

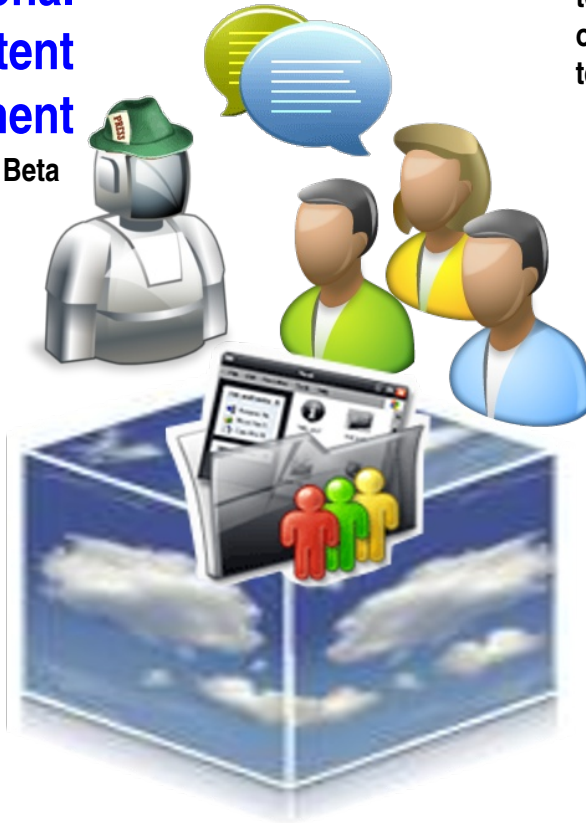
The Making Of...

## Conversational Content Management

The Beta

Learn how this software was put together using nine open source technologies.

With Conversational Content Management, a chat reporter listens in on your team's discussion. When you are done, the reporter interviews you for any remaining information then publishes a summary of the actionable content from the discussion.



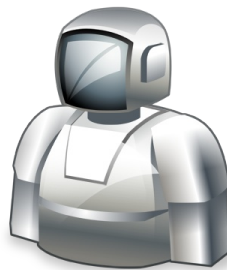
A chat based collaborative document creation experience is more intuitive as it leverages our naturally social characteristics. We are much more likely to enter in to a real time group discussion than we are to fill out a form or post to a forum.

## Under the Covers with the Chat Reporter

Nothing is more clarifying than an example. In this document, we introduce the internals to Conversational Content Management by taking you through a sample conversation with the chat reporter. In this conversation, you see the chat interaction followed with how the chat session state is resolved after each message is sent.

```
[11:52:36] User: testchat1@muc.ec2-107-20-72-176.compute-1.amazonaws.com/glenntest
<Tester> Document: who=open source companies big players
Most of the participants IBM Oracle active too
are open Answered: who
source Template: incidentNew Style
companies ChatGroup: testchat1
but big Topic: who
players like Convo: Most of the participants are open source
IBM and companies but big players like IBM and Oracle are
Oracle are active too.
active too. Title:
Mode: 2
```

## Interview with a Chat Reporter



## Part 1 Templates and Attributes

User: testchat1@muc.ec2-107-20-72-176.compute-1.amazonaws.com/glenntest  
Document: why=IT more savvy prefers commodity licensing |who=open source companies big players IBM Oracle active too  
Answered: who|why  
Template: incident  
ChatGroup: testchat1  
Topic: why  
Convo: Most of the participants are open source companies but big players like IBM and Oracle are active too.|The reason is that IT is more savvy and prefers commodity licensing.  
Title:  
Mode: 2

[11:54:07]  
<Tester> The reason is that IT is more savvy and prefers commodity licensing.

The secret to summarizing actionable content out of a social discussion is that the chat reporter models the discussion as a document type whose template contains relevant attributes. As messages continue to be posted to the chat room, the reporter is keeping an eye out for the document type and its attributes.

Certain special words (i.e. synonyms) key the reporter to associate the relevant part of the discussion to the corresponding attribute of the template.

[11:54:56]  
<Tester> There is a growing trend to use open source software for big data is the conclusion.

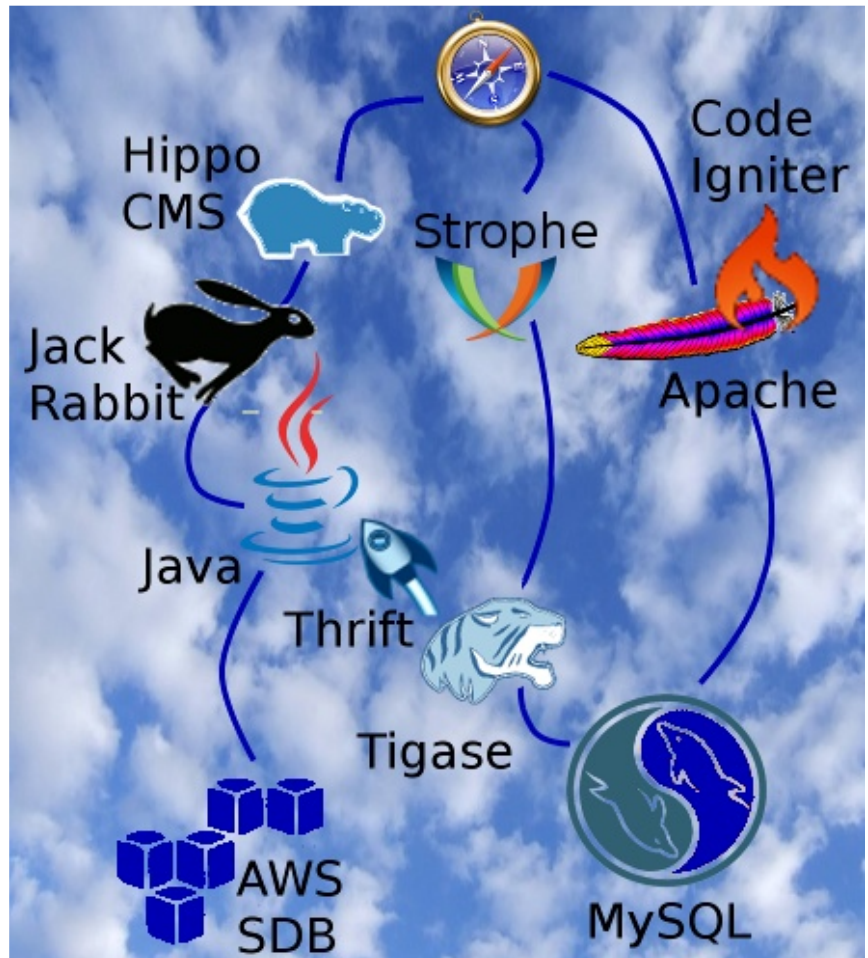
User: testchat1@muc.ec2-107-20-72-176.compute-1.amazonaws.com/glenntest  
Answered: who|why|what  
Document: what=growing trend to use open source software big data |why=IT more savvy prefers commodity licensing |who=open source companies big players IBM Oracle active too  
Template: incident  
Topic: what  
ChatGroup: testchat1  
Convo: Most of the participants are open source companies but big players like IBM and Oracle are active too.|The reason is that IT is more savvy and prefers commodity licensing.|There is a growing trend to use open source software for big data is the conclusion.  
Title:  
Mode: 2

## The Anatomy Of Conversational Content Management

We'll return to the chat reporter in a minute but first, let's take a tour of the inner workings of this application.

When you register and validate your email, you are using a PHP app written with Code Igniter. Your registration information is stored in a MySQL database

When your group is ready, they join the discussion. This causes the web app to connect to the Tigase XMPP server via Strophe and communicate with each other via Multi-User Chat. The chat reporter also joins the chat room.



Tigase uses MySQL to store all of its information but the chat reporter communicates with a middle tier component via Apache Thrift to maintain the state of each of its team discussions. That component maintains the chat session state in Amazon Simple DB.

Once the team believes that the conversation is over, someone asks the chat reporter to publish the document. The chat reporter interviews the group for any remaining information that it needs in order to complete the document. After the chat reporter has all the information that it needs, it creates the document in a JCR compliant repository implemented by the Apache Jack Rabbit open source software.

This repository is surfaced on the Internet for public access by another open source project called Hippo CMS.

Hippo, Jack Rabbit, the chat session manager, Thrift, and Tigase are all written in the Java programming language. The web application is written in PHP and uses the Code Igniter framework. Strophe runs in the web browser and is written in Java Script.

The Java programs run in the OpenJDK virtual machine. All server side components run in Amazon's EC2 environment.

## Interview with a Chat Reporter

### Part 2 Document Type Transformation

User: testchat1@muc.ec2-107-20-72-176.compute-1.amazonaws.com/glenntest  
Document: which=Hadoop R Mondrian Jaspersoft |what=growing trend to use open source software big data |why=IT more savvy prefers commodity licensing |who=open source companies big players IBM Oracle active too  
Answered: who|why|what|which  
Template: incident  
ChatGroup: testchat1  
Topic: which  
Convo: Most of the participants are open source companies but big players like IBM and Oracle are active too.|The reason is that IT is more savvy and prefers commodity licensing.|There is a growing trend to use open source software for big data is the conclusion.|The details are Hadoop, R, Mondrian, and Jaspersoft.  
Title:  
Mode: 2

[11:56:56]  
<Tester> The details are Hadoop, R, Mondrian, and Jaspersoft.

[11:57:20] <Tester> This is a project.

User: testchat1@muc.ec2-107-20-72-176.compute-1.amazonaws.com/glenntest  
Document: status=IT more savvy prefers commodity licensing |goal=Hadoop R Mondrian Jaspersoft |name=growing trend to use open source software big data  
Answered: status|name|goal  
Template: project  
ChatGroup: testchat1  
Topic: goal  
Convo: Most of the participants are open source companies but big players like IBM and Oracle are active too.|The reason is that IT is more savvy and prefers commodity licensing.|There is a growing trend to use open source software for big data is the conclusion.|The details are Hadoop, R, Mondrian, and Jaspersoft.|This is a project.  
Title:  
Mode: 2

Identifying a different document type queues the chat reporter for a one time conversion to a different template with different attributes.

What follows next is the human asks the reporter to publish and the reporter interviews the human about the remaining attributes to the underlying template that the reporter has not encountered previously in this discussion.

[11:57:58] <Tester> Publish this reporter.

[11:57:58] <reporter> How would you introduce this project to someone new?

User: testchat1@muc.ec2-107-20-72-176.compute-

1.amazonaws.com/glenntest

Document: status=IT more savvy prefers commodity licensing  
|goal=Hadoop R Mondrian Jaspersoft |name=growing trend to use open  
source software big data

Answered: status|name|goal

Template: project

ChatGroup: testchat1

Topic: intro

Convo: Most of the participants are open source companies but big  
players like IBM and Oracle are active too.|The reason is that IT  
is more savvy and prefers commodity licensing.|There is a growing  
trend to use open source software for big data is the  
conclusion.|The details are Hadoop, R, Mondrian, and  
Jaspersoft.|This is a project.

Title:

Mode: 3

[11:58:34] <Tester> Get to big data faster and cheaper.

[11:58:35] <reporter> Who does this project serve?

User: testchat1@muc.ec2-107-20-72-176.compute-

1.amazonaws.com/glenntest

Document: status=IT more savvy prefers commodity licensing  
|goal=Hadoop R Mondrian Jaspersoft |intro=Get to big data faster  
and cheaper.|name=growing trend to use open source software big  
data

Answered: status|name|goal|intro

Template: project

ChatGroup: testchat1

Topic: audience

Convo: Most of the participants are open source companies but big  
players like IBM and Oracle are active too.|The reason is that IT  
is more savvy and prefers commodity licensing.|There is a growing  
trend to use open source software for big data is the  
conclusion.|The details are Hadoop, R, Mondrian, and  
Jaspersoft.|This is a project.|Get to big data faster and cheaper.

Title:

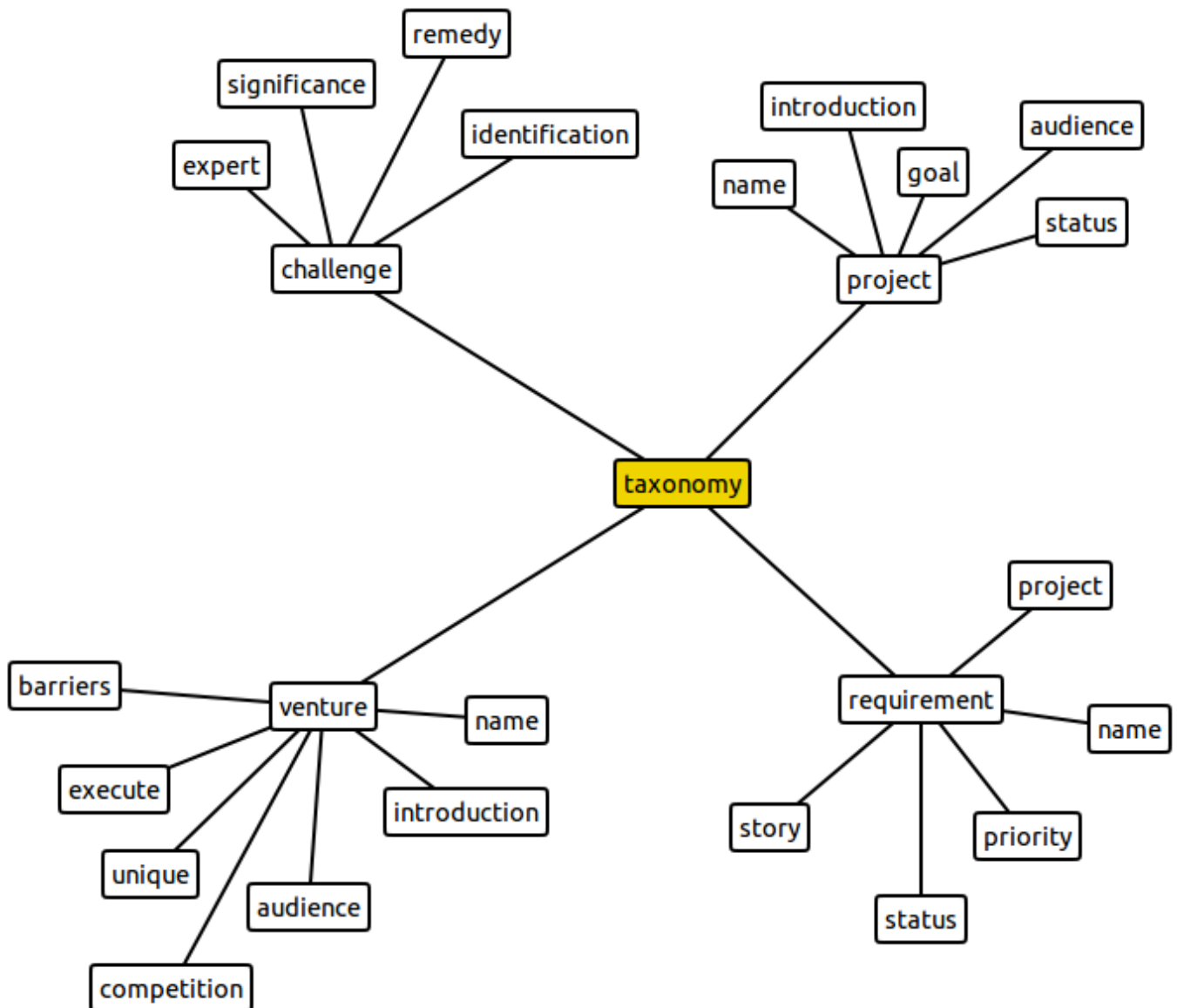
Mode: 3

[11:58:49] <Tester> IT professionals

[11:58:50] <reporter> Your document has been created and is ready  
for viewing.

# Conversational Content Management Template and Attribute Taxonomy

Here is the current taxonomy that the chat reporter knows about. When you start a discussion, the chat reporter defaults to the more generic template of incident (not shown) which has the following attributes; why, what, where, who, which, when, and how.



About the Author: <http://www.linkedin.com/in/gengstrand>



Glenn Engstrand has served in the roles of engineer, manager, teacher, and entrepreneur. His current interests include real-time communications, big data, and search engine technology.

