

Kofax Kapow™ 10 Training Preparation

What you should know before training

Overview

Robots developed with Kofax Kapow can interact with websites, databases, web services, Microsoft Windows desktop applications, terminal applications, and certain file and document formats to either execute tasks or extract data.

While it is not required for Kapow robot developers to have computer programming experience, being knowledgeable of the technologies used for web page development is a helpful starting point for developing successful robots. Many of these technologies will be used or referenced during training. In order to fully embrace the value of your Kapow training, it is recommended that you have at least a basic understanding of these technologies.

HTML

HTML is an acronym for **H**yper **T**ext **M**arkup **L**anguage and is the language used for describing web pages. HTML documents contain HTML tags and plain text. The HTML tags, sometimes referred to as elements, describe document content. In more formal terms an HTML element is everything between the start tag and the end tag, including the tags.

Tags

HTML tags are keywords surrounded by angle brackets like `<html>` and come in pairs like `<html></html>`. The first tag is known as the start tag, the second tag is known as the end tag. The end tag is written like the start tag with a forward slash before the tag name. The start and end tags are also referred to as opening and closing tags.

```
<tagname>content</tagname>  
Start tag           End tag
```

Elements

HTML elements are everything between the start tag and end tag including the tags. For example, you would say everything between the start tag and the end tag is the element.

```
<tagname>content </tagname>  
Start tag           End tag  
<<<<<<<< element >>>>>>>>
```

Attributes

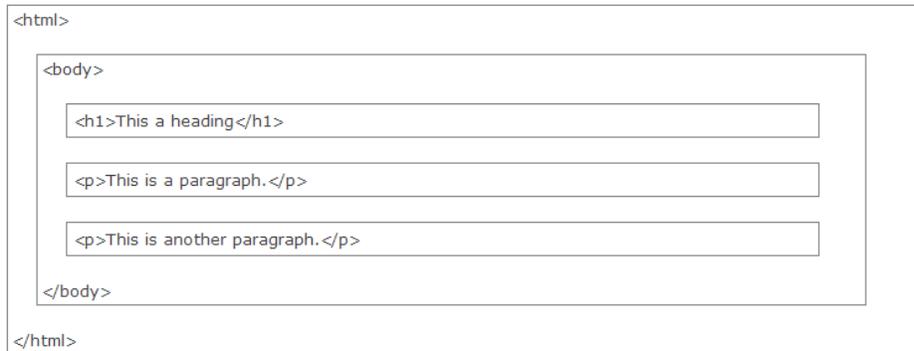
HTML tags can have attributes, which provide additional information about an element. Attributes are always found in the start tag and come in name\value pairs like: name="value".

```
<tagname attribute=attribute_value>content</tagname>  
Start tag                               End tag
```

Some common attributes you may see in HTML tags are: id, class, style, src, href.

Structure

The structure of a web page is composed of several HTML elements. The HTML elements can be “nested” inside another HTML element. This is a simple visualization of the structure of an HTML document or web page.



CSS

Known as **C**ascading **S**tyle **S**heets, this technology can define how to format and display HTML elements based on tag attributes such as id, class and style. Understanding how CSS is used to handle the display of HTML can provide benefits when developing robots and accessing elements that may not be visible when the page is loaded in the browser.

JavaScript

JavaScript is the scripting language for web pages. Modern HTML pages use JavaScript in many different ways to manage interaction with the page or dynamic presentations.

If desired, Kapow robot developers can write their own JavaScript functions to manipulate data or perform other custom functionality.

Advanced JavaScript

While it is very rare for robots to need to do so, Kapow allows you to develop direct interactions with JavaScript events as well as inject custom JavaScript to execute in the loaded page.

JavaScript is also used for “behind the scenes” executions called AJAX. Having a comfort level to read and understand JavaScript will allow you to be able to manually execute and extract data from these kinds of AJAX calls in your robot development.

Kapow Native File Formats

In addition to HTML, Kapow has data types for storing text, numbers, dates, etc. But it also supports several complex file formats as data types.

Each of the following file format data types are first class citizens in Kapow, meaning they all have a native view (syntax highlighting/presentation) and actions which supports reading, updating, inserting and deleting elements for the given data type.

The native file formats are listed below.

XML

XML stands for **eXtensible Markup Language** and is designed to transport and store data.

Robots can extract data directly from XML files or retrieved XML data packages of web services, REST or AJAX executions.

JSON

JSON stands for **JavaScript Object Notation** and is a syntax for storing and exchanging text information similar to XML, but with less declaration.

JSON data format may be received from web services, REST or AJAX calls, allowing robots to extract specified data elements.

Excel

Kapow can be used to open and extract data from Excel files. To be proficient in using Excel files you should understand how cell referencing works in Excel. For example, Sheet2!A:A is a reference to all rows of the first column in the second sheet.

CSV

Comma-separated values or character-separated values files store tabular data in plain-text form. Kapow can utilize CSV files as both input data and as output data. Native looping capabilities provide an efficient way to parse and extract data from CSV input.

HTTP

HTTP is a network protocol for the Web and stands for **Hypertext Transfer Protocol**. The Kapow browser engine is an HTTP client because it sends requests to an HTTP server similar to your normal web browser. Kapow also supports HTTP over SSL/TLS, also known as HTTPS.

SQL/Relational Database

SQL stands for **Structured Query Language** and is an industry standard database access language. Kapow robots can communicate with databases using SQL to retrieve data, create new data and alter data by updating or deleting records.

When working with relational databases it is important to have a solid understanding of database concepts and the specific database with which you are working.

String Concatenation

String concatenation is the operation of joining two character strings end-to-end. For example, the concatenation of "Kapow" and "Robots" is "KapowRobots", written in terms of an expression as "Kapow" + "Robots".

Kapow provides built-in functionality through expressions and converters that allow you to transform and concatenate strings from retrieved or extracted information.

Regular Expressions

Regular expressions, also known just as regex, is a sequence of characters that forms a search pattern that is used for pattern matching with strings. Pattern matching is a powerful tool that can be utilized when building Kapow Robots. It can provide a more accurate and robust way to find tags and extract data.

Web Services (SOAP & REST)

Web services are web application components that can be published, found and used on the web. Kapow provides mechanisms to easily call SOAP or REST style web services.

SOAP, or **S**imple **O**bject **A**ccess **P**rotocol, is a protocol specification for exchanging structured information of web services. SOAP relies on XML for its message format.

REST, or **R**epresentational **S**tate **T**ransfer, is a simple stateless architecture that generally runs over HTTP. REST services rely on HTTP verbs (GET, POST, PUT and DELETE) which dictate the actions that are requested. REST services transfer data using XML or JSON formats.

Device Automation

Device automation is used within Kapow for automation of Microsoft Windows desktop applications or terminal applications.

This works by the robot interacting with the applications as though it were an actual user using a keyboard and mouse. Device automation utilizes a specialized subset of the regular robot steps, and therefore can be integrated with all the features and functions referenced in this document.

Robot developers leveraging device automation should have strong familiarity with the applications being automated.

CSS Selectors

As an extension of how CSS is used for styling in HTML, CSS also comes with selectors (rules) for when to apply a given style. This mechanism is used within device automation for finding an element in a desktop or terminal application.

Business Process Modeling

Kapow automates standardized business processes as implemented by the robot developer. It is valuable for the robot development team to have subject matter expertise in the business processes to be implemented. It is also helpful for the team to include people who are experienced in analyzing and documenting business processes to thoroughly define desired handling for all the expected, alternate and exception paths.

Additional Learning Resources

If any of the above concepts are new to you and you'd like additional familiarity with them to gain the most benefit possible from your upcoming Kapow training, feel free to explore some of the following resources:

HTML

- W3Schools - <http://www.w3schools.com/html/>
- Codecademy - <http://www.codecademy.com/tracks/web>
- HTML Dog - <http://htmldog.com/guides/html/beginner/>
- Lynda.com - <http://www.lynda.com/HTML-training-tutorials/224-0.html>

CSS

- W3Schools - <http://www.w3schools.com/css/>
- MDN - <https://developer.mozilla.org/en-US/learn/css>
- Lynda.com - <http://www.lynda.com/CSS-training-tutorials/447-0.html>

CSS Selectors

- http://www.w3schools.com/cssref/css_selectors.asp

JavaScript

- W3School - <http://www.w3schools.com/js/>
- MDN - <https://developer.mozilla.org/en-US/docs/Web/JavaScript>
- Lynda.com - <http://www.lynda.com/JavaScript-training-tutorials/244-0.html>

XML

- W3School - <http://www.w3schools.com/xml/>
- Lynda.com - <http://www.lynda.com/XML-training-tutorials/334-0.html>

JSON

- W3School - http://www.w3schools.com/js/js_json_intro.asp
- Lynda.com - <http://www.lynda.com/JSON-training-tutorials/1551-0.html>

SQL

- W3School - <http://www.w3schools.com/sql/>
- SQLCourse.com - <http://www.sqlcourse.com/index.html>

Regular Expression Resources

- Regular-Expression.info (<http://www.regular-expressions.info/>)
- RegExOne (<http://regexone.com/>)
- The RegEx Coach - <http://www.weitz.de/regex-coach/>

SOAP/REST

- W3School - http://www.w3schools.com/xml/xml_services.asp
- Learn REST: A tutorial - <http://rest.elkstein.org/>