



Ershad Kunnakkadan &lt;ershad92@gmail.com&gt;

---

**GSoC'14**

9 messages

---

**Vidhun Vinod** <xvidun@gmail.com>  
To: ershad92@gmail.com

Wed, Mar 5, 2014 at 7:01 PM

Hey Ershad,

I'm not sure if it is a good idea contacting you privately, it was the last resort after futile attempts at trying the irc channel, maybe you could help me move in the right direction.

So anyway, I'm looking at contributing to SMC for this GSoC. I have been looking at the codebase for the Grandham and did some self research on the MARC21 bibliographical data formats.

I've been working with rails for sometime now, so I have some relevant experience :).

Also, is there some additional aspects of Grandham that require improvements I can work additionally for this summer(if there is some time left on my summer).

My [github profile](#), don't have some major contributions yet, this might be a start :)

I enjoy [stackoverflow](#).

Thanks for support,  
Vidun K.

---

**Ershad K** <ershad92@gmail.com>  
To: Vidhun Vinod <xvidun@gmail.com>

Thu, Mar 6, 2014 at 11:43 AM

Hi Vidun,

MARC21 is a widely accepted data format used by many library software. It's difficult to understand and parse such records, see if you can come up with some simple solutions to read and write MARC21 efficiently. It's okay to use existing libraries, but you should know how to handle different fields and leader variables. Good luck.

Cheers.

[Quoted text hidden]

—

Regards,  
Ershad K  
<http://ershak.com>

---

**Vidhun Vinod** <xvidun@gmail.com>  
To: Ershad K <ershad92@gmail.com>

Thu, Mar 6, 2014 at 6:40 PM

hey,

Oh it looks like there is a ruby [gem](#) for parsing MARC. Currently i'm looking into more resources to understanding the MARC-21 formats. As of now it looks very unstructured, hopefully I can understand something before I begin making my proposal.

[Link](#) seems like a good resource to start with.

So I'm in doubt of the efficiency, some of the [archives](#) of Marc-21 bibliographic data is over 3GB in size and also some of the archives I've seen seem to provide API, so I think it would be a better feature to also provide an API for access apart from importing Marc data.

Is there any additional resources that you can provide me with that would be helpful?

Regards.

[Quoted text hidden]

---

**Ershad K** <ershad92@gmail.com>  
To: Vidhun Vinod <xvidun@gmail.com>

Sat, Mar 8, 2014 at 1:38 PM

Hi Vidhun,

Grandham has APIs which provides basic information on books. The idea was to import MARC21 records from various sources and make that data available as an API. Also we need to have a MARC21 export feature for other library systems to use the data from Grandham.

I had tried that gem a few months back and it worked with the basic fields. I'm not sure how it would respond to the actual MARC21 data like the one you pointed in Harvard library.

The scope of the project includes building a MARC21 import/export feature in Grandham. That would require changes in the Ruby on Rails application we made.

Have a look at the tutorial ([loc.gov](http://loc.gov)) you mentioned and tell me what you think.

[Quoted text hidden]

---

**Vidhun Vinod** <xvidun@gmail.com>  
To: Ershad K <ershad92@gmail.com>

Sat, Mar 8, 2014 at 10:31 PM

I was looking in this, [authority](#) has the up to date complete documentation with examples on using MARC-21.

Key notes on features for Grandham.

1. Ability for Grandham user to create his MARC-21 reference for use in the software.

Eg. 100 \$a in MARC-21 refers to name, we keep some default value in database 'Name' while giving the flexibility for the user to call this Personal name, name etc, similarly for every thing other fields.

2. Using MARC data within Grandham:

As you mentioned "changes in the Ruby on Rails application", this project should make some workaround to migrate the current cataloging structure to use MARC-21 data structure in the database completely.

3. Export:

If we are to use MARC data inherently in the application, exporting will not be a problem. User can input what fields to export.

4. Import:

The number of fields used differ with libraries, so possible solution would be to analyze marc data to identify the fields present, then list the fields to user for selection as to what fields to import. Then import can be handled.

I was [playing around](#) with this gem with some sample MARC-21, I'm not sure of what kind of problems that can come with using this gem.

Please let me know if there are some errors or corrections regarding my approach.

Thanks!

[Quoted text hidden]

---

**Ershad K** <ershad92@gmail.com>  
To: Vidhun Vinod <xvidun@gmail.com>

Mon, Mar 10, 2014 at 10:16 PM

On Sat, Mar 8, 2014 at 10:31 PM, Vidhun Vinod <xvidun@gmail.com> wrote:

I was looking in this, authority has the up to date complete documentation with examples on using MARC-21.

Key notes on features for Grandham.

1. Ability for Grandham user to create his MARC-21 reference for use in the software.

Eg. 100 \$a in MARC-21 refers to name, we keep some default value in database 'Name' while giving the flexibility for the user to call this Personal name, name etc, similarly for every thing other fields.

Okay, what about the fields which are not present in the database, but exists in the MARC21 record?

2. Using MARC data within Grandham:

As you mentioned "changes in the Ruby on Rails application", this project should make some workaround to migrate the current cataloging structure to use MARC-21 data structure in the database completely.

Yeah.

3. Export:

If we are to use MARC data inherently in the application, exporting will not be a problem. User can input what fields to export.

I'm not clear about this. Can we choose the fields to export in MARC21?!

4. Import:

The number of fields used differ with libraries, so possible solution would be to analyze marc data to identify the fields present, then list the fields to user for selection as to what fields to import. Then import can be handled.

Our aim is to store all the data we get, we should have a solution to go about the extra fields. Thoughts?

[Quoted text hidden]

[Quoted text hidden]

---

**Vidhun Vinod** <xvidun@gmail.com>  
To: Ershad K <ershad92@gmail.com>

Mon, Mar 10, 2014 at 11:06 PM

Okay, what about the fields which are not present in the database, but exists in the MARC21 record?

There is a lot of metadata in Marc21 records that are mostly left unused, also it is mentioned in MARC21 docs from [loc.gov](http://loc.gov) that only 10% of the data fields is commonly used by libraries. So I thought it would make sense to allow flexibility to decide on what fields to use, this would also ensure resource and performance safety. Maybe the admin can add additional fields later on to use.

If a field is not present in DB we will not import. I'm open to discussions regarding this.

I'm not clear about this. Can we choose the fields to export in MARC21?!

I pointed earlier about the usability problem, so export only necessary fields leaving others empty.

Our aim is to store all the data we get, we should have a solution to go about the extra fields. Thoughts?

Although the admin can choose to import all the fields, it would be better if we gave the admin this choice as most of the meta-data the other library's MARC21 uses will not be relevant to Grandham in its context. Even at the worst case the admin might have to import all data which the system can also allow.

Although it would be easier to just use all of MARC fields but I feel that there is just too much meta-data that will probably never be used.

My approach poses another problem which is to allow dynamic database fields(tell user to rake db:migrate when required to add new field, maybe some other workaround?) or create DB with all of MARC fields([commonly used MARC21 fields](#) by itself is a lot). This has to be discussed.

If you feel that these points doesn't make sense please correct me.

With your feedback I have take the relevant points that will make it to the proposal.

Thanks for your time and feedback,

Best.

---

**Ershad K** <ershad92@gmail.com>  
To: Vidhun Vinod <xvidun@gmail.com>

Tue, Mar 11, 2014 at 1:43 AM

On Mon, Mar 10, 2014 at 11:06 PM, Vidhun Vinod <xvidun@gmail.com> wrote:

Okay, what about the fields which are not present in the database, but exists in the MARC21 record? There is a lot of metadata in Marc21 records that are mostly left unused, also it is mentioned in MARC21 docs from [loc.gov](http://loc.gov) that only 10% of the data fields is commonly used by libraries. So I thought it would make sense to allow flexibility to decide on what fields to use, this would also ensure resource and performance safety. Maybe the admin can add additional fields later on to use. If a field is not present in DB we will not import. I'm open to discussions regarding this.

Are you saying the admins will configure the fields first and the importer would import data for these configured fields?

I think we should import all the data in the given MARC21 record, store them efficiently in a system which supports new fields(a key-value pair may be) and show only the relevant information to the user for now. We need a system to store and process the complete MARC21 information.

I'm not clear about this. Can we choose the fields to export in MARC21?! I pointed earlier about the usability problem, so export only necessary fields leaving others empty.

I'm confused about this. If it's valid in MARC21 to export only the fields which have data, it's completely fine. We need to confirm this.

Our aim is to store all the data we get, we should have a solution to go about the extra fields. Thoughts? Although the admin can choose to import all the fields, it would be better if we gave the admin this choice as most of the meta-data the other library's MARC21 uses will not be relevant to Grandham in its context. Even at the worst case the admin might have to import all data which the system can also allow.

As I mentioned before, I think we need to store all the information available even if we are not using them for now. We will confirm this too with other people in the project.

Although it would be easier to just use all of MARC fields but I feel that there is just too much meta-data that will probably never be used. My approach poses another problem which is to allow dynamic database fields(tell user to rake db:migrate when required to add new field, maybe some other workaround?) or create DB with all of MARC fields([commonly used MARC21 fields](#) by itself is a lot). This has to be discussed.

We could keep a dynamic fields list (in db) and use them with the values in a different table with field's foreign key. We don't have to keep on adding new columns to a table.

If you feel that these points doesn't make sense please correct me.

With your feedback I have take the relevant points that will make it to the proposal.

Thanks for your time and feedback,

Best.

Can we CC this discussion to [student-projects@lists.smc.org.in](mailto:student-projects@lists.smc.org.in) ? Other people might like to pitch in with new ideas.

Cheers.

—

Regards,  
Ershad K  
<http://ershadjk.com>

---

**Vidhun Vinod** <xvidun@gmail.com>

Tue, Mar 11, 2014 at 2:14 AM

To: Ershad K <ershadjk@gmail.com>, student-projects@lists.smc.org.in

| I think we should import all the data in the given MARC21 record.  
This makes sense as it would be easier and more appropriate to import all data, or if he needed to list additional data later on he would have to import them all over again. We have to probably restrict the data to the view.

| I'm confused about this. If it's valid in MARC21 to export only the fields which have data, it's completely fine. We need to confirm this.  
As far as I know I think this is possible, anyway I'll look into this and let you know. Should probably play around a little more with the ruby-marc.

| We could keep a dynamic fields list (in db) and use them with the values in a different table with field's foreign key. We don't have to keep on adding new columns to a table.  
Great, thought of this but as I'm not sure how large the number of fields in MARC21 will be, probably do a little more digging. Will have to confirm with other people whether to use complete MARC21 or the most used parts first. I think we'll have to use everything considering that we have to import every information.

I will have to start working on my proposal, any suggestions or things I should know before I begin?

It's probably a good idea to CC [student-projects@lists.smc.org.in](mailto:student-projects@lists.smc.org.in) to get some fresh ideas and suggestions, isn't the IRC active?, it was not when I tried.

Best.