

Sustainability Guide

Last Updated November 2019

What We Will Host	1
Key Points From Minimum Hosting Requirements	1
Triage Process	2
Sustainability/Preservation Processes	3
Contacting Site “Owners”	3
Remediation of Sites	3
Flattening of Sites	3
Archiving/Preserving of Sites	4
References	4

What We Will Host

RRCHNM commits to hosting the following types of websites, under the conditions laid out in the Server Sustainability Policy and Minimum Hosting Requirements.

1. RRCHNM projects
2. RRCHNM-associated research projects
3. RRCHNM-associated teaching projects
4. Contracted projects under development
5. Contracted projects paying for hosting
6. Legacy projects

RRCHNM team members should keep a running log in the Basecamp sustainability project of all actions taken on a project (including annual re-evaluations, as well as looking at a project and assessing it as currently not in need of remediation or other action)

Key Points From Minimum Hosting Requirements

1. finished projects must have documentation deposited in the Basecamp sustainability project, which includes the following information
 - a. project contact(s)
 - b. project purpose
 - c. essential project functionality
 - d. CMS/technical stack description
 - e. external dependencies
 - f. funders/grants list

- g. sustainability commitment given to funders/other legal entities
- 2. all sites must either meet WCAG 2.0 and Section 508 standards or provide justification for why portions of the site cannot be made completely accessible; alternatively, project owners can demonstrate that they are actively working towards making their sites as accessible as possible
- 3. new content must have been posted to the site within the past 12 months; sites that are not being updated at least once a year may be flattened at the senior staff's discretion
- 4. sites that cannot be updated for more than 6 months are subject to being triaged or otherwise removed from RRCHNM servers

Triage Process

When triaging projects for sustainability, the following factors will be taken into account:

1. Annual traffic to site
 - a. sites receiving over 100K unique visitors a year will be high priority sites for remediation when functionality begins to deteriorate
 - b. sites receiving 10-99K unique visitors a year will be medium priority sites for remediation when functionality begins to deteriorate and may instead be flattened for sustainability
 - c. sites receiving 1-9K unique visitors a year will be considered for flattening, rather than remediation, when functionality begins to deteriorate
 - d. sites receiving less than 1K unique visitors a year will be considered for archiving, rather than remediation, when functionality begins to deteriorate
2. Difficulty of remediation

When a site has been identified as possibly needing remediation, the first step in this process should be having an RRCHNM team member audit the site to determine what steps will need to be taken to remediate it; the remediation steps will be turned into an internal project estimate (time/money) for consideration by the senior staff. Part of this consideration will include determining whether only part of the site is worth remediating.

 - a. sites that require less than 10 hours of FTE work will be high priority sites for remediation when functionality begins to deteriorate
 - b. sites that require 10-40 hours of FTE work will be medium priority sites for remediation when functionality begins to deteriorate
 - c. sites that require 40-120 hours of FTE work will be low priority sites for remediation when functionality begins to deteriorate
 - d. sites that require 120+ hours of FTE work should be assessed for grant-funding rather than self-funded remediation
3. Security risk
 - a. sites that are urgent security risks will be the highest priority for remediation and should be temporarily removed from public access
 - b. sites that have security risks will be medium priority sites for remediation and should be closely monitored and/or temporarily flattened
4. Intellectual merits of the project

5. Project prestige

Sustainability/Preservation Processes

Contacting Site “Owners”

Good faith efforts will be made to contact site owners before their sites are remediated, flattened, archived, or otherwise deleted.

1. All projects going forward should have contact information included in their documentation. It is the site owner’s responsibility to ensure their documentation is up to date.
2. For previous projects, if Googling names fails to turn up an email address or other method of contact for a project PI/owner or other responsible contact person, we may first update the landing page of the project with a notification indicating what is about to happen to the site and providing the RRCHNM email for contact in case of concerns. Whenever possible, this should be left up for 4-6 weeks before work begins.
3. When the owner cannot be found, extra care should be made in preserving the original code, database, and content of the site, so that they can be provided to the owner should that person later surface.

Remediation of Sites

1. If a site is slated to be remediated, its current codebase may or may not be archived, depending on the scope of the work to be done.
2. The site’s most important features should be identified (see documentation requirement above) and prioritized in the remediation process.
3. Remediated sites should ideally be designed to degrade gracefully, to reduce long-term technical debt.
4. Containerization may be a solution for sites that are security risks vs. sites whose functionality has begun to deteriorate.

Flattening of Sites

1. If a site is slated to be flattened (HTML/CSS/JS only), its current codebase may first be archived (see below).
2. The active dates of the project will be added to the homepage of the site and any database functionality (e.g. log-in links for CMS sites) should be deactivated.
3. The site will be flattened using wget/web-crawling, plug-ins (e.g. WP sites), or some other method chosen by RRCHNM team members.
4. The old site will be removed and replaced by the flattened version of the site.

Archiving/Preserving of Sites

If a site is slated to be archived/preserved, we will create and upload the following materials to the MARS repository:

1. TIFF or JPEG screenshots of its homepage and any other significant pages
2. Webcrawled WARC file of the site
3. Zipped folder with flattened version or original codebase of the site depending on code complexity
4. Database dump, if applicable
5. Information from any “about” pages should be added to the deposit metadata

References

[Digital Documentation Process](#) (External project, in-progress)

[Project Endings](#) (External project)

their [resources list](#) is also very helpful

[The Socio-Technical Sustainability Roadmap](#) (External project)