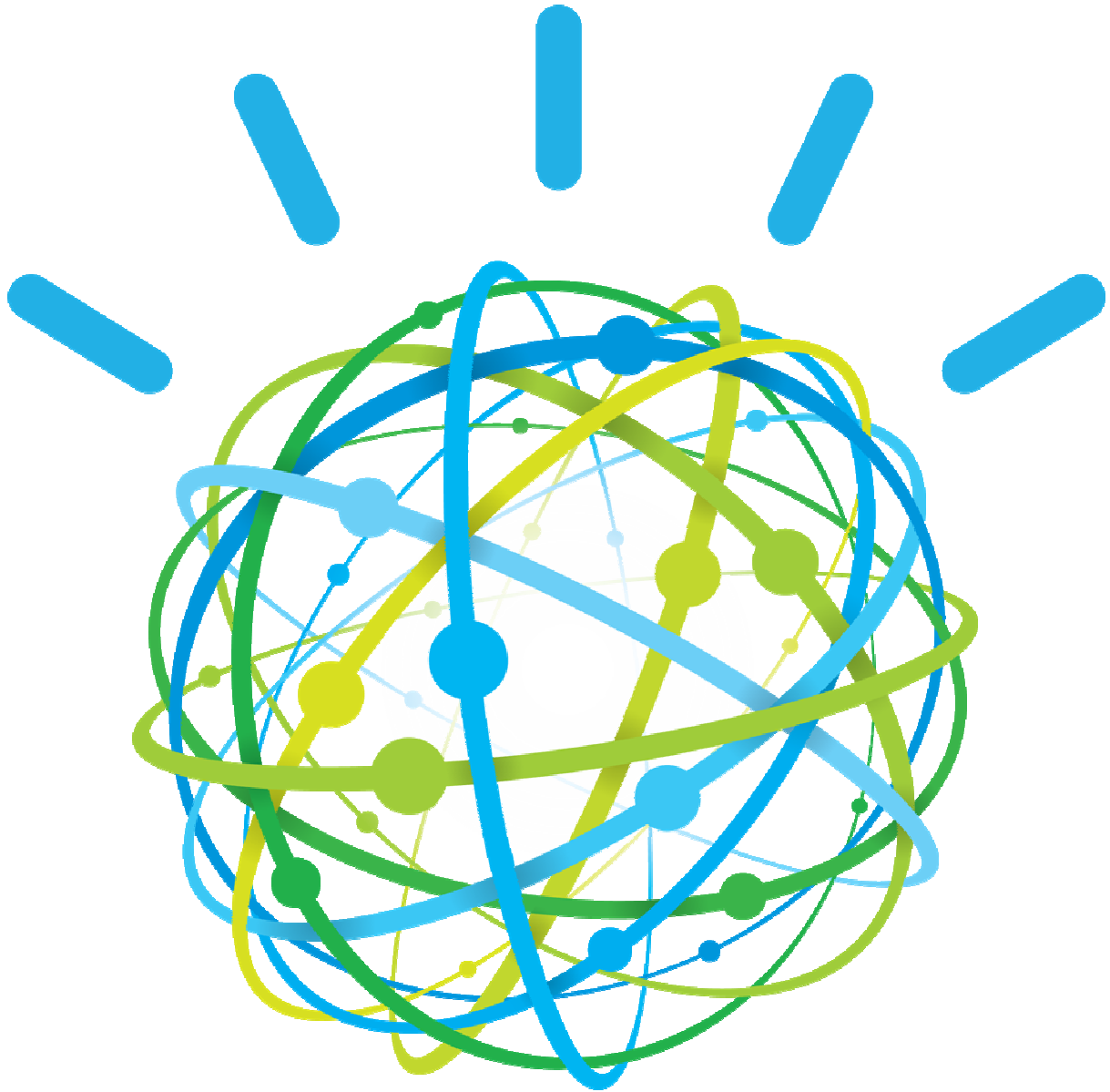


IBM Watson Ecosystem

Getting Started Guide



Version 1.1 – July 2014

Table of Contents:

I. [Prefix](#)

[Overview](#)

II. [Getting Started](#)

A. [Prerequisite Learning](#)

III. [Watson Experience Manager](#)

A. [Assign User Roles](#)

B. [Manage Corpus](#)

C. [Train Watson](#)

D. [Configure Watson](#)

E. [Test Watson](#)

IV. [Sample Reference Mobile Applications](#)

A. [Watson Photography App](#)

B. [Watson Question Answer API](#)

V. [Frequently Asked Questions](#)

VI. [Glossary](#)

[Acronym Definition](#)

I. Prefix

Overview

As a partner in the IBM Watson Ecosystem, you will have a unique opportunity to build a one of a kind application using the cognitive computing power of IBM Watson!

This document will introduce you to the tools and concepts you'll need to build your *powered-by-Watson* application.

To begin, you'll learn about The IBM Watson Experience Manager web portal. This portal allows you to upload and test your data on the IBM Watson instance assigned to your team.

Then, we'll cover the RESTful service API that you will use to connect to your data and provide you with the source code and sample applications for: a hybrid mobile, Android native and iOS native reference application.

Let's get started!

II. Getting Started

A. Prerequisite Learning

It is highly recommended that you view the following videos before starting to use Watson and the Experience Manager.

- Watch the “IBM Watson Experience Manager Overview” tutorial which illustrates how to use IBM Watson Experience Manager for: (i) uploading content to create a ‘Corpus’, (ii) testing Watson by asking questions and exploring answers, (iii) configuring Watson to change the look and feel of Watson, and (iv) training Watson by providing a set of question-answer pairs.

<https://www.youtube.com/watch?v=aBYSEJ-cu1w>

- Watson tutorial videos:
 - Module #1: What is IBM Watson? This video presents an overview of cognitive computing, IBM Watson and IBM's strategy to proliferate cognitive computing in the marketplace with Watson.
<https://ibm.biz/watson-tutorial-overview>
 - Module #2: What is the IBM Watson Ecosystem? This video describes the IBM Watson Ecosystem, how it enables partners to develop cognitive

computing applications, and the different types of partners within the ecosystem (Application Partners, Content Partners, Talent Partners). It also discusses how to get started, and what partner companies get when they become part of the IBM Watson Ecosystem.

<https://ibm.biz/watson-tutorial-ecosystem>

- Module #3: Roadmap for Building Your "Powered by Watson" Cognitive Application. This video outlines the roadmap (including key activities and milestones) for planning your IBM Watson application including the Engage, Prototype, Build and Deploy phases.
<https://ibm.biz/watson-tutorial-plan>
- Module #4-1: Designing and Developing Your "Powered by Watson" Cognitive Application. This video covers the process of building a "Powered by Watson" cognitive application, describes the Watson Experience Manager, (which is a browser based tool to access Watson) and touches on the kinds of content that are accepted by Watson.
<https://ibm.biz/watson-tutorial-build>
- Module #4-2: Enrich Your "Powered by Watson" Cognitive Application With Content. This video focuses on how to enrich Watson applications with most relevant content and describes how to upload and manage content using Watson Experience Manager.
<https://ibm.biz/watson-tutorial-content>
- Module #4-3: Train your "Powered by Watson" Cognitive Application. This video discusses training of Watson and illustrates one way to assist in training Watson by providing a set of question-answer pairs.
<https://ibm.biz/watson-tutorial-train>
- Module #4-4: Test and Deploy your "Powered by Watson" Cognitive Application. This video describes how to test the Watson application using Watson Experience Manager and highlights various considerations to think about when deploying the application including the user base, rollout plan, security provisions, and promoting the application to the production zone in the Watson Developer Cloud.
<https://ibm.biz/watson-tutorial-testanddeploy>

More videos will be posted on the IBM Watson Ecosystem YouTube channel as they become available: <https://ibm.biz/watson-ecosystem-youtube>

For more background material, please consult the IBM developerWorks Watson community: <https://developer.ibm.com/watson/>

III. Watson Experience Manager -- Connecting to and Configuring Your Watson Sandbox

Watson Experience Manager (WEM) is a browser-based tool that allows you to interact with your Watson instance (delivered via the Watson Developer Cloud) while developing your *powered-by-Watson* app.

To start, you will use WEM to assign user roles to users. Next, you will use WEM to upload your application specific content and create a corpus. Once you have created a corpus, you can use the WEM UI test window to pass test questions to Watson and retrieve answers. The main objective of this early testing is to verify that the returned answers come from the uploaded documents thus confirming that the content has been properly ingested by Watson. Please note that at this point, no content specific training has been done and as such, you may not get back the exact answer you are expecting as the first answer.

To access your unique instance, replace the *<your-unique-instance-number>* in the URL links below with the Watson instance number that was provided to you in the Getting Started email.

The Watson Experience Manager has a comprehensive set of documentation that is hosted on your sandbox environment: <https://watson-wdc01.ihost.com/instance/<your-unique-instance-number>/predeploy/index.jsp> but the following instructions are tailored to the process you should follow during your prototype building phase.

To access your Watson instance via WEM, access the following URL from a web browser: <https://watson-wdc01.ihost.com/instance/<your-unique-instance-number>/predeploy/WatsonExperienceManager#>

When prompted, enter your assigned username and password. Once authenticated into WEM, you will be redirected to the homepage as shown in Figure 1:



Figure 1: Watson Experience Manager Home Page (screenshot from version of the tools as of date of this document)

Watson Experience Manager enables you to execute the following steps:

A. Assign User Roles – To assign roles for different users

There are several well defined personas which represent different user roles and contexts in a Watson solution. Start by first understanding the different roles and identifying the people who can fill those roles in developing the *powered-by-Watson* application.

From WEM homepage, use the home button to access the System Administration tool as shown in Figure 2.

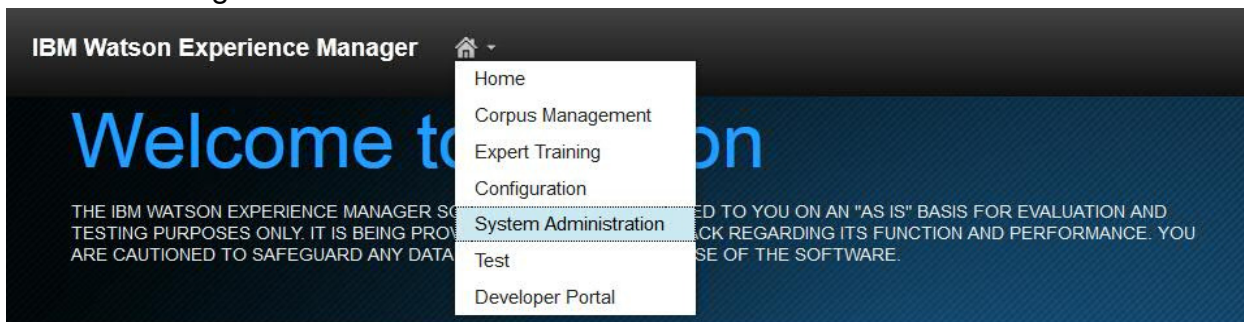


Figure 2: System Administration Access (screenshot from version of the tools as of date of this document)

Under System Administration → User Role Administration, you can assign user roles to users by clicking “New User” button as shown in Figure 3. **Please note that you can**

only assign roles to Users that have been already created by IBM and communicated to you in your Getting Started email.

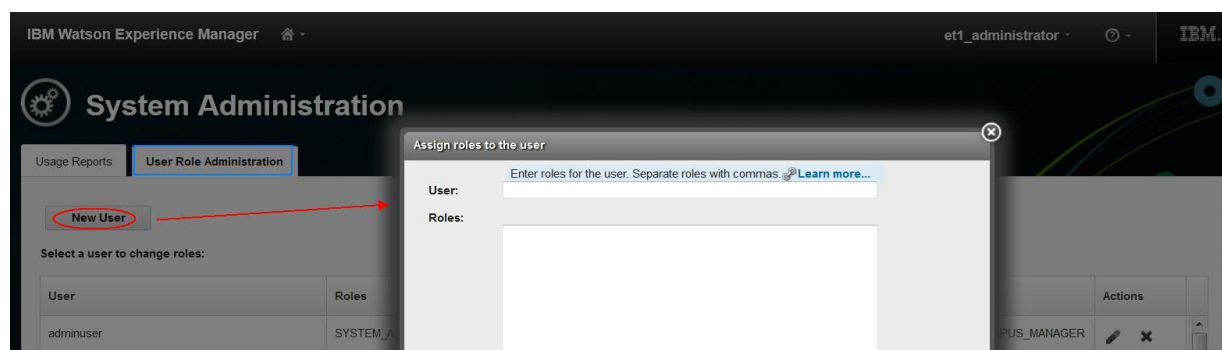


Figure 3: System Administration New User Role Assignment (screenshot from version of the tools as of date of this document)

The available user roles and associated actions they can execute are outlined in Table I. A user can have multiple roles and can access only the tools and areas that apply to the roles they are assigned. For example, you might have a subject matter expert who identifies and uploads documents from which Watson finds answers. This same person might also help create training questions and answers but not approve the questions and answers. In this example, assign the user the roles of GTT_AUTHOR and CORPUS_UPLOADER.

Table I: User Roles and Actions

User Role	Actions
CORPUS_MANAGER	<ul style="list-style-type: none"> ✓ Create a corpus and deploy it to Watson ✓ View information and reports about a version of the corpus ✓ Upload and delete content ✓ Test Watson
CORPUS_UPLOADER	<ul style="list-style-type: none"> ✓ Upload and Delete Content in Corpus Management tool
CUSTOMER_ADMIN	<ul style="list-style-type: none"> ✓ Access Configuration tool and Customize Watson ✓ Assign User Roles in System Administration tool ✓ Generate usage reports in System Administration tool ✓ Test Watson
GTT_APPROVER	<ul style="list-style-type: none"> ✓ Review, approve, and reject questions and associated answers in Expert Training tool
GTT_APPROVER_EXT	<ul style="list-style-type: none"> ✓ Reassign questions to different users in Expert Training tool
GTT_AUTHOR	<ul style="list-style-type: none"> ✓ Create and edit questions in Expert Training tool
GTT_END_USER	<ul style="list-style-type: none"> ✓ Create questions in Question Input (users assigned this role can access only this tool even if they have other roles assigned to their user id)

B. Manage Corpus – To Upload Content and Create a Corpus

The first step you need to take to configure your Watson instance is to upload your content. It is very important that this content should be relevant to your application domain and represents the data set you would like Watson to use when providing answers to questions. Valid content can be in any of the following supported formats: ***.doc**, ***.docx**, ***.pdf**, ***.html**, ***htm**, ***mht** and ***.zip**. Please note that nested folders are not supported and the ***.zip** can only include files of the other valid formats specified above. Figure 4 illustrates guidelines for good content which includes documents in the valid formats listed above with well formed semantics (titles, sections, headings, etc.). Enriching Watson’s dictionaries with glossary and abbreviations relevant to your application domain help improve Watson’s responses.

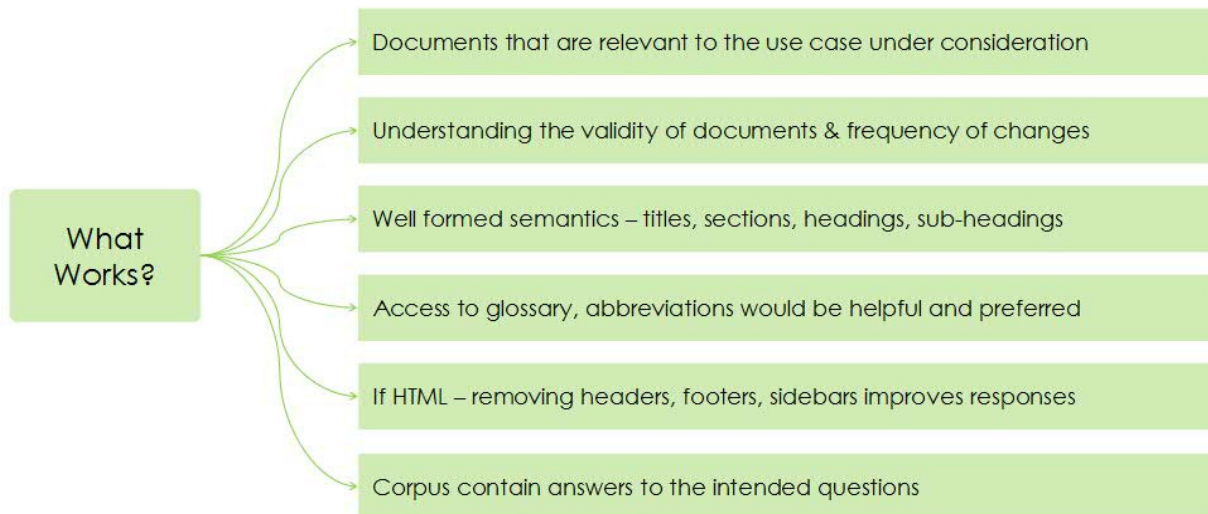


Figure 4: Guidelines for good content

Please note that at this time, Watson does not understand graphics (images, videos, sounds, ...), non-English documents, or data with complex relationships such as nested tabular data, call data records, sensor logs, meter data, etc. Figure 5 illustrates examples of content that is “not good” such as scanned documents (OCR), charts, password-protected documents, noisy chat logs, nested tables and image-rich documents with no description. To reiterate, documents whose text content is not relevant to the application domain or don’t include answers to end users’ questions would not be very useful for Watson.

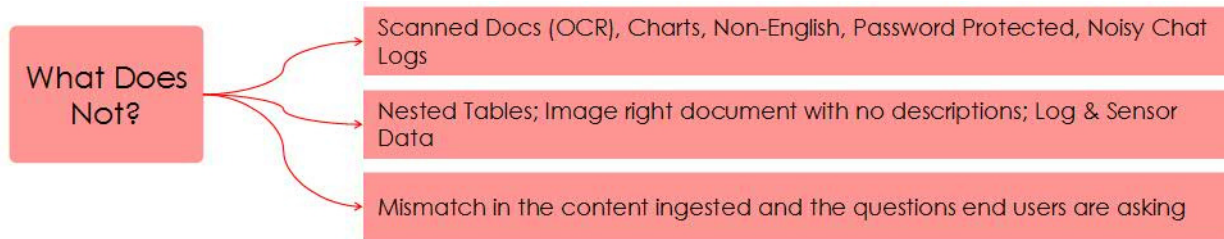


Figure 5: Guidelines for content that is “not good”

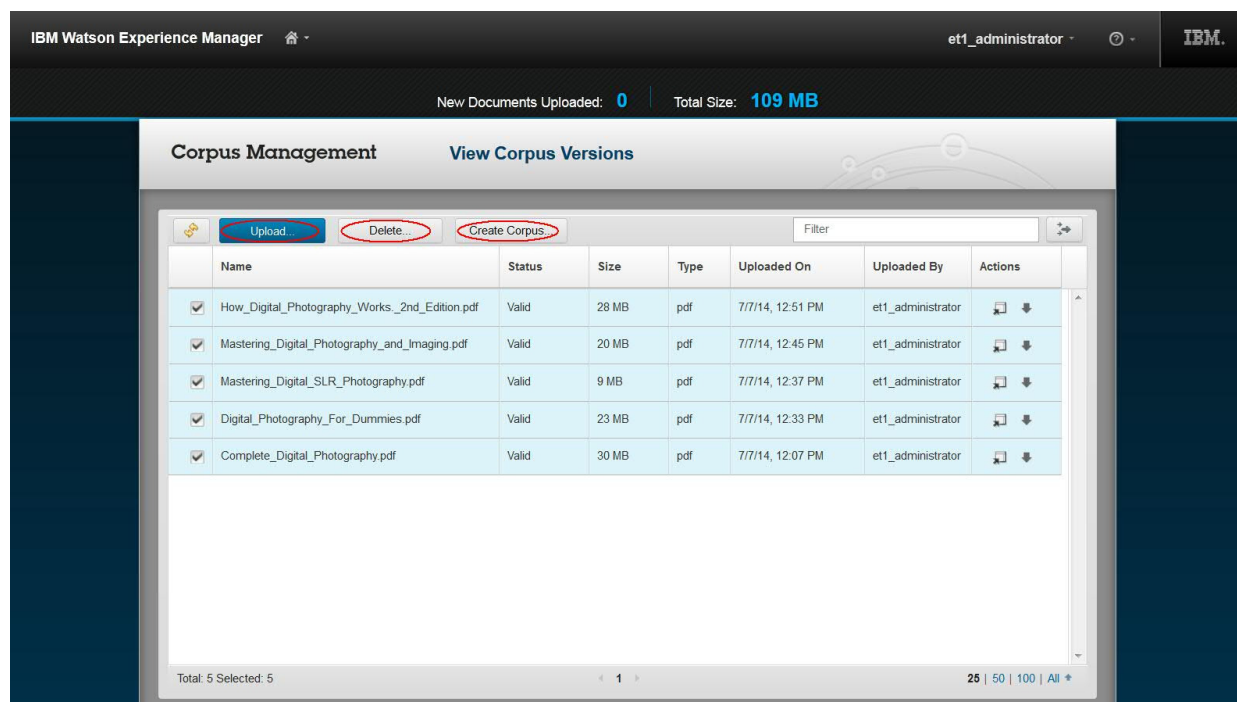


Figure 6: Corpus Management (screenshot from version of the tools as of date of this document)

- **Upload:** As shown in Figure 6, use this button to upload selected files that represent relevant content to the Watson instance.
 - When you click the **Upload** button a file browser window will allow you to select files individually. If you would like to upload multiple files at one time, create a zip archive of those files and select the '.zip' file to upload.
 - Once the files have been uploaded, Watson begins analyzing and ingesting the content. You can upload multiple files concurrently.
- **Delete:** Select one or more uploaded documents and use the **Delete** button to remove the selected documents.

After you have uploaded all of the content it's time to create your corpus

- Use the “**Create Corpus**” button to create and deploy the corpus, which will be used by Watson for finding the best answers to the questions.

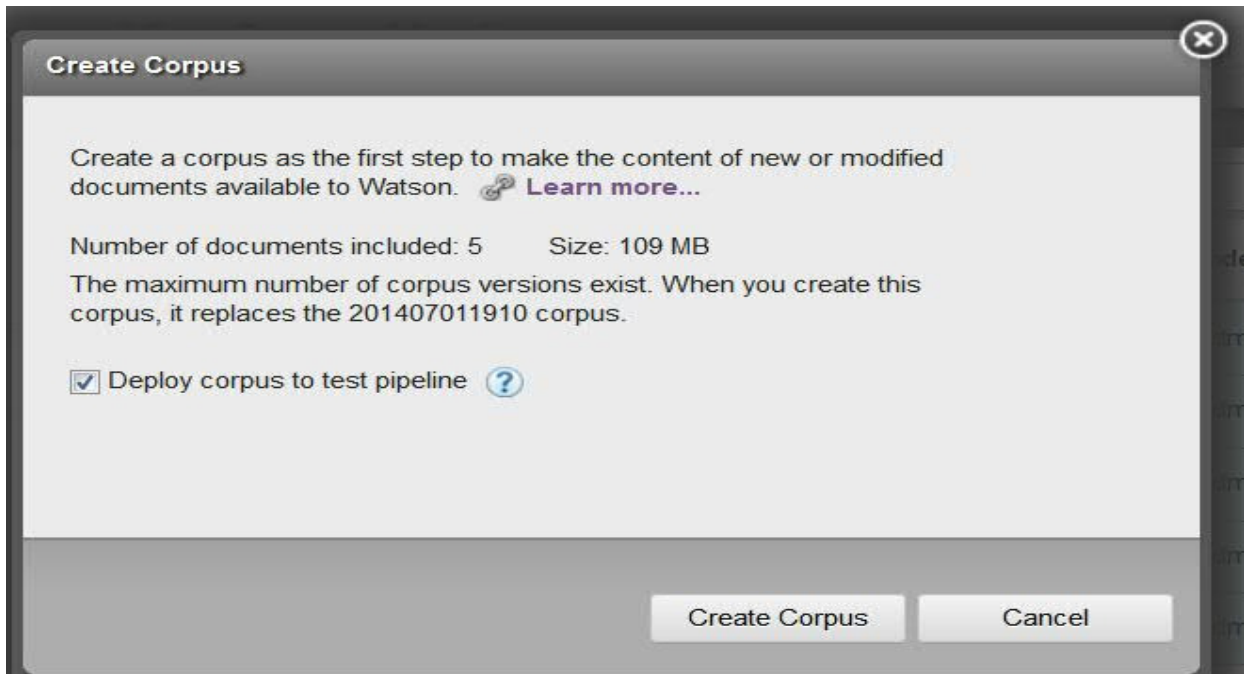


Figure 7: Create Corpus pop-up (screenshot from version of the tools as of date of this document)

- After pressing “**Create Corpus**” you will be presented with the popup shown in Figure 7. Make sure the “Deploy corpus to test server” checkbox is checked, then click on the “**Create Corpus**” button in the popup.
- Watson is now analyzing your documents and creating your corpus. Depending on the amount of data, this could take up to a few hours.

C. Train Watson

This activity involves defining “ground-truth”, or a set of question-answer pairs (also referred to as answer key) to train the machine learning models which power Watson’s cognitive engine. In defining these “ground-truth”, keep in mind that both quality and quantity are important metrics.

- **Quality:** It is critical that the “ground-truth” questions are representative of true interaction patterns with Watson. In addition to enlisting the help of subject matter experts (SMEs) to identify representative questions, it is highly recommended to collect questions suggested by potential end users of your application.
- **Quantity:** It is important to have a large enough sample of question-answer pairs for adequate training of the machine learning models. A typical prototype application within the Watson Ecosystem requires roughly 200-800 question-answer pairs. For a production application, it is recommended to have the “ground-truth” consist of roughly 3000-7000 question-answer pairs. While the suggested numbers for QA pairs provide a rough estimate, in practice, the actual number of QA pairs depends on the content size, the interaction patterns, and

the current algorithms being used; and this number could evolve or change over time.

Note that training Watson is an iterative exercise and the accuracy of the responses from Watson improves as more training is performed.

Identifying Answers

To train Watson, users can leverage the Expert Training tool. The tool is accessible from the WEM home page by clicking on “Train Watson”. The Expert Training tool enables users to match questions with answers thus creating the “ground-truth”. **Note that the answer(s) for the representative questions must exist in the uploaded content.**

After entering a representative question as shown in Figure 8, the Expert Training tool enables a user to match answers to that question in one of three ways:

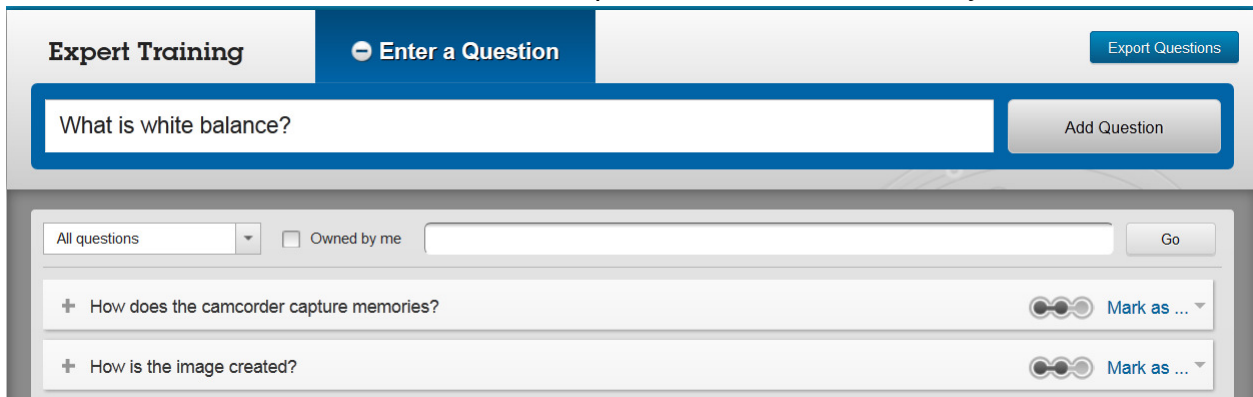


Figure 8: Expert Training Question Entering (screenshot from version of the tools as of date of this document)

- **Match a Similar Question** (pane highlighted with red ellipse in Figure 9): Matching questions create a question cluster. By matching a question to another question-answer pair (matching question is highlighted with green ellipse in Figure 9), you help Watson learn to associate questions and answers that may contain different terms, idioms, phrases and other natural language nuances that are unique to your domain.

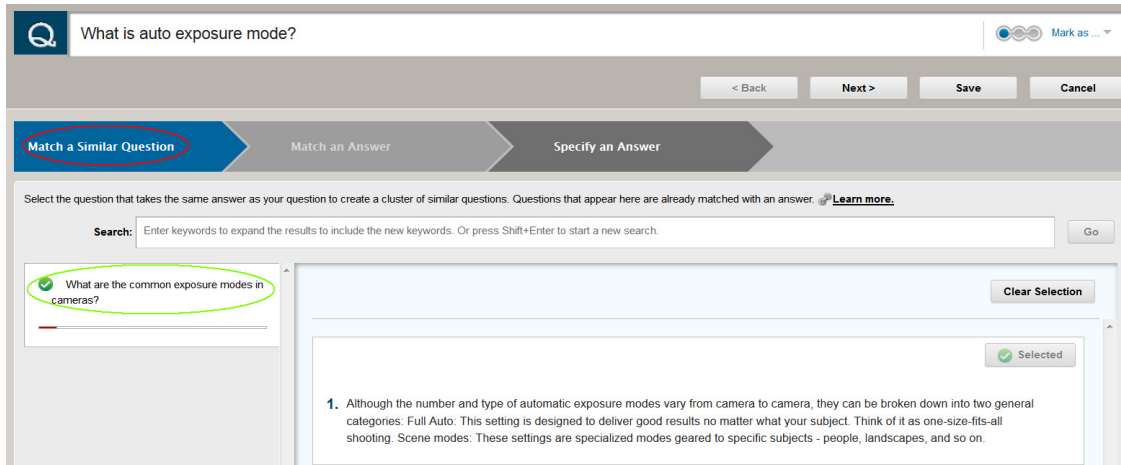


Figure 9: Match a Similar Question pane (screenshot from version of the tools as of date of this document)

- **Match an Answer** (pane highlighted with red ellipse in Figure 10): Although creating clusters of questions is the preferred way to train Watson, you might not find a similar question or you may be just starting your training. In that case, try to find a formatted answer that matches your question. Watson extracts formatted answers from the documents that are uploaded in the corpus management tool. Below each answer or document in the list is a bar (highlighted with green ellipse in Figure 10) that indicates how relevant the answer is to your original question. The longer the bar the more relevant the result is (as measured by Watson). You can restrict the answer source to a single document by entering a file name and/or you can include keywords to help locate a particular answer.

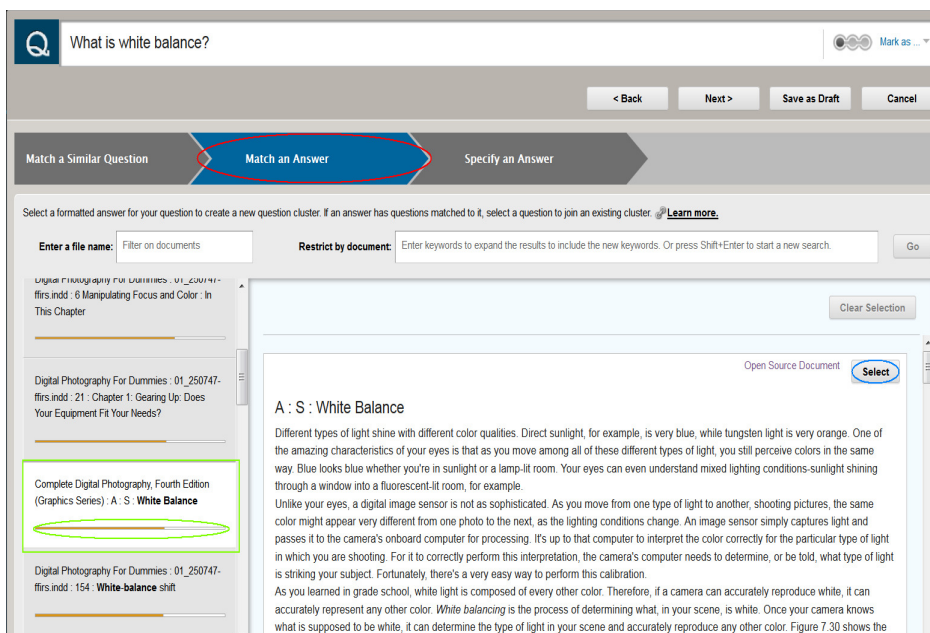


Figure 10: Match an Answer pane (screenshot from version of the tools as of date of this document)

- **Specify an Answer** (pane highlighted with red ellipse in Figure 11): If you can't match the question with a formatted answer or a similar question, you can specify answer passages. You specify an answer passage in the “Specify an Answer” pane by highlighting the section of text that contains the best answer to your question (highlighted with green ellipse in Figure 11). If you find more than one appropriate answer, specify separate answer passages.

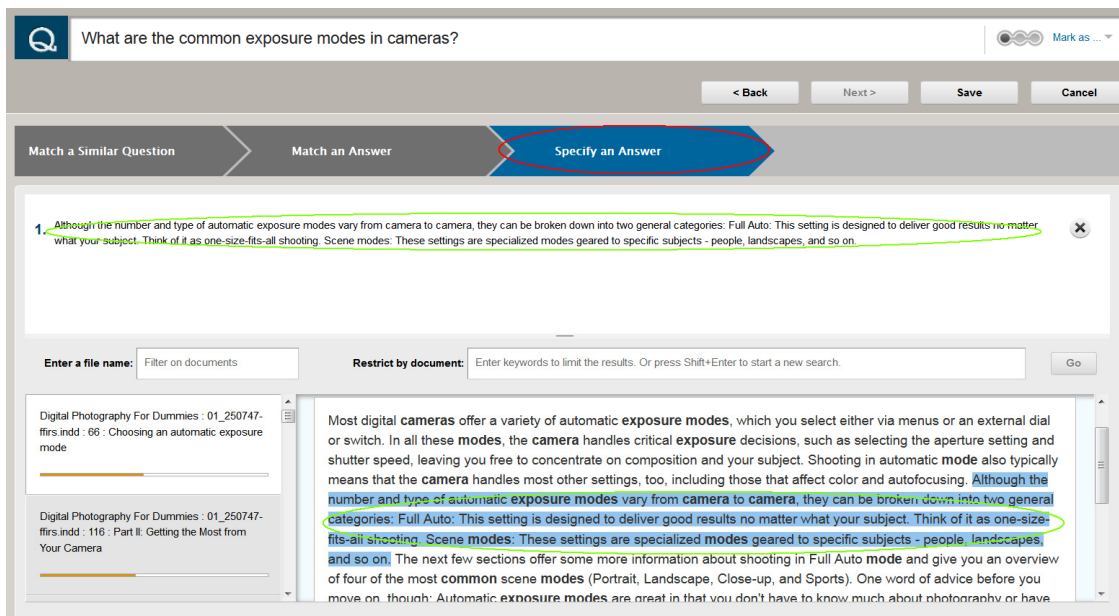


Figure 11: Specify an Answer pane (screenshot from version of the tools as of date of this document)

Matching a similar question or matching an answer is referred to as **low-impact** training while specifying an answer is part of a training pathway called **high-impact** training.

For low-impact training, the training process involves entering the question(s), identifying the correct answer(s), verifying and approving the answer(s). When all QA pairs are approved, the corpus needs to be re-created (go back to Corpus Management and Create Corpus as explained earlier) and then the training will have an effect on Watson's responses.

High-impact training, on the other hand, requires an additional set of offline steps before the training can have an effect on Watson's responses. The QA set which forms the “ground-truth” is exported and a series of experiments and training runs are performed. These training runs create new machine learning models which are migrated back to the production instance and inserted into the pipeline. Until these steps are completed, specifying an answer will have no effect on Watson's responses.

Working with Questions

The Expert Training tool offers several features to help the users manage the process of creating, reviewing, updating, and approving questions and associated answers.

- Specify one or more of the following options from the question list view to filter the list of questions:
 - Select **Owned by me** to restrict the list of questions that you have added or that were assigned to you.
 - Filter the questions by choosing a condition from the list shown in Figure 12:

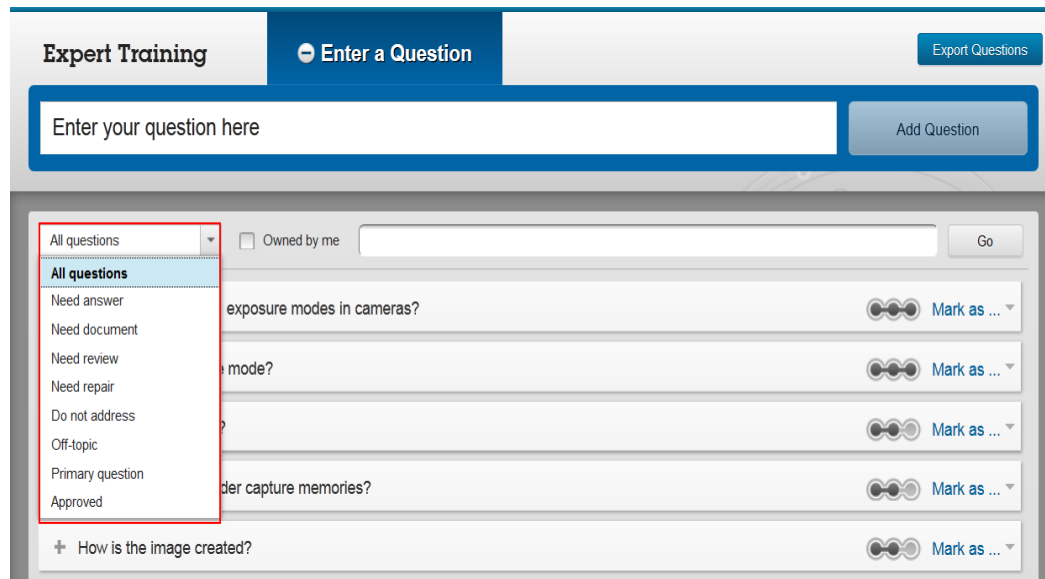


Figure 12: Questions Filter (screenshot from version of the tools as of date of this document)

All questions – All questions that are available in the Expert training tool.

Need answer – Draft question that need answers.

Need document – No answers were found for these questions. Upload more content that has answers to these questions.

Need review – Questions with answers or marked for special handling that are awaiting approval.

Need repair – Questions for which associated answers were rejected. Revise the answers for these questions to get them reviewed again for approval.

Do not address – Questions that are not relevant to the subject matter that Watson is being trained for.

Off-topic – Questions that Watson is restricted from providing answers to. These types of questions include questions that require legal advice or that are better answered by a person.

Primary question – Questions that are not matched to other similar questions. You can match a primary question to an answer, mark it for

special handling, or identify that it has no answer from the current content. A question that is matched to a primary question becomes part of a cluster of similar questions.

Approved – Questions for which the associated answers have been reviewed and approved.

- Restrict the list of question by entering keywords in the filter box that occur in the text of the questions.

- Expand a question by clicking the question text.
- Choose an action for a question (possible actions are highlighted in red boxes in Figure 13 below). Depending on your role and on the condition of the question, you can choose one of the following actions:

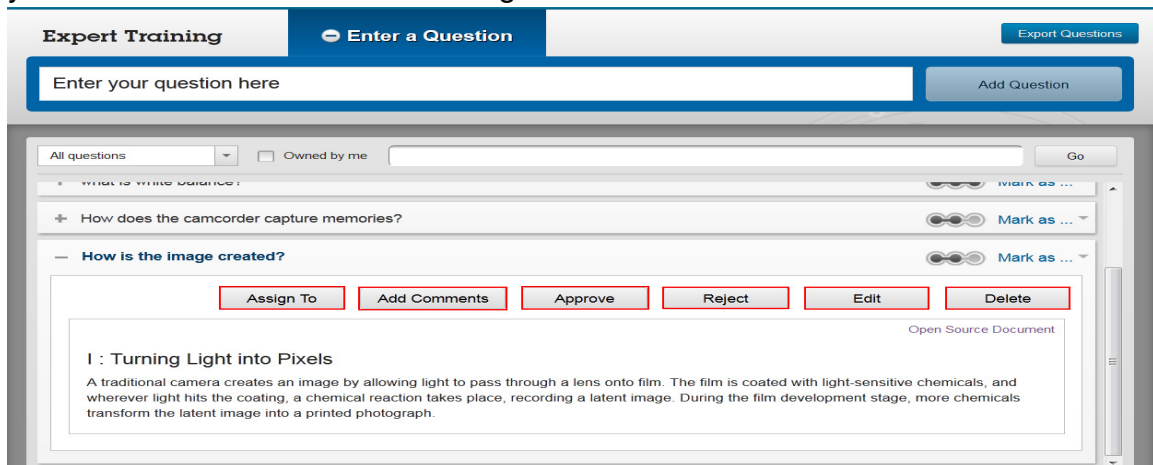


Figure 13: Question Actions (screenshot from version of the tools as of date of this document)

- **Assign To** – to assign a question-answer pair to another user.
- **Add Comments** – to add a comment to a question-answer pair.
- **Edit Comments** – to modify a comment on a question-answer pair.
- **Approve** – to approve a question-answer pair.
- **Reject** – to reject a question-answer pair.
- **Edit** – to edit a question-answer pair.
- **Delete** – to delete a question-answer pair.

Training is a critical step in improving the accuracy of the responses returned by Watson. Training can be less or more involved and can be automatic or semi-automatic where IBM runs experiments and improves machine learning models. Numerous factors influence the quality of training including: (i) correct classification of answers to training questions, (ii) the quality of the uploaded content in generating correct answers, and (iii) the ability of the training question set to represent true interaction patterns with Watson.


Please make sure these requirements are considered throughout the development process; this will maximize the impact and success of your cognitive application.

D. Configure Watson

WEM allows you to generate and customize an iFrame interface to the Watson Question & Answer API (QAAP). In regards to this prototype, this activity is out of scope since you will be interfacing with Watson Question & Answer API directly via a RESTful service connection (described later in this document).

E. Test Watson

WEM provides an interface for you to test your Watson corpus by asking questions and receiving the best answers identified by Watson.

Once your corpus has been loaded, click on the Home button  and select "Test". Try asking a few questions and explore the answers returned by Watson. Watson can only answer questions that have answers in the corpus.

If you are not getting the answers you are expecting, try asking the question phrased differently. For example:

"What is White Balance?"

"What is the definition of White Balance?"

"What does White Balance mean?"

"What are some common uses of White Balancing in photography?"

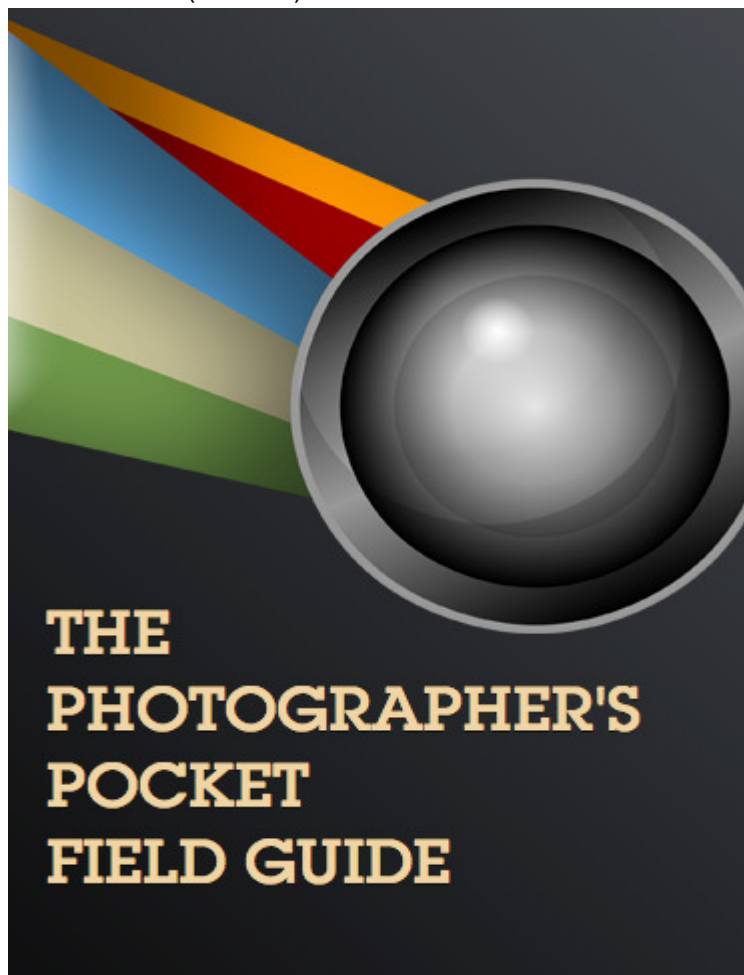
IV. Sample Reference Mobile Applications

In this section, we describe the three sample reference mobile applications that were designed and built by IBM to help illustrate the process of building a *powered-by-Watson* application.

The reference mobile application, “Watson Photography,” has been created for:

- Native Android
- Native iOS
- Hybrid (HTML, JavaScript, CSS, Apache Cordova) using Apache Cordova that can be deployed to both Android and iOS platforms.

The reference applications connect over https to Watson using the Question and Answer API (QAAP).



A. Watson Photography App

The Photographer’s Pocket Field Guide is a mobile reference application that answers users’ questions related to photography.

These reference applications will show you how to connect to your Watson instance, ask questions, and display the results.

Sample code and platform specific documentation can be found on the following public Github repositories:

iOS: <https://github.com/IBMMobileCoC/watson-photography-ios>

Android: <https://github.com/IBMMobileCoC/watson-photography-android>

Hybrid: <https://github.com/IBMMobileCoC/watson-photography-hybrid>

B. Watson Question Answer API

Once your Watson instance is trained to your application domain, it is ready to be integrated into your application. In this guide, we will focus on using the Question Answer API (QAAPI) to interact with Watson. The Watson QAAPI exposes a REST service interface that allows applications to access Watson's cognitive computing capabilities. You can access your API endpoint at the following url: <https://watson-wdc01.ihost.com/instance/<your-unique-instance-number>/deepqa/v1/question>

In what follows, we describe how to use the QAAPI to submit questions and retrieve answers. Although the Watson Question and Answer REST service accepts many parameters, only one parameter, ***questionText***, is required to post a question. The following shows a sample JSON defining the question to post:

```
{
  "question" : {
    "questionText" : "What is White Balance?"
  }
}
```

When you post a question, include an HTTP header. The header must include the content-type. The header must also include a value for X_SyncTimeout since we are using synchronous mode. The X_SyncTimeout value represents how long the server waits after the question is submitted until it times out. The value does not represent how long the client waits for a response from the server. The header should also include the authorization (base64 encoded) required to access your Watson instance.

Accept: application/json

Content-Type: application/json

X-SyncTimeout: 30

Authorization: Basic username:password

The following sample Java code illustrates how to prepare and post the question:

```
// create the question portion
@SuppressWarnings("serial")
HashMap<String, Object> question_info = new HashMap<String, Object>() {{
    put("questionText", "What is aperture?");
}};

// given the collection we created above, serialize it into a JSON object
JSONObject question = new JSONObject().put("question",question_info);

// construct a resource using the REST Client. The URL is the URL of the QA API
Resource resource = restClient.resource(QUESTION_POST_URL);

ClientResponse response = null;

// HTTP header that designates the REST service to be synchronous
// The service will wait 30 seconds after receiving the request and giving it
// to the Watson pipeline. After 30 seconds, the service will return with
// a timeout in the status field. A -1 indicates to wait indefinitely.
resource.header("X-SyncTimeout", "30");

String auth = new String(Base64.encode("username:password".getBytes()));
resource.header("Authorization","Basic " + auth);

// post the question to the Q/A API
// Accept : application/json
// Content-Type : application/json
// HTTP POST
try{
    response = resource.accept(MediaType.APPLICATION_JSON)
        .contentType(MediaType.APPLICATION_JSON) .post(question.toString());

    // check HTTP response code
    if (response.getStatusType() != javax.ws.rs.core.Response.Status.OK) {
        System.out.println("HTTP Response code returned an error:
"+response.getStatusCode());
        System.exit(1);
    }
}
catch(Exception e){
    System.out.println("Unable to connect to server on " + QUESTION_POST_URL);
    System.exit(1);
}
```

After you submit the question, the response includes an HTTP status code. For the synchronous mode, the successful HTTP status code is 200 Created. The successful response includes the original question and the reported answers sorted in decreasing order of confidence. The answers can be obtained from the ***evidencelist*** array where each entry in that array includes the answer text and the confidence of that answer

(highlighted in red rectangles in the example question below). Each evidence item also includes id, title and reference.



The following sample Java code illustrates how to obtain the response from WATSON REST service:

```

// marshal the entity response from the POST into a JSON Object that we can process
JSONObject payloadResponse = response.getEntity(JSONObject.class);

// pull the question from the JSON Object
JSONObject questionResponse = payloadResponse.getJSONObject("question");

// We can also see that the question was Accepted if the 'status' is queued. Other possible
values include ERROR
String status = questionResponse.getString("status");

if (status.equals(STATE_COMPLETE)) {
    // Here we have gotten an answer back and we are going to display some results
    System.out.println("----- Portions of the Answer Payload shown as an example -----");

    //display various Example answer values
    System.out.println("status :"+questionResponse.getString("status"));
    System.out.println("questionText:"+questionResponse.getString("questionText"));

    //loop through the answers displaying some information
  
```

```

JSONArray answers = questionResponse.getJSONArray("evidencelist");
for (int i=0;i < answers.size();i++ ) {
    System.out.println("answer id      "+answers.getJSONObject(i).getString("id"));
    System.out.println("  answer text    "+answers.getJSONObject(i).getString("text"));
    System.out.println("  answer confidence
:"+answers.getJSONObject(i).getString("value"));
}
} else {
    // Question failed or timed out
    System.out.println("Question response returned status: "+status);
    System.exit(1);
}
}

```

Please consult more detailed documentation on the QA API and full Java examples in Watson Experience Manager by going to Home → Developer Portal and then following the API Reference link:

https://watson-wdc01.ihost.com/instance/<your-unique-instance-number>/predeploy/index.jsp?scope=gtthelp&topic=/com.ibm.watson.xmgr.gt.trident.doc/topics/sdk_qaapi_overview.html

V. Frequently Asked Questions

1- Can we delete a corpus that we have previously created?

Yes, you can delete previously created corpora. However, there should be one deployed corpus for the Watson pipeline to be active. For controlling which corpus is deployed, check question 4 below. If you run into any issues deleting a corpus, please post a message to the forum in the community and IBM technical support will delete the corpus for you.

2- Can we delete documents that we have uploaded?

Yes, you can delete uploaded documents. If you run into any issues deleting documents, please post a message to the forum in the community and IBM technical support will delete the documents for you.

3- We have uploaded documents and created corpus already but we need to add new documents, what should we do?

You can delete previously uploaded documents and/or you can upload new documents. Once you have uploaded the relevant documents, you can create a new corpus. Figure 14 shows that under “View Corpus Versions” tab, you can see all the created corpora and which is the one currently deployed.

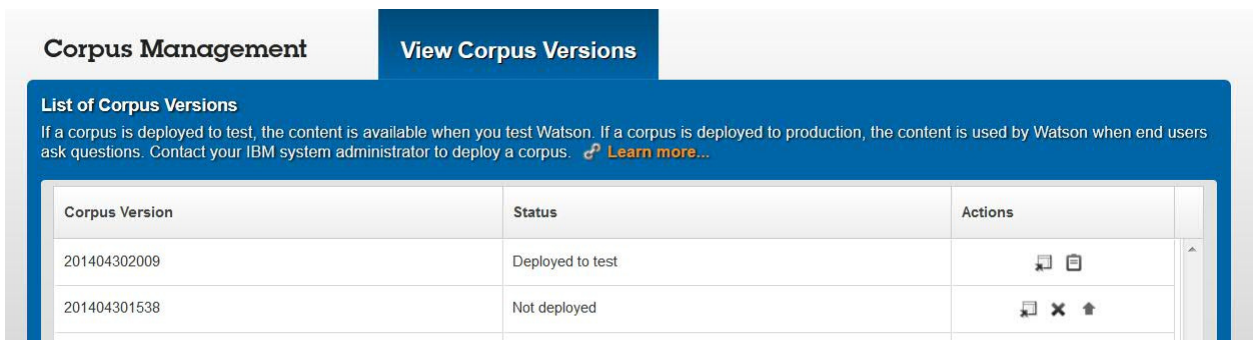


Figure 14: Corpus Versions (screenshot from version of the tools as of date of this document)

4- We have created multiple corpora, can we choose which one gets deployed?
 Yes, once you have created multiple corpora, you can choose which corpus to deploy to test by selecting the “Deploy” button under Actions column under the “View Corpus Versions” tab as shown in the Figure 15.

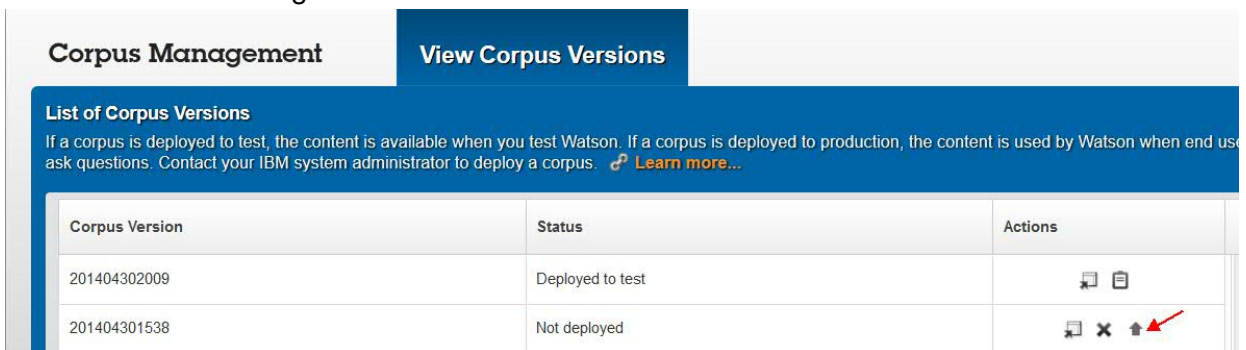


Figure 15: Deployed Corpus (screenshot from version of the tools as of date of this document)

5- Is there a limit on the number of corpora we can create?

Yes, there is a limit of 4 corpora that you can create.

6- During this prototype phase, will using the “Useful” or “Not Useful” buttons in response to Watson’s answers have any effect?

No, using “Useful” or “Not Useful” buttons will not have any effect on Watson in this prototype phase.

VI. Glossary:

Acronym Definition

Corpus – Collection of documents representing the content uploaded to Watson.

QAAPI: Question Answer Application Programming Interface

JSON: JavaScript Object Notation

REST: Representational State Transfer

WDC: Watson Developer Cloud

WEM: Watson Experience Manager

GIT: Source code control management system