W3C Internationalization Checker - An API rewrite in Java

Joseph J Short

3 Swansea Road, Norwich, NR2 3HU, United Kingdom
Joseph.J.Short@gmail.com Tel: +44 (0) 7837 668932
jjs1 on irc:irc.freenode.net

Abstract

The World Wide Web Consortium (W3C) Internationalization (i18n) Checker aides web developers in the creation of language-sensitive web pages by providing a detailed summary of character encoding, language declarations, and so on. It is one of a number of free services that comprise the W3C's 'Validator Suite' for web developers. The project embodies W3C's objective of "making the web world-wide", and their slogan, "Leading the web to it's full potential".

The i18n Checker is currently a 'pre-final release' web-application written in PHP. As it stands, the logic that comprises the validation and summary is integrated into the logic of the web service. The whole web application must therefore be deployed and run on a PHP supporting server for its functionality to be used.

The purpose of this project is to bring the i18n up to date with it's counterpart validators by rewriting it in Java, and separating the validation logic in to an API. This will leave the W3C with a modularised, easy to deploy API, which can be run by itself or integrated in to other applications.

The project also has a focus on robust and maintainable code. An existing test-suite must be ported in to Java, and the new Java code must conform to these tests.

The deliverables of this project, in decreasing order of priority, are: A Java API version of the i18n Checker; A Java version of the test suite and related integration; project documentation and integration with the www interface; and further development of the i18n Checker by adding support for HTML5 tags and so on.

Contents

1 Project proposal

`TODO: Section introduction based on abstract`

1.1 Major deliverables

A rewrite of the i18n Checker source code in Java

A rewrite of the test suite in Java

Project documentation This includes, code commenting and documentation using Javadoc, project description using a long README file, and relevant writing in LaTeX.

Extension of the i18n Checker

1.2 Skills and technologies

My mentor has suggested that I choose a selection of appropriate tools to use. The tools I plan to use are listed in this section.

Apache Maven I shall use a Maven Project Object Model (POM) to describe the software project, and to handle the dependencies and build tasks.

Git and github.org I shall be using Git to track changes to the source code. Github.org shall be my remote host, and the remote repository will be created using W3C's user account.

JUnit I shall use JUnit for the ported test suite. JUnit integrates well with Maven, providing me with automated testing.

async-http-client I'll use the `async-http-client` library to process HTTP requests as part of the API.

Jenkins CI Jenkins is integration testing server that focuses on build automation and a variety of deployment tasks. I'll use this server, either in a public or private setting, to continuously test and build snapshots of my code.

Other The project shall be documented using Javadoc, Markdown and TeXwhere appropriate.

1.3 Time-line

Week 1 - June 17th - Familiarisation The first week shall be used to: download and run the existing code; set up a development environment with the right tools and technologies; conduct research about the purpose and domain of the i18n Checker; and create a rough design for the new Java API. The deliverables are:

- a short summary of my domain analysis, including a glossary of domain-specific terminology; and
- a short design document with a proposed class diagram and/or architectural diagram.

Week 2 - June 17th - Design and preparation After week 1 my mentor will be able to. Deliverables are:

- a template software project tracked using a SCM tool linked to a remote backup; and
- 'stub' classes and method definitions.

Weeks 3 to 6 - July 1st - First sprint As soon as the above duties are completed, work on implementation can begin. In this period I shall implement the API and begin porting the test suite in to JUnit. I shall talk to my mentors for guidance, and update them at regular intervals.

Week 7 - July 29th - Week of mid-term evaluation deadline Early in the week, me and my mentor shall meet to discuss the progress of the project. This is in preparation for the mid-term student and mentor evaluations (deadline on August 2nd).

Weeks 8 to 14 - August 5th - Second sprint The second period of implementation. Initially I shall act upon any suggestions and advice from the mid-term evaluation. The rest of the second implementation shall be focused on sequential completion of the major deliverables.

Week 15 - September 16th - Last week The 16th is Google's "suggested 'pencils down' date". In accordance, I shall use the last week to make sure that I leave the i18n Checker in a good, robust condition. Activities for the last week are: 'scrubbing code', integration and acceptance testing, updating the documentation and preparing the project for future developers.

Further details I shall spend about 30-40 hours a week on the project, and manage my breaks in a flexible way. I shall establish a steady presence on the W3C IRC channel (`irc://irc.w3.org/#w3c`)

1.4 Mentors' details

My mentors from the W3C are:

- Alexandre Bertails (`bertails@w3.org`); and
- Thomas Gambet (`tgambet@w3.org`).

2 Personal profile

I'm a bright and hard-working undergraduate student at the School of Computing Sciences at the University of East Anglia (UEA). I'm excelling at the undergraduate degree course of Master of Computing in Computing Science and expecting to graduate with a first in the summer of 2015.

I've worked in part-time jobs since leaving school, and have a professional and driven attitude towards my work. Soon I'll be at the start of a new career in computer science!

2.1 Background and education

2009 to 2012 Catering Assistant (part time) After leaving school I found work as a catering assistant. I gained experience in: working long hours in a busy environment, working in a team to a demanding schedule, and customer service. This job helped me to develop a professional attitude to work, which I think has been indispensable in my studies. And it has given me the drive to apply myself, excel, and achieve more in life.

2010-11 Access course *Applied Computing Science with a Foundation Year.*
A year of introductory theory, programming, maths, writing and teamwork.
Finished with an overall grade of 68%.

2011-Current *Master of Computing in Computing Science.*
An undergraduate master's course with a wide coverage of topics in the field of computer science. Finished my first year with an overall grade of 75%.

I've studied:

- Technical writing
- Presentations
- Teamwork
- Advanced programming
- Theoretical computing
- Software engineering
- Data structures and algorithms
- Databases
- Architectures and operating systems

And in the next two years I'll study:

- Research methods
- Dissertation
- Master's project
- Machine learning
- Information retrieval
- Further software engineering
- Networks

2.2 Suitability

`TODO - this section needs filling out.`

Suitability of the project Working with W3C - standing on the shoulders of giants. Opportunity to benefit the organisation.

Suitability of me I have substantial experience with all of the tools and technologies. Work at uni. I learned to program with PHP.

2.3 Open source

I do not have any experience in *contributing* to open source projects; this is very much to my detriment. As I become a professional in the field I'm very keen on creating a presence in the open source community. Indeed, the opportunity to contribute to open source is one of the driving factors behind this proposal; this project may very well override another application that I've already submitted.

I have a great admiration for open source projects, especially those that are developed in a public and collaborative fashion. I feel that I owe the most to members of: The Apache Software Foundation, W3C, the Software Freedom Conservancy, and the developers of GNU.

2.4 References

`TODO: I shall ask my professors for recommendations and provide them here in the form of a broad question coupled with the professor's answer, and their email address.`

Is Joseph a capable programmer? Is he familiar with Java and object-oriented principles?

Is Joseph familiar with principles of software engineering? Is Joseph familiar with a wide selection of tools of software development?

2.5 Other commitments

Summer research placement I have applied for funding for an 8 week summer research placement in my school. It is likely that I will get this funding. If that were the case, there would be a four-week overlap with the start of my GSoC project. I don't think this is a problem. I won't be able to keep regular office hours for this period, but I'll still be able to put about 30 hours in a week. I have experience with long hour days and weeks for my previous part-time work. That said, I will respect the wishes of W3C and my mentors. Moreover, if it comes to an ultimatum, this project is definitely my preference (i.e. If I had to choose, I'd choose this project).

GSoC Other than the above I have no commitments this summer. In fact I'm very enthusiastic about putting a lot of time in to this project; whereas in previous summers I'd have spent about 35 to 60 hours a week working in my dull catering job, now I'll have the opportunity to work on this fantastic project, and to garner the respect of members of the open source community!