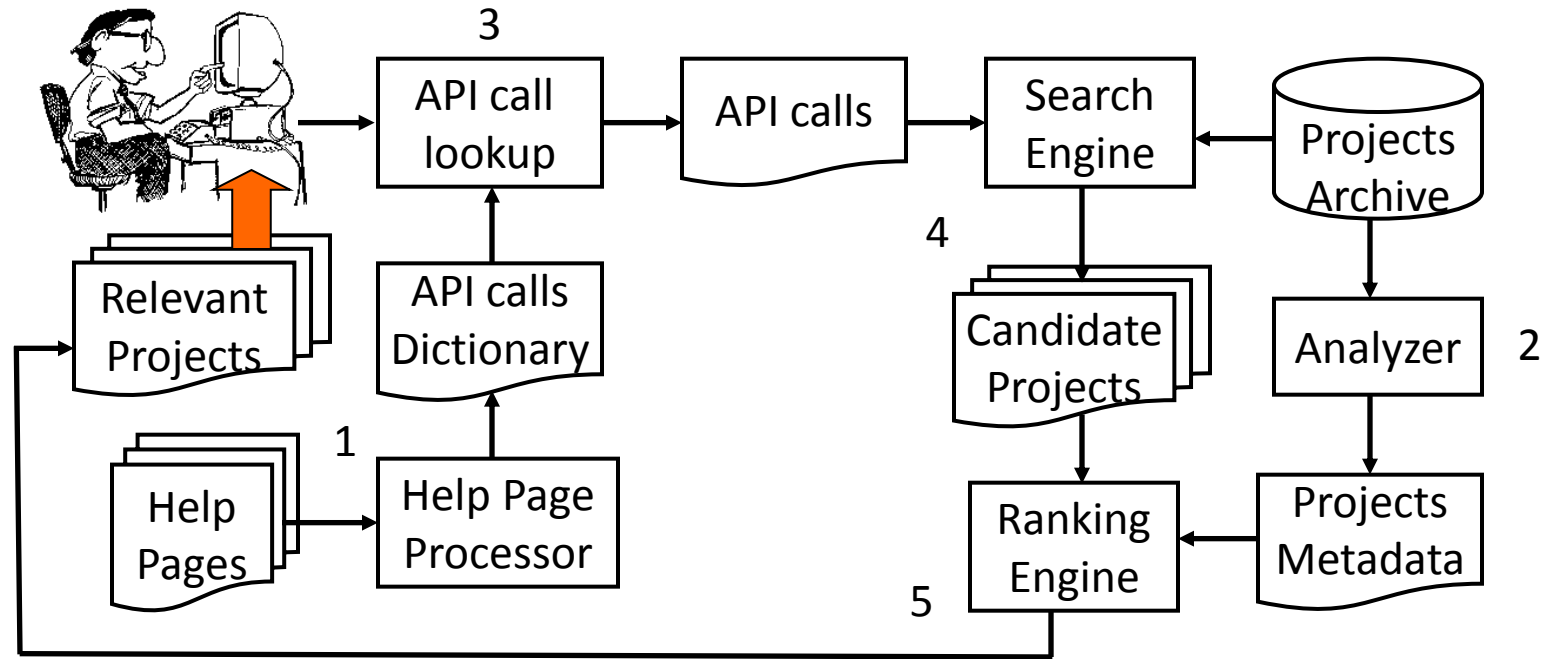


Our idea is to augment standard code search to include API documentation.

Exemplar



"record midi file"



javax.sound.midi.MidiDevice.getReceiver()

... Obtains a MIDI IN receiver through which the MIDI device may receive MIDI data ...

javax.imageio.ImageWriter.write()

... Appends a complete image stream containing a single image ...

java.awt.geom.AffineTransform.getScaleY()

... scaling element (m11) of the 3x3 affine transformation matrix ...

AffineTransform.getScaleY()

AffineTransform.createInverse()

Jazilla

ShortMessage.ShortMessage()

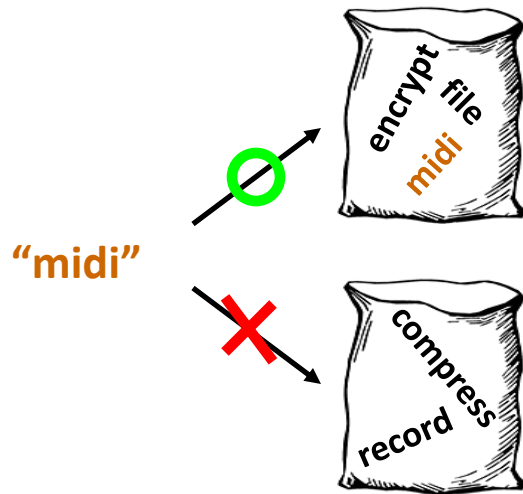
MidiDevice.getReceiver()

MidiEvent.MidiEvent()

Tritonus

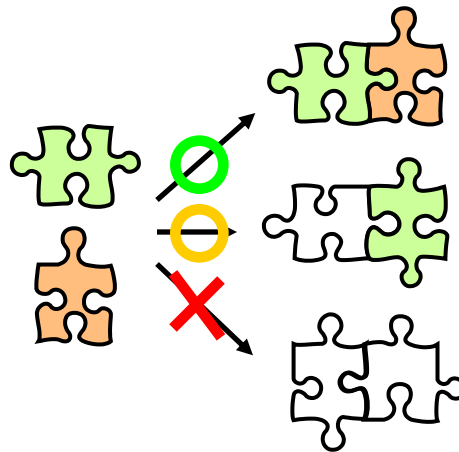
There are three components to compute scores in Exemplar's ranking system.

Word Occurrences (WOS)



Exemplar ranks applications higher when their descriptions contain keywords from the query.

Relevant API Calls (RAS)



An application's RAS score is raised if it makes more calls to relevant methods in the API.

Dataflow Connections (DCS)

“record midi file”

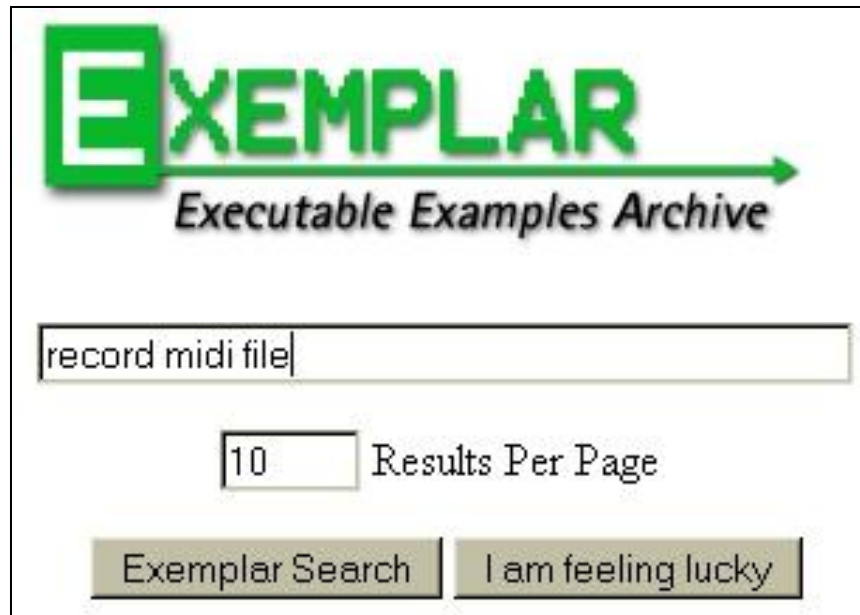
```
String dev = getDevice();  
String buf[] = A.readMidi(msg);  
B.write(buf);
```

If two relevant API calls share data in an application, Exemplar ranks that application higher.

Exemplar Demonstration

<http://www.xemplar.org/>

The user enters a high-level query.



The image shows a screenshot of the Exemplar website's search interface. At the top left is the logo, which consists of a green square containing a white stylized 'E' followed by the word 'XEMPLAR' in green, with a green arrow pointing to the right underneath. Below the logo is the text 'Executable Examples Archive'. Below this is a search input field containing the text 'record midi file'. Underneath the search field is a dropdown menu showing '10' and the text 'Results Per Page'. At the bottom are two buttons: 'Exemplar Search' and 'I am feeling lucky'.

<http://www.exemplar.org/>

The search returns a list of projects, their descriptions, and their scores.

Project Name	Relevance Score	Description
MidiQuickFix	100%::45.59%	MidiQuickFix allows you to directly edit the events in a Midi file. It is intended to make it easy to find and fix problems, such as setting volume and pan values for a track, without the need for a complex Midi sequencing program.
Saiph	100%::30.71%	Java-based (multiplatform) tool for algorithmic musical composition. Saiph generates sequences made of tracks made of segments with musical events, currently notes and MIDI controllers. It supports MIDI and MusicXML file output.
PJLMidiParser	100%::0%	PJLMidiParser provides efficient parsers, written in Java, for MIDI files. It is like XML SAX parsers in that it is event-driven; the parsing is initiated and then triggers callback handlers in response to events in the MIDI file.
Tritonus	0%::100%	Tritonus is an independent implementation of the Java Sound API (http://www.javasoft.com/products/java-media/sound/index.html).
TuxGuitar	0%::82.8%	TuxGuitar is a multitrack guitar tablature editor and player written in Java-SWT, It can open GuitarPro, PowerTab and TablEdit files.

The programmer can view a list of API calls and their locations within projects.

File	Line No	API Used
0.3.0/tritonus-0.3.0.tar.gz/tritonus-0.3.0/src/javax/sound/midi/MetaMessage.java	93	javax::sound::midi::MetaMessage::MetaMessage
0.3.0/tritonus-0.3.0.tar.gz/tritonus-0.3.0/src/javax/sound/midi/Track.java	54	javax::sound::midi::MidiEvent::getTick
0.3.0/tritonus-0.3.0.tar.gz/tritonus-0.3.0/src/javax/sound/midi/Track.java	54	javax::sound::midi::MidiEvent::MidiEvent
0.3.0/tritonus-0.3.0.tar.gz/tritonus-0.3.0/src/javax/sound/midi/Track.java	98	javax::sound::midi::MidiEvent::getTick
0.3.0/tritonus-0.3.0.tar.gz/tritonus-0.3.0/src/javax/sound/midi/Sequence.java	79	javax::sound::midi::Sequence::createTrack
0.3.0/tritonus-0.3.0.tar.gz/tritonus-0.3.0/src/javax/sound/midi/MidiSystem.java	60	javax::sound::midi::MidiSystem::getMidiDeviceProviders

Exemplar: EXEcutable exaMPLes ARchive

For more details, please visit our technical paper presentation on Friday, May 7th at 2:00pm in room 1.4 (1/2):

M. Grechanik, C. Fu, Q. Xie, C. McMillan, D. Poshyvanyk, and C. Cumby, "A Search Engine For Finding Highly Relevant Applications," *Proc. of 32nd ACM/IEEE International Conference on Software Engineering*, p. 10, May 2-8 2010.

<http://www.xemplar.org/>

