Group 4

Forager Meeting Minutes : Sprint Planning

October 18, 2012

Opening:
The first meeting of *Group4* was called to order at 6:00 pm on October 18, 2012 at SPSU by The Team.

Present:
Matthew Powell
Robin Mays
Thomas Couture
Samuel Hall

Present:
A. Approval of Agenda

The agenda was unanimously approved as discussed.

B. Approval of Minutes

The minutes of the previous meeting are available for viewing.

C. Opening Issues

Complete risk mitigation and sprint planning.

D. New Business
Review user stories/requirements for Forager's two sprints. Discuss and update risk mitigation for the upcoming sprints.

User Stories/Requirements:
ID: 1
As a user, I would like to be able to visit and access all pages of my website.
Priority: 10, Cost: 20

ID: 2
As a user, I would like this program to record any resources that are unavailable, including dead links,
missing images, scripts or css files.
Priority: 10, Cost: 20

ID: 3
As a user I must be able to view a report of a given scan. This report should show all broken links and
missing images that fall under my domain. These scans should be listed and accessible from a website.

Priority: 10, Cost: 30

ID: 4
As a user I would like to view scans and start new scans from a website.
Priority:10, Cost 20

ID: 5
As a user, I would like to be able to sort the reports that are presented to me. Useful sorting functions
would be: by the order in which the pages were visited, alphabetically, and alphabetically by the parent
page.
Priority: 7, Cost: 10

ID: 6
As a user viewing a report, I should be able to generate that report in a printer friendly format.
Priority: 6, Cost: 10

ID:7
As a user, I would like to be able to select two scans and show only the items that have changed.
Priority: 6, Cost: 6

ID: 8
As a user viewing a report, I should be able to view reports from scans that are in progress.
Priority: 6, Cost: 4

ID:9
As a user, I would like to be able to limit the run time of a scan when
I start it, either by time, or by distance from the start page.
Priority: 6, Cost: 12

ID: 10
As a user I would like to select a scan, and run a new scan that will check if the previous errors have
been corrected.
Priority: 5, Cost: 10

ID: 11
As a user, I would like for reports to include pages that are accessible
over secured links.
Priority: 5, Cost: 8

ID:12
As a user, I would like to sort a report based on the subdomain.

Priority: 3, Cost: 14

ID: 13
As a user, I should have to login before initiating a scan or viewing a report.
Priority: 3, Cost: 2

ID:14
As a user, I might like to pause a scan that was currently in progress.
Priority: 2, Cost: 4

ID: 15
As a user, having Scans that were automatically run at regular intervals.
Priority: 1, Cost: 6

ID:16
As a user I would like to see page load times in my reports.
Priority:1 Cost: 2

Risk mitigation plan:
We could have library problems dealing with SSL pages.
This can be mitigated by research, and is not a huge issue as it's a
relatively low priority story.

We could have communications problems between the report viewer and the
web crawler.
This would be a huge issue. A backup plan might be to use python to
generate JSON directly from pickled object from the web-crawler, or
potentially generate one shot reports directly in the web-crawler, those
these options are both less flexible.

Dealing with javascript could make me so angry I stick forks in my eyes.
This might make typing difficult, and I should probably keep a bottle of
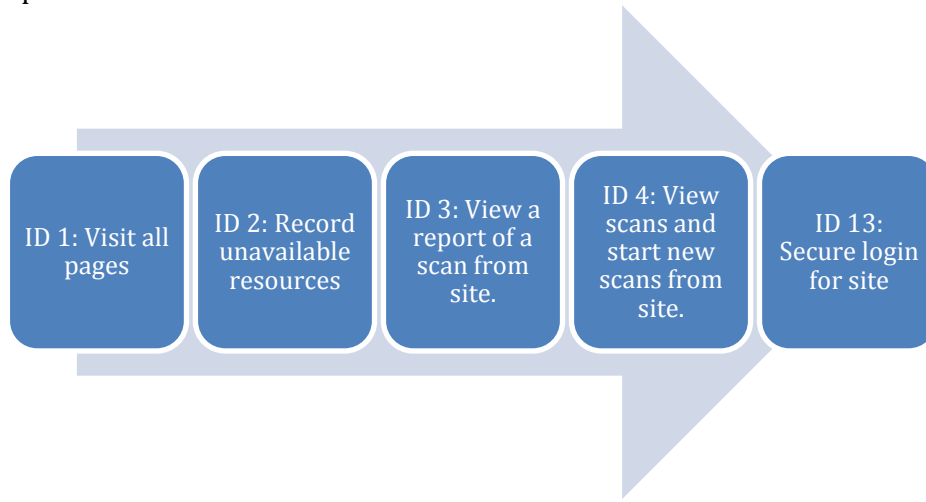medication uncorked and nearby to prevent such an occurrence.

System Architecture:
We will be writing the web-crawler in python utilizing libcurl and communicating with a postgres
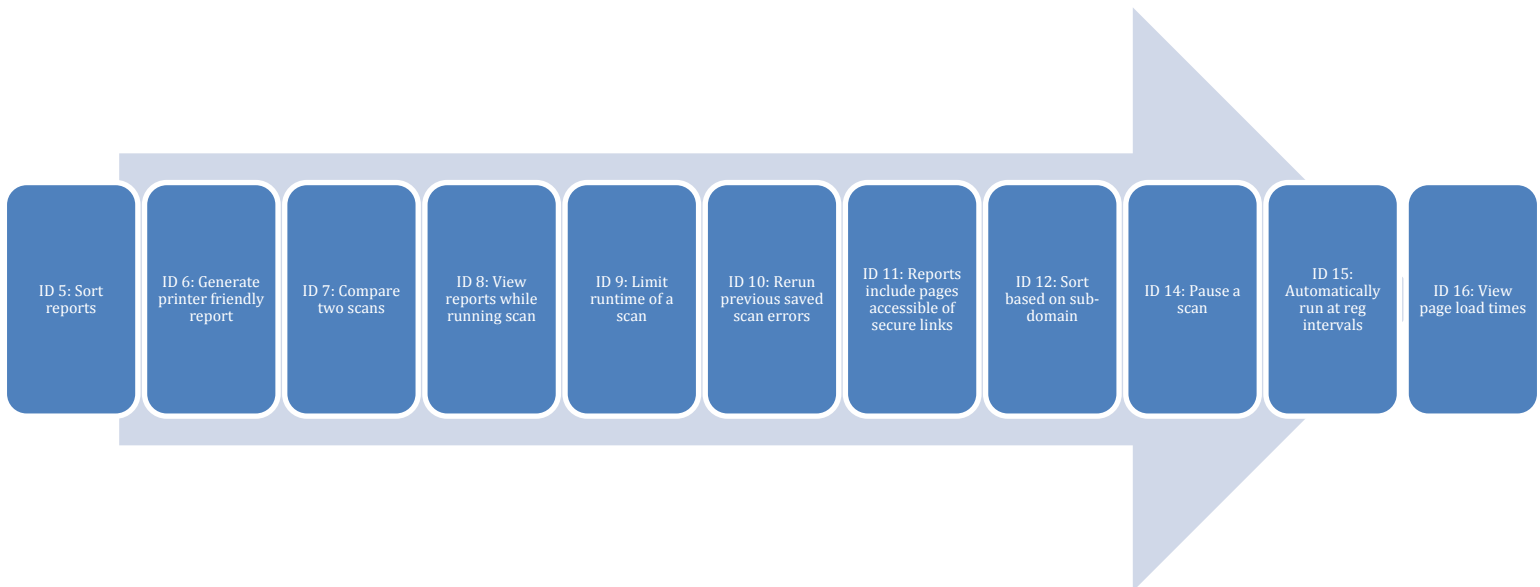datastore using PsychoPG2.

The front end will be in PHP, but to provide user interaction (Sorting lists
in sane ways, etc), we're going to have to use javascript, which means using jQuery.

Sprint Planning : Cycle Plans

Sprint 1:

ID 1: Visit all pages | ID 2: Record unavailable resources | ID 3: View a report of a scan from site. | ID 4: View scans and start new scans from site. | ID 13: Secure login for site

Sprint 2:

ID 5: Sort reports | ID 6: Generate printer friendly report | ID 7: Compare two scans | ID 8: View reports while running scan | ID 9: Limit runtime of a scan | ID 10: Rerun previous saved scan errors | ID 11: Reports include pages accessible of secure links | ID 12: Sort based on sub-domain | ID 14: Pause a scan | ID 15: Automatically run at reg intervals | ID 16: View page load times

E. Agenda for Next Meeting

Design architecture for Forager and review and update risk mitigation plan.

Adjournment:
This meeting was adjourned at 7:30pm by The Team. The next general meeting will be at 6:00pm on October 25, 2012 at SPSU.

Minutes submitted by: Robin Mays

Approved by: Matthew Powell