

What to do when your repository enters middle age

Don Brower, Mikala Narlock

Publication Date

14-12-2023

License

This work is made available under a CC BY 4.0 license and should only be used in accordance with that license.

Citation for this work (American Psychological Association 7th edition)

Brower, D., & Narlock, M. (2020). *What to do when your repository enters middle age* (Version 1). University of Notre Dame. <https://doi.org/10.7274/r0-e32v-2h81>

This work was downloaded from CurateND, the University of Notre Dame's institutional repository.

For more information about this work, to report or an issue, or to preserve and share your original work, please contact the CurateND team for assistance at curate@nd.edu.

WHAT TO DO WHEN YOUR REPOSITORY ENTERS MIDDLE AGE

Don Brower and Mikala Narlock,
University of Notre Dame,
Samvera Connect 2020

DB: "What to do when your repository enters middle age"

Don Brower, Digital Projects Lead, Co-Product Owner
Mikala Narlock, Digital Collections Strategy Librarian, Co-Product Owner
Hesburgh Libraries
University of Notre Dame

OUTLINE

- Introductions
- Our repository's mid-life crisis
 - Refining the scope
 - Focusing our work
 - Centering the repository
- Conclusion
- Q&A

DB: Here is what we'll talk about:

- Introductions
- Our repository's mid-life crisis
 - Refining the scope
 - Focusing our work
 - Centering the repository
- Conclusion
- Q&A



CURATEND

DB: Our institutional repository CurateND. Or, as we call it, our beast of burden.

The cutting edge (~2013)

Fedora 3

Solr 4

Sufia (pre Hyrax)

Custom ETD Workflow
(Sipity)

Custom Preservation
Storage (Bendo)

THE TECHNOLOGY

DB: Our technology was cutting-edge Samvera stuff circa 2013: Fedora 3, Solr 4. Some Samvera, then Hydra, gems. The thing is, this is still our tech stack in 2020. Nothing has really changed.

We do have some custom pieces, one for ETD the deposit workflow and one for preservation storage.

pre-2012: Idea
conceptualized

2013: First code commit

2014: Soft-launch

2015: Oversight team

2016: First PO appointed

2017: PO leaves; another
appointed

2019: Don and Mikala
become co-POs

THE TIMELINE

DB: The idea for an institutional repository was around for a while, but got traction in 2013 when the development started.

The code was originally based of Penn State's ScholarSphere, and quickly changed to use Sufia once Justin Coyne extracted it.

The self-deposit repository was launched in 2014. Originally development was overseen by the development team, but in 2015 a larger team was formed from interested people in the library.

In 2016 the first product owner was appointed to prioritize and guide the developers to making useful features.

We have both been co-product owners since the beginning of 2019.

Two developers

Two product owners

Innumerable stakeholders

THE TEAM

DB: Our current team is 2 developers, 2 product owners, and (>10) stakeholders in the library.

But over the life of the service there have been many more developers (6), product owners (2), and stakeholders (???).

MID-LIFE CRISIS

DB: Like any system, once it was created, the repository has grown organically and in places has become complex, confusing, and/or contradictory.
This was definitely our situation.

- Scholarly materials
 - ETDs
 - Patents
 - Datasets
- Digital Collections
 - Lantern slides
 - EAD Finding Aids
- Miscellaneous
 - “Just put it in Curate”
 - E-Resources: e.g. journals, videos

REPOSITORY SCOPE

DB: To set the stage for this mid-life upheaval was an overly-broad scope that left nearly everything in scope. Our repository was housing scholarly materials, like ETDs, Patents, Datasets, and other more traditional IR content like journal articles and journal citations; we also had digital collections, including a massive collection of thousands of architectural lantern slides; and, because CurateND had the infrastructure, some purchased electronic resources. Really anything that needed to be made available online was “Just put in Curate” because it was one of the few places we as an institution had easy access to make content accessible.

One of the original design goals was to accommodate high-energy physics data (~60 TB!!!). This goal was never reached.

CURATEND: JACK OF ALL TRADES

MASTER OF NONE

DB: Which meant that the repository handled lots of things OK, or less than OK, but nothing well.

And also, some simple (conceptual) tasks were really hard to do with the software.
(e.g. change a record type.)

MID-LIFE TURMOIL

- Limited adoption or campus awareness
- Technical debt
- Limited internal buy-in
- Limited outreach
- Organizational upheaval

DB: which all led up to, as we have called it, mid-life turmoil.

In preparation for this talk, we researched mid-life crises; one thing some researchers note is a difference between mid-life crises and mid-life stressors; mid-life stressors is a big life-event that happens on top of the build-up of daily life stressors for the past ~40 years; straw that backs the camel's back.

In this instance, we had several straws break the back.

- Campus wide, there was little adoption or even awareness of the system;
- increasing amount of technical debt that we were defaulting on;
- limited buy-in from internal stakeholders, like library faculty;
- limited outreach;
- and there was organizational upheaval in the form of a Mellon grant for digital collection site, and start of library wide reorganization & competing for resources.

HOW FORWARD?

MRN: With all of this mid-life turmoil, how should the system move forward? As Don said, this system was not widely adopted, and maintaining the system was emotionally and technically laborious. Setting aside the sunk cost of time and energy, we thought carefully about what role CurateND would play in the emerging digital library landscape at Hesburgh Libraries.



MAKE IT USEFUL.

MRN: And we eventually came to the realization that, in order for it to continue we needed to make CurateND useful to our campus community.

FOCUSING OUR EFFORTS

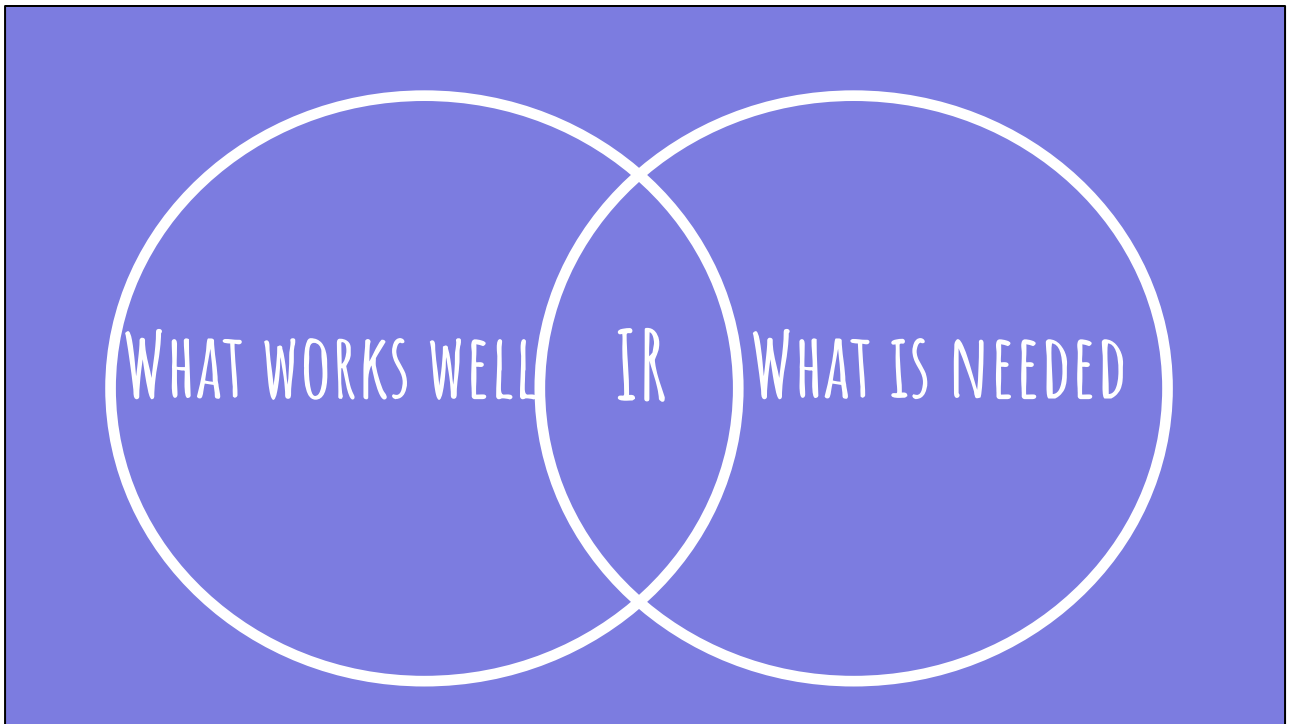
- Refined the scope
- Focused development effort
- Centered the repository

MRN: To make the system useful, we employed a three-pronged approach: refining the scope; focusing our development efforts; and centering the repository.



~~NEW AND
IMPROVED?~~
OLD AND REFINED

MRN: Our system is not new and improved-- instead, it's old and refined



MRN: To refine the scope, we started by critically evaluating CurateND, and examining the strengths of the system as it stood: without any improvements, what about the system and associated services works well? Access and permissioning controls-- flexible at both record and file levels. Additionally, CurateND handles persistence incredibly well: in addition to locally create purls, we also offer DOIs for users. Moreover, self-deposit for average files works well. The last thing our little system does well is preservation: our home-grown backend (that Don can discuss in more details for those curious) is extremely robust, and has actually proven itself on a few occasions.

After we went through the exercise, we then discussed what was needed in our current digital library landscape. With the in-process Mellon Grant (now called MARBLE) targeting digital collections, what was missing across campus was a discipline-agnostic repository.

So, we decided to refine our original scope from Digital Collections, Scholarly materials, and Misc, to just focus on Scholarly Materials (and misc came along for the ride) and become an IR.



USEFULNESS

EVERYTHING ELSE

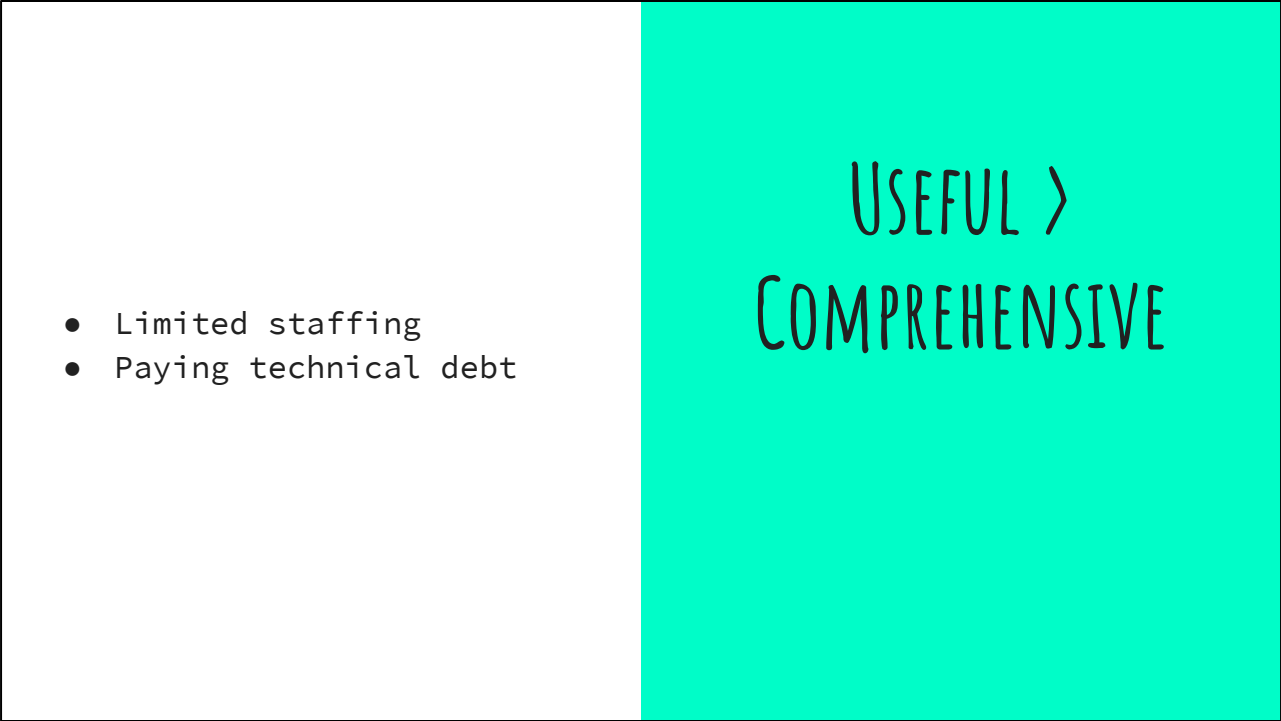
MRN: As we've already hinted at, our guiding principle was centered around usefulness, which was more important than just about everything else in our system.

By usefulness, we mean that, as product owners of the institutional repository, our job is to help our local community do their job and do it well. We're not selling them a repository: there are dozens of discipline specific (and agnostic) repositories they can turn to; we are helpers, and we have (and will continue to!) refer folks to other repositories or services we think are more appropriate.

We considered very carefully what services we are in the best position to offer, either because they're unique to us OR because they are of such value that our local community would be more inclined to use our services and system.

Focused scope, we further refined our efforts on articles and article deposits: not flashy, but the focus made it easy to conduct outreach-- datasets are a little more squishy (e.g., qualitative v. quantitative), but articles are much more tangible. This also made it easy to collaborate with our librarian colleagues, as we could team up to discuss scholarly communication issues.

(We also articulated non-goals, or things that we were not going to focus on in the immediate future, which includes databases, faculty profiles pages and AV support-- not enough of a demand.)



USEFUL >
COMPREHENSIVE

- Limited staffing
- Paying technical debt

MRN:

This also meant we quickly abandoned any notion of making the repository comprehensive.

First and foremost, with limited staffing to support capturing and ingesting new content, so we would never be the definitive source of research output for the campus.

Also, we knew the system could never have support for every possible use case; with a focus on paying technical debt, meant there was no room to develop 'wow' features or fully support niche or edge cases. (80-20 rule)



FOCUSED DEVELOPMENT

MRN:

This focus on 80% of users need want 20% of the services/features allowed us to focus our development efforts. We could explicitly drop support for unused services, allow other services to limp along, and improve our existing strengths.

INCREMENTAL CHANGES

- Articulated our boundaries
- Projects with a high return on investment
- Paying down technical debt

MRN:

With a refined scope, we had more teeth to say NO to features. We had articulated the boundaries of the system, which meant we no longer had to consider requests that were just for digital collections.

Our developers targeted low-cost, high-benefit projects. This included small changes to make the site more useful or easier to navigate; and large changes such as integrating with our library search interface. We also finished in-progress features.

We also started paying down technical debt by prioritizing maintenance and gem upgrades. While this is difficult to spin in a sexy, upbeat, look at us kind of way, we decided it was critical. We gave ourselves a 3 year window (we are 1 year in!) to pay technical debt down enough to start looking at what our next steps for the repository might be.

EASIER TO USE ->

INCREASED ADOPTION ->

MORE RESOURCES

MRN: By focusing on minor improvements and incremental changes that make the system easier to use, we could show users why this is a helpful tool, increase campus adoption, and eventually garner more resources (namely time and money) to continue improving the system.



CENTERING THE REPOSITORY

MRN:

Our last area of effort post mid-life crisis was centering the repository: which I define as meeting users where they are and finding our champions.

- ~~Faculty and graduate students~~
- Outreach and marketing
- Administrative professionals
- Internal advocates and allies

SUPPORT NETWORK

MRN:

This required identifying our supporters-- which luckily was an unintended side effect of a needs assessment: after interviewing faculty, students, and staff at the university to see what their needs were and any services we could easily provide (Or already provide!), we found that faculty and graduate students really weren't our support network (which we can discuss more if you'd like). Instead, our allies were outreach and marketing folks, administrative professional, and internal library advocates, primarily new hires (faculty and staff) who only knew CurateND as an IR.

MEET CHAMPIONS WHERE THEY ARE

MRN:

After identifying our support network, we found the centers and institutes on campus were our champions. And to meet them where they were, we articulated collaborative workflows and made modest improvements that made their lives easier. One example of this is an institute that, although faculty members are distributed across campus, wants to showcase all of their publications in one location. We facilitated the upload of content, exposed the records via an OAI-PMH endpoint, and our campus IT helped them harvest the records and make a feed to point back at CurateND.



SO WHAT?

MRN:

Ok, so what do you do when your repository enters middle age?

- Strong internal advocate(s)
- Mid-life crises can be an opportunity
- Use > Sunk costs
- Marketing 101

MID-LIFE
~~TURMOIL~~
REBIRTH

MRN:

First and foremost, strong internal advocates, like product owners but also at the administrative level, are crucial for revitalizing a system.

Instead of a mid-life crisis, or mid life turmoil, CurateND had a midlife rebirth. Even though the tech was difficult to maintain and had limited adoption didn't mean this tool wasn't useful or needed by campus partners.

Use is more important than sunk costs; sometimes, as was said in the presentation about design thinking on monday, you have to throw out the trash and keep what works. Even if it is painful.

Lastly, we inadvertently discovered Marketing 101 and learned the concepts everyone knows about-- Know your audience and know your champions and 2: product concept-- don't sell to them, show them how you can help/meet their needs. We used this to find our champions and keep them happy



THANK YOU, ND FAMILY

MRN: Before we sign off, thanks to everyone at ND (and in the samvera community) who has contributed to this system and our success (whatever that means); too many people to name, but hundreds of thousands of hours have gone into this system, and we are grateful for everyone's contributions

QUESTIONS?

Don Brower (dbrower@nd.edu)

Mikala Narlock (mnarlock@nd.edu, @mikalaraethelib)