

Building Community in the Rosetta Working Group



History

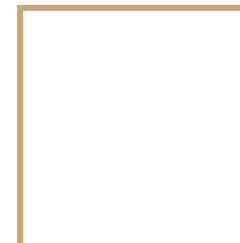


In the Beginning

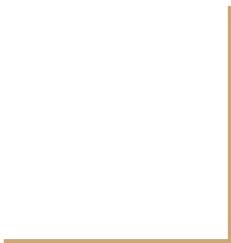
- Previously, users were supported by the Rosetta Advisory Group.
- Is there enough interest to sustain a users group?
- Identify its purpose
- Should it remain an independent group?
- First Steps
 - The Rosetta Users Group will meet before the Advisory Group meeting.
 - A wiki was established in 2013.

Guiding principles for the group

- Define and agree on strategic goals for the user community.
- Provide a platform for sharing information about the product and practices to address community needs.
- Serve as a shared vehicle for requesting Rosetta enhancements and information.
- Contribute to the broader objective of digital preservation.
- Offer a platform to share experiences with Rosetta and digital preservation.



Structure



How do we group ourselves

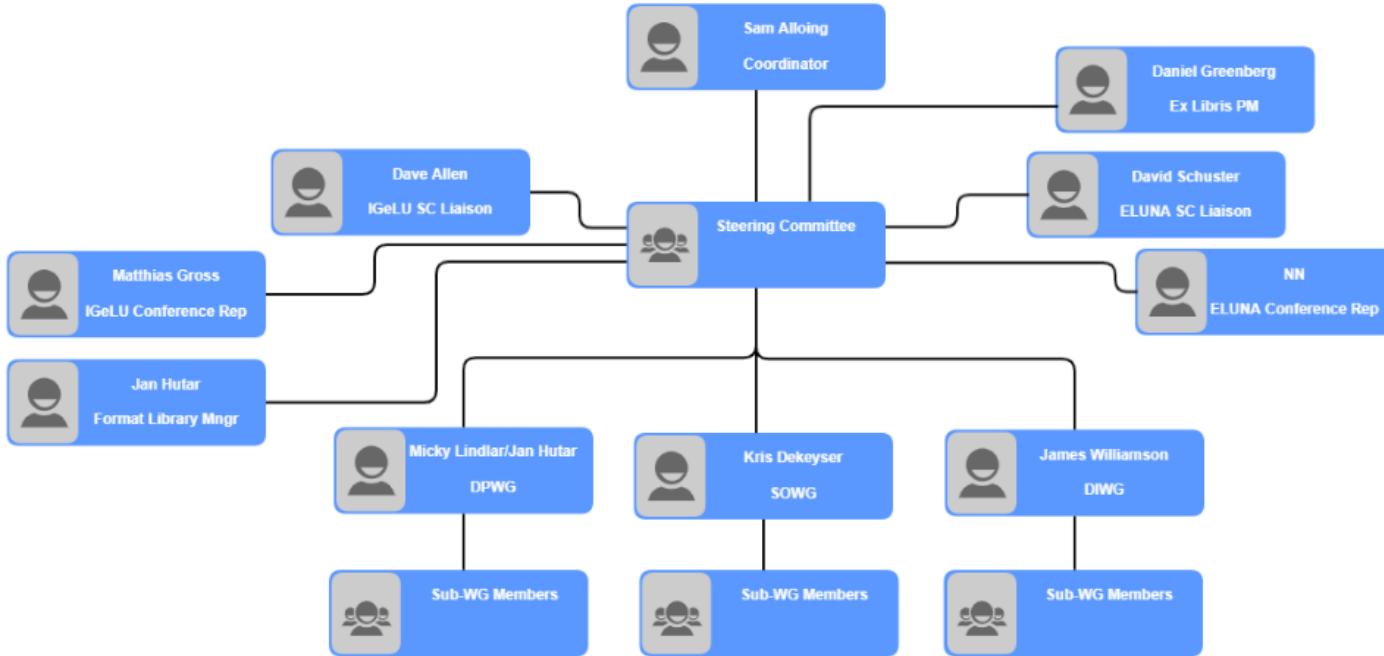
- Should groups be organized by type (e.g., academic, public, private)?
- Are there sufficient differences among users to warrant these divisions?
- Is the product primarily driven by institutions with established expertise?
- Should a Chair/Co-Chair volunteer to serve as the primary line of communication?

Rosetta Users Group (Independent Era)

- Concerns of holding multiple meetings (money and time)
- Funding meeting before the Advisory Group is pulled
- Papers and ID'd Pressure Points
 - Points Gathered and describe in detail for Ex Libris
 - Divide up responsibilities in focus groups
- Line up with Roadmaps

Rosetta Working Group (Major Label Era)

- Joined as a joint ELUNA/IGeLU Working Group in 2017.
- Developed subgroups around system components:
 - Digital Preservation (validation, metadata extraction, preservation planning)
 - Format Library
 - System
 - Delivery and Integration
- Topics are generated by members and assigned to the appropriate area of interest or request.
- Term-limited coordinator position.



How We Are Today

Enhancement Requests

Balancing

- Major requirements vs small changes
- Immediate needs vs long term work

Enhancement Request Challenges

- Struggle to lineup with the Rosetta Roadmap
- Limited transparency on support cases and multi-year struggle to understand what users were asking
- Problems on how to gather and rank requests from a variety of sources
- Emails from customers, internal emails, feedback at meetings were the only sources

RWG Enhancement Process

- Drafting of ER
 - Enhancement Request Template on the wiki
- Collection of use cases within the RUG
- Defining of requirements
- Feedback Phase
 - Sub Working Group, Wider RUG, and Clarivate (On Basecamp)
- Implementation and Testing

DI07-Thumbnail Generation
Created by Kris De Meyer last updated by James Williamson 14 minutes ago • 1 minute read

BUG WG	Open
Enhancement Request Call	October 2023
Document status	Submitted
Enhancement Request Lead	Kris De Meyer (LIBS)
Enhancement Request Contributors	Dave Allen (SLQ), Teresa Seneau (Getty), James Williamson (SMU), Michael Huber (SSB)
Implementation Status	Scheduled for Rosetta 9.0
QA	
QA Status	NOT STARTED

Introduction
Currently the process of automatic thumbnail generation in Rosetta is documented, but not 100% predictable and very rigid. The process is explained in the Rosetta Configuration Guide. In short, Rosetta will select a representation, preferably a Derived Copy or any other REP without access rights. If the user has no access to the REP, a 'no access' thumbnail is shown. Otherwise Rosetta will create a thumbnail by following the Thumbnail Task Rules set. If no such rule exists, a generic thumbnail icon will be returned instead. A thumbnail will be generated for only a limited number (5 by default) of files per REP. There are a number of problems with this:

- 1. The thumbnail task rules do not determine which DC representation will be chosen when multiple DCs fit the requirements. The algorithm explicitly states a random REP will be chosen when there is no DC or REP with no access to choose from. In the Delivery context, a thumbnail is generated on an asset that is part of the DC as well as any other representation and the unrepresentativeness of the thumbnail generation is cumbersome.
- 2. The thumbnail is generated on the fly and Rosetta then relies on caching for storage. The thumbnail is not part of the DC itself.
- 3. The Thumbnail Task Rules only has the extension as input parameter. It is not possible to generate different thumbnail sizes based on other input parameters (e.g. Entity type).
- 4. The thumbnail task rules only has the extension as input parameter. It is not possible to generate different thumbnail sizes based on other input parameters (e.g. Entity type).
- 5. Settings that influence the thumbnail algorithm (like granting access rights and the thumbnail task rules) are done on a global (consortium) level. No differentiation on institutional or collection level is possible.

It is possible to run tasks that generate thumbnails with specific settings that are more customized as part of a task chain during maintenance operations. But still relying on the caching makes this rather hazy and inconsistent. The lack of control on the cached storage is problematic. The solution proposed in this ER is to have a dedicated representation in which custom thumbnail files can be stored and kept alongside the other representations of the DC. The existing thumbnail algorithm is left undisturbed, but preference has to be given first and foremost to the thumbnail representations.

Use Cases
LIB Leuven (LIBS)

- A collection of images DC has no presentation master and most derived copies protected by access rights. One Derived Copy is unprotected, but it heavily watermarked. Rosetta will choose the unprotected and watermarked DC to generate a thumbnail from, which turns out to be more watermark-irreversible. We want to be able to specify which DC rep is to be used for the thumbnail generation.
- A selection of photos is archived and presented in Rosetta. The photos are not public, but within the compounds of the institution. It should be possible to view the thumbnails of the photos. We need to be able to define access rights on these thumbnails. Outside of the institution location, only a generic icon should be displayed.
- In the delivery context, we want to be able to generate a thumbnail for an asset that is part of a collection. For example we have an image that has not given their consent to be published outside of the archive. The thumbnails, which are relatively large in this case, need to have those files blurred in order to be published within the compounds of the institution.
- Thumbnails can be used in different contexts and these contexts often have different requirements. Different contexts may be search result, detail page or an image carousel. Or it may be an web site requiring a preview of the document or an inventory list that prefers a generic icon. On web pages with a large amount of thumbnails retrieving a pre-generated thumbnail of the correct thumbnail is a time-saver and improves the application performance.
- For certain material types automatic thumbnail generation is simply not possible. Digitized old books typically are bound in similar or the same covers and users want to see a thumbnail of the first meaningful page. For movies, the preferred thumbnail could be the poster or a thumbnail with the first frame. We want to be able to specify which thumbnail should be used.

BV/BSB

- We use IIP Manifests to show digitized work including rather large volumes (up to 1000 pages). In Germany, state funded digitization projects are required to have a standardized footer added to every digitized page. The automated thumbnail generation of Rosetta creates a thumbnail including this footer, which we would rather avoid. This is why we want to use our own thumbnails (without the footer).

SLQ
SLQ/H1 - SLQ publishes access to Rosetta assets via Primo - In Primo we would like to display a thumbnail that is representative of the digital asset.
Refer https://digital.slq.qld.gov.au/delivery/DeliveryManagerService?doi=6160682&change_jpg=on for the asset that is to be delivered - the thumbnail that is currently in use is https://digital.slq.qld.gov.au/delivery/DeliveryManagerService?doi=6160682&change_jpg=on&thumb=1



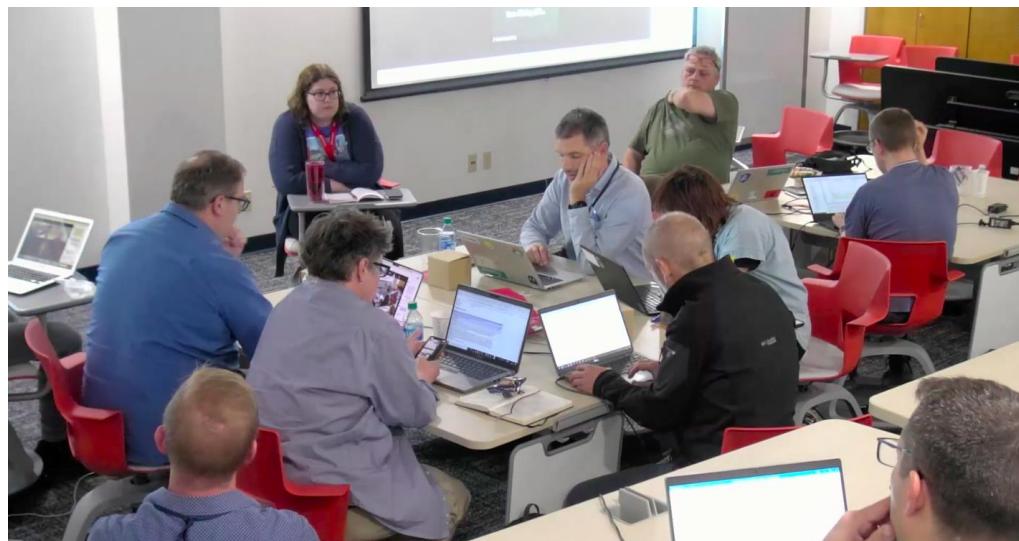
Instead of the front page which is the binding that SLQ has done for the original see image above – we would like to use a page within the publication which is more representative of the work



Making Global
Connection Feel
Local

Resource Sharing

- Rosetta Working Group Wiki
 - Customer Information (Workflows/Integrations)
 - Presentations/Publications/Documentation
- Hack-A-Thon at the RWG Annual Meeting
 - Format Issues of Rosetta Customers
 - Property Mapping



Open Discussions

Last 4 years, the RWG has run multiple open discussion sessions

Short presentations from RWG members/Discussion amongst attendees

- Preservation Planning
- Risk Management
- Web Archiving
- Workflows
- Storage
- Monitoring and Reporting

16 Open discussions in total



Food and Fun

User Group Gathering

- Try to gather outside the meeting venue at least once
- Event sponsored by the host and Ex Libris
- My UnTapped stats are very grateful for the RUG/RWG meetings



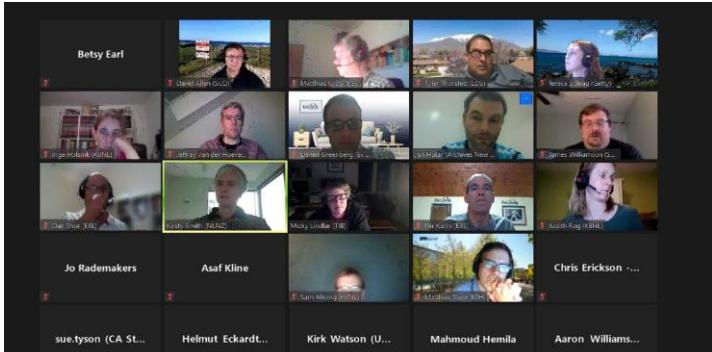




Building Through Crisis

Rosetta Working Group (The COVID Years)

- Shift to online Rosetta Working Group meetings for three years
- Tested splitting our meetings to accomodate time zone difference
 - Dedicated regional sections of the meetings
- Highest attendance for the three year the meeting was online
- Socialization and coffee break opportunity throughout





Where We Are Today



Contact Info

James Williamson

Digital Preservation Librarian

Southern Methodist University

214-768-2323

jdwilliamson@smu.edu

