DuraSpace Sponsor Summit   **Day 2, March 12, 2014**
DSpace Breakout

<u>Introductions</u>

Governance and Membership model

DSpace more of an outlier in that it conforms most to the new membership model?

Alternatives for representative governance

<u>Impressions from the group on membership and governance:</u>

- Asked for example of working groups – JM provided example from Fedora – steering group meets every other week, very demanding schedule for senior members of orgs, formed working groups to get specific tasks done. Now a resources working group – money, support, grants, also had a working group to work on governance. VIVO has a working group for ontology and one for apps and tools, they report back to steering group on a monthly basis
- Where would DSpace community (DCAT) report in this model? - JM – up for discussion, need to also consider vision group although that was considered temporary, these are hubs of communication
- Asked for impressions of workability
- Sarah Shreeves: representation from developing countries, membership could be a barrier. How do we bring those voices to the table? Not only third world, some smaller institutions in Canada who have contributed a lot to code.
- One member doesn't think this serves the DSpace community, leadership group has to be representative of the different layers, those who spend more get a place at the table but need representative at each level, elections at each level – pyramid shape, more money more seats.
- Have to understand how to represent the diversity of the community.
- Need to establish a mechanism to measure in kind contributions to balance in kind with monetary contributions.
- Need a 'tin' membership – in essence just get the communications benefits, we track you, you give us $500, includes smaller groups, should be able to submit requests.
- What control over the benefits does the steering group have? How do we superimpose this?
- Will not exclude anyone who contributes code. Steering group doesn't have resources to divert to development
- Fedora – if you donate more than a half time person that person will be directed by the project, a more managed process, something the steering committee could do. IF a group doesn't have money can contribute developer for same benefits.
- Try trial memberships for free and perhaps in a couple years will contribute.
- Tin Memberships – is this a way to engage third world institutions?
    - Could have consortial memberships for these groups for a small contributions
    - Benefits would have to be very clear, compelling
    - Could we use UN List or similar to recruit these members?
    - $500 still a lot for smaller institutions

- o In kind – training and education still missing. If people don't have developers on staff could they contribute other skills? Translation, training
  - o Also organizations like OpenAIRE – COAR(?) – aimed at developing countries, reach out to these groups
  - o QUESTION – what is the right amount and who should be allowed to come in at reduced amount?
  - o Resources within the UN that might want to support community building in these developing countries
  - o World Bank – could do a membership but not a sponsorship
  - o Need to focus on community building to engage International groups
- How to engage:
  - o Products
  - o Preservation and access of information in third world
  - o Capacity building – who is able to use what technologies
  - o Status – sometimes appealing to governments
  - o EIFLnet – good group to work with
  - o Role for consortia
  - o EIFLnet has one person on DCAT, already participating in conversations
  - o What is the role of DCAT?
- Is the intent of the model to increase monetary contributions to use as directed by the steering group or are we trying to engage technical contributions by more groups? Need to understand intent of governance model
- JM – DSpace community more diverse, complex – need to put some effort into investigating needs, could see how the project could benefit from project manager or director but don't have enough money to support this. What should we do about DSpace? What kinds of work does it need? Vision Group – many see themselves straining against the bounds of DSpace, majority of institutions not feeling that. We may find from survey, a large contingent that is satisfied to use DSpace in a more modest way. Does this imply more development or two separate strands? DSpace will require more money. At a disadvantage because we don't know what the 'blah' is.
- Have an opportunity to group some users together like OPENAIRE in Europe, could be a development working group based on contributions from that group. Could engage other international groups
- Some development activities not undertaken – ex. Implementing hibernate, has been attempted three times, requires big changes and testing, requires lots of human resources. DSpace considered older technologies because these kinds of features not implemented.
- DSpace organic and awesome but doing larger projects, cleaning things up, things that take more than three developers, doesn't happen. Need larger institutions to champion. We understand roadmap but don't have control over human resources, a team to devote to projects.
- Being able to fund something big, like Fedora, is a good model.
- Spin off a managed project
- All this work needs a project director or manager, GOAL to fund this position

- World Bank – example of group that could be engaged, may be interested in a generic solution for data preservation, would be interesting to talk to them, may be able to fund some kind of development work that could benefit community
- Project director – in kind support opportunity? Institution could pay salary, a two year stint.
- Model in PKP – consortia decided on one developer at a member institution worked with PKP to determine course and then wrote code.
- Fedora example – a person is donated, then the local institution's needs take precedence, work in two week sprints
- What would it take if we needed to do some significant re-architecting? Devote ppl resources. Would we be able to manage large changes? Yes, but would be better to have a project director to oversee development, avoid one-offs and create a longer-range vision for development. Need a roadmap for project director to sell this position.
- Project Director – business and community development role, reach out to groups like UN try to establish groups in developing countries
- Goal – raise enough money to hire a PM. Staff on project only 50-60%, never full time.
- How much money to make Tim full-time and hire PM - $200k, up to $500k total, some overhead involved too, JM's time, marketing, etc.
- Everything up to Tim's time is covered, need additional funds
- Need to structure governance to incentivize community to raise the $500k needed.
- Haven't come to terms with what we need in the way of development. Reconcile vision groups' needs with developing nations' needs - they need a voice in this. OCLC uses a regional model, one way to get full representation.
- How is communications about membership carried out beyond North America? – Targeted mostly first world outreach. Japan – issue getting communications constructed and translated in an effective way. Not clear that Japan does this sort of thing in a business/culturally. Is there a Japanese consortium? Same reps at Open Repositories, have spoken to them but it's unclear how to appeal to them. JM – could talk to SAKAI, they have an organization in Japan. SPARC also has a rep. Low level of sponsorship. A project director could be tasked with this.
- To get engagement from Japan, work with existing groups – get our governance right and show that it works.
- What do we expect new governance model to do? What are deliverables –
    - Outline these and tie back to better service, better code
    - More importantly think about how these tools work together in the larger ecosystem
    - Very rare for committers to think strategically, no one has time or wherewithal to do this
    - SHARE and other things going on in Europe, how does DSpace integrate, need strategic direction
    - See larger clients moving away from DSpace because there is no three year roadmap, unlike BePress – @mire
    - Need to define concretely the project manager's role
    - Who would hire PM and what is the timeline? – Like other projects – the role is in DuraSpace, dedicated to work steering committee determines, doesn't preclude a university funding salary, hiring process
    - Steering committee does oversight

- Need to define what we would do with PM position and then raise funds
- People want to know that if they give money it will be directed to what they want done

Have to agree on a governance model – interim steering committee should talk about these issues and figure out how to inspire other membership, talk about other revenue streams, DuraSpace can work with Steering Committee

- Action ITEMS:
  - Work with ranges of three levels of monetary commitment
  - Need a group – Steering Committee – to come up with model
  - Can't have three different models for each project, could have extensions for DSpace, needs a certain normalcy to structure
  - Steering group will be in charge of projects, interim piece, what will this look like when it's permanent
  - Still need to consider those who are under-resourced – three tiers with ranges, could work across all three projects, need agreement on in-kind contributions and how that compares with monetary contributions, every philanthropic group has these kinds of levels
  - Good governance more important than representation.

Second Session – How DSpace works – TD

How things work, where are the gaps?

- DCAT, Vision team, loose set of ambassadors that communicate work, don't meet, Steering Committee, people doing most work are the committers – diverse group around work, mostly in Europe, a meritocracy – committers vote you in, although anyone can nominate, DSpace built around same model as APACHE
- Code is submitted by contributors and prioritized by committers,
- Supporting Committers by DCAT, mostly repository managers, committers ask them to review requests to make sure they address needs
- DCAT helps with testing of releases
- DuraSpace – mostly Tim, his lead is mostly a facilitator – coordinates work of 19 institutions, does outreach in absence of project director, Val organizes DCAT group
- Who writes the code? - changed over the years, a growing number of contributors, only about half the contributors are committers, committers make sure their code is secure, good
- Last three releases most of the code given by non-DuraSpace committers, larger number of community members contributed to 4.0 release, process has become more collaborative, number of committers up but their contributions are down, more community contributions means a more organic process
- Github – a more social way of engaging code contributions
- Before Github, a more linear process, committers had to approve and integrate patches, with 4.0 more community members can comment on code and say if they've used code successfully, speeds process up.

- Community members contributions has increased to 25-30% - may not grow beyond that but it's important to the organic process
- Are there priorities from DCAT? – Yes, tell us what improvements are desired. Institutions vote by devoting resources to development, Tim can ask for institutions to devote resources for priorities set by DCAT
- May take several releases for change requests to bubble up to the top of the priorities – there is a running list of requests by DCAT, metadata example – DCAT long process as they mull over standardization, one of the challenges is that there is not a strong connection between DCAT and committers, modes of communication – international group and modes of communications make it difficult to communicate effectively – OR face to face really helpful but group size can be unwieldy
- Infrastructure gaps make metadata piece difficult to address – one piece that didn't require infrastructure change got done, others more complicated, speaks to issue of strategy
- Is revisiting the code base part of the strategy or should it be?
- Scaling? Can be difficult depending on how big the institution, not many experience this because collections aren't that big
- UI and infrastructure much different endeavors
- How does code make it to production?
  - Committers trust one another
  - New features must be approved by at least 2 members
  - Majority wins if there is a disagreement on priorities but committers have veto power on code
  - Committers split on UI code because different groups working on it
- Possible process benefits
- Dspace is truly community based
- Developers empowered
- Highly dependent on volunteerism
- Long term planning and implementation difficult
- Those with developers and committers have greater influence
- How would stronger governance model change this? TD – difficult to say – need to see what is more truly broadly needed to direct resources effectively
- What types of institutions contribute? US more larger public, MIT example of exception, but some smaller groups come out of woodwork, increasing number of European committers, a growing number in Europe, where activity occurs but not an indication of divide
- Would need to get buy-in from smaller institutions who contribute ppl sources for them to influence changes
- Committers glad for steering committee and governance, they would welcome a roadmap
- Need a group tasked with creating a roadmap

- 112 started survey, 83 completed it on SurveyMonkey, nothing surprising about geographic response rate, representative of contribution
- Demographics – 60% of users have used DSpace for over 5 years, version used most was 3.x (about half respondents), don't know how many contributors this includes
- How do you use DSpace, open access system for things like theses, etc., a third of institutions putting research data in DSpace
- Quarter for educational, anther for administrative materials, significant number using for digital collections
- Other platforms used (Primo, Summon, EDS integrations)
- Interest in ArchivesSpace, Fedora, Vivo, and a small group of Vireo users
- List of features important to respondents – create and manage metadata (already does this), don't support these features but should we (metadata for all levels of hierarchy of collections, relationships between objects, better support for hierarchical metadata) how are priorities being decided? Need strategy,
- What features should DSpace provide, relationships between items, better preservation, dataset archiving, ability to import and export data to spreadsheet (already exists), need to work on selling existing features ppl don't know about
- More configuration options
- Better API so we can use DSpace as a content management system but expose it elsewhere
- What does DSpace mean to you? – criticisms of old technology BUT open access, standard stuff, reliable, lot invested, frustration with metadata
- Biggest pain point – upgrading, no compound objects, old UI, difficult to install
- What now, define vision and project/product roadmap, plan to make it happen

DIVERSITY of DSPACE COMMUNITY

How does this drive or not drive development

- 20% installation in US, next Europe, Asia, South America (highest growth) some activity in contributor, biggest growth in contributions in Europe, "committership" greater in Europe, lots of money in European Union for building infrastructure.
- Decisions being made not representative – one camp wants an enterprise system, another wants an out-of-box system – how do we reconcile/manage the two. How do we know this? Many institutions install and leave it, older versions (these users didn't respond to the survey) do we need to worry about this group, if we evolve the roadmap to work for the more strategic organizations we want to bring some along and leave others (who are not interested) behind. Drupal example – can bring up a corporate website and leave or participate in development, OPENAIRE model – is this possible? Would have to make it really simple for one part of community and more complicated for another. Make DSpace more simplistic for installations but have plugins available for more sophisticated use.
- Could Tim kick-start position of PM? Still a funding issue. Tim doesn't feel he's the best fit for the role, could kick-start it. Needs to be focused on DSpaceDirect.

- What is the vision for the Project Manager – coordination of work, address the fragmentation in code, facilitate conversations about priorities, identify resources and organize communities around grant-seeking.
- Two camps – what are the characteristics of these two?
    - One is extending DSpace from original use – data, multimedia, (MIT), want to handle all rich content types, don't want the same UI to view content, Need for better API (validated by survey finding), enterprise group wants to do everything enterprise so there is one big repository to manage all content (Fedora attractive because it's more robust), DSpace reliable but not as flexible. These groups may trust that Fedora and Hydra will evolve to address needs. This shows why a roadmap is important so groups can determine if they can commit.
    - Some enterprise groups using a hybrid (Fedora, DSpace) system
    - Need to revisit code base – Fedora process for rewriting code could work – Andrew working on it full time, stakeholders commit a half an FTE, sign up for min 3 sprints, your FTE works fulltime on 2 week sprints, developers go in and out but remain consistent within the sprints, decision-making done by steering group, contributed $20-30k and one half FTE, technical decision-making Andrew + Steering committee
    - Fedora Model could work – critical difference, those involved with both have a heavy commitment to Fedora, this would mean institutions with more human resources would have greater influence, may not reflect all the voices of the community
    - If we had a roadmap would there be enough committers? Group thinks so, might mitigate only big institutions having influence
    - Some things would matter to committers but some things involved in re-writing infrastructure may be a more difficult sell because they are more complicated and might have a negative impact on use at first
    - Need a shared vision
    - Implementing hibernate could be done as a one-off
    - May be an opportunity for a planning grant – ARL, Sloan, Mellon
    - There are a ton of things we could do, we just have to organize
    - Who is the right group to pull in? Have sponsored vision groups but not always result in a consensus, need committers and higher level participants like managers, some senior people, MK – need institutional decision-makers
    - None of the vision groups have been tasked with writing a roadmap, need to be given authority
    - After roadmap, need to decide what needs to be done to the guts of DSpace, this discussion should be done at an in-person tech group meeting to determine how much work needs to be done.
    - DSpace is ten years old – a 50-60k planning grant could help to convene group, developing funds to implement is another, larger issue
    - Can get to $400-500k total membership with reasonable membership dues (90 members at different levels)
    - A good way to get code to institutions who can't afford it
    - Capacity for larger one-time funds and longer commitments (Gates Foundation?)
    - Need to justify funds to senior leadership at smaller organizations

- Smaller institutions want to keep doing work and have other options as long as the addition of newer features doesn't interrupt current work.
- DSpace for day-to-day stuff, Fedora for more complicated work
- Steering committee – no dollars attached, meant to be representative, if they are charged with devising roadmap how should the committee be organized? Current committee should decide on governance structure to make sure all voices are heard,
    - Planning grant
    - Roadmap – current committee draft a high level roadmap and get input from community and DCAT – timeframe – one piece needs to be presented at OR (group at OR will care more about this) goal to generate buy-in, share vision document to inform roadmap
    - Governance – need the roadmap first – a draft of this at OR
    - What about product manager role? Could we present a document that drafts their responsibilities and how much money we will need to collect? Also at OR, use to recruit members.
- URGENT NEED – steering committee plus some representation from DCAT, draft a roadmap – 2 day meeting to draft roadmap, need more than 5-6 points, some examples of tasks and what resources would be needed, explain what it means and why it matters
- How to tie DSpace in as a low-cost system into a more enterprise solution

JM=Jonathan Markow
TD=Tim Donahue
MK=Michele Kimpton