

2014年度NRDF作業部会報告

Annual Report 2014 of NRDF Working Group

北星学園大学経済学部
片山 敏之

KATAYAMA Toshiyuki
School of Economics, Hokusei Gakuen University

Abstract

The activities of the fiscal year 2014 on the Working Group of the Nuclear Reaction Data File (NRDF) of our data center (JCPRG) are reviewed. The XML format of the NRDF coding format and the new NRDF coding editor system are discussed this year and some prototype format and coding editor system are proposed.

1 はじめに

NRDF 作業部会は、本センター（略称は JCPRG）が管理する核データのデータベース NRDF を、構築、運営、利用、公開するために必要な作業および研究活動を行っている。NRDF 作業部会は昨年度の5月から活動を開始した。昨年度に引き続き、今年度も月に2,3回を目標に金曜日の午後2時から4時までを定例の集まりとし、その会合の名称を NRDF セミナーと呼んでいる。今年度は、第1回が4月25日で、8月前後の夏季休暇や学会関係の出張期間を除いて、3月中旬まで24回ほど開催された。

今年度の NRDF 作業部会の構成、主な活動内容およびいくつか提案された成果について報告する。

2 NRDF 作業部会の概要

NRDF 作業部会の目的および中心的な課題については昨年度の年次報告 [1] にある通り、「核データの新しい採録書式 (coding format) の開発、採録作業を支援する coding editor の開発、および NRDF の品質向上のために既存の登録済みデータベースのデータの正誤検査」であり、本年度もそれを継続している。NRDF 作業部会に関する連絡はメーリングリスト「nrdfwg@jcprg.org」で行われ、毎回の NRDF セミナーは本センターに集まって行われる。

2.1 今年度の構成員

2014 年度の構成員は以下の 7 名とオブザーバー（2 名）である。

北海道大学大学院理学研究院・本センター（JCPRG）

合川 正幸，江幡 修一郎，今井 匠太朗，加藤 幾芳

札幌学院大学

千葉 正喜

北星学園大学経済学部

能登 宏，片山 敏之

他に，オブザーバーとして，元北海道大学理学部の吉田 ひとみ，同じく大型計算機センターの芦沢 貴子が参加している。

2.2 NRDF マスターファイルに登録済みの論文のチェック作業

既存の登録済みデータベースのデータの正誤検査の作業として，各人コーディング済の論文 1 編程度を分担する。通常は，1 回のセミナー作業時間に数編をチェックし確認する。これに関しては，CHEN プログラムを実行して警告ログが出たデータファイルに対してのみ重点的にチェックした。

分担内容と分担者

物理確認：主にヘディング等，HENDEL で NRDF 独自の項目がある部分

片山，加藤，能登

形式確認：NRDF フォーマットのチェック

千葉，（吉田，芦沢）

HENDEL 反映・CHEN 確認：修正依頼を HENDEL に反映

江幡，今井

マスターファイル更新：CHEN 警告ログを確認し，NRDF マスターファイル更新

合川，（または NRDF セミナーの共同作業として実施）

3 2014 年度に検討された課題

NRDF 作業部会で検討された課題別に，NRDF セミナーの検討事項を述べていく。NRDF セミナーには都合で欠席する場合もあるので，その詳細についてはメーリングリストの記録を参考に作成した。

3.1 NRDF の XML 化

XML はデータ記述を含む文書を構造化し、機械可読にするための汎用マークアップ言語であり、XML で記述された文書は、異なるアプリケーションプログラムの間での中間的なデータ保存および検索の対象として一般的に利用されている [2]。

昨年度は NRDF の XML 化について、BIB セクション部分の XML 化を議論したが、今年度は、まずその検討結果の確認から作業を開始し、REACTION セクション部分と DATA セクション部分の議論を行った。これで一応 XLM 化の全体像ができたことになる。

(第 2 回 5/2～第 5 回 5/30, 第 13 回 10/24)

付録 1 に検討結果の XML ファイル (原案は合川) を掲載する。

なお、この課題についてはこれまで本センターで既にいくつか検討 [3] がなされている。

3.2 新しい NRDF エディタ・コーディングシステムの作成

NRDF 作業部会では、新しいエディタをローカルシステムと Web システムの両方で利用できることおよびユーザ側の OS に依存しないことのため、このシステムを作成するプログラム言語として Java を採用する。IEEE Spectrum のプログラム言語ランキング [4] に依ると、Java は、成長の度合い、企業側の必要性、オープンソース開発プロジェクトやソーシャルメディアの 4 つのそれぞれの指標で人気度評価がトップにランクされている。このことは、実用性やメンテナンス (即ち将来性) の観点でも Java によるシステムの開発は適切であると考えられる。

新 NRDF エディタの概念設計および作成手順を検討した。概念設計を 3 つのステップに分けて図式化したもの (原案・合川) を付録 2 に示す。

そして、作業手順の「Step2」に相当する作業として、XML 形式のファイルおよび関係データベースシステムから核データを検索する Java プログラムのサンプルシステムを作成した。また、採録作業のユーザインターフェースの検討を目的として、ローカルシステムとして稼働する Java プログラムも開発している [5]。

(第 6 回 6/13～第 10 回 7/25, 第 20 回 1/23, 第 21 回 2/6～第 23 回 2/27)

3.3 マスターファイル更新

今年度は、月 1 回程度、NRDF セミナーの時間内に、合川の主導で NRDF マスターファイルの検査プログラム CHEN を使い、その警告ログを NRDF セミナーの共同作業として確認し、即座に NRDF マスターファイルを更新する作業を行った。その場で訂正できる警告の他に、後でまとめて議論すべき問題もいくつか出てきた。

(第 14 回 11/7, 第 16 回 12/12, 第 19 回 1/16, 第 24 回 3/13)

3.4 その他

革新的研究開発推進プログラム (ImPACT) [6] 中の「核変換による高レベル放射性廃棄物の大幅な低減・資源化」(プロジェクトマネージャーは東芝・名大特任教授の藤田 令子氏) における本センターの役割と本作業部会の課題が時間的に優先すべきものとなった。(第 11 回 9/12～第 12 回 10/10)

NRDF セミナーではその他に、2014 年次報告の編集および原稿依頼、進捗状況の確認についての打合せ (第 15 回 12/5, 第 17 回 12/19), 本センターの外部評価に向けた自己評価の確認事項 (合

川) および DOI リストの作成状況の報告 (合川, 中川) (第 18 回 1/9, 第 17 回 12/19) などが議論された。

4 おわりに

今年度は新しい NRDF エディタ・コーディングシステムの作成について、実際の論文の NRDF コーディングデータを XML 化したり、エディタシステムのサンプルプログラムの作成など、かなり具体的な議論を積み重ねてきた。最後に、検討成果を箇条書きにしてまとめ (原案・江幡、今井、千葉) とする。

書式について

データベースを構築するにあたって、ラベルの用意とラベル付けを規定する。

データベースの意味づけは書式で決まる。

我々は基本的に物理的な意味でラベルを作りたい。

マスターデータベースからユーザーに合わせたデータベースの再構成などの利用も考慮する。

NRDF と EXFOR の書式スタイルについて再考

核データで最も重要なのは反応式と物理量を特定する物理的な条件である。

これ等をどの様に扱っているかを明示する。

NRDF の扱い

反応式 [Target(Projectile,Emitted)Residual] と物理条件を分けた書式を採用している。

- メリット
論文から読み取ったまま採録できる。採録がデータに特化した論文の要約になる。
採録したデータテーブルは各々の物理量で書く為、任意の物理量を採録できる。
- デメリット
データの扱いに難がある。すべての条件を読み取らないとデータを特定できない。

EXFOR の扱い

反応式と物理条件を分けない書式を採用している。物理量の特徴をすべてコード化している。[T(P,E)R,SIG,,, など]

- メリット
反応式を見たときに採録しているデータの物理量がわかる。データの取り扱いが容易。
- デメリット
物理量の特徴がコードで厳格に決められているので、新しい物理量を採録するときはコード辞書の再定義が必要。その為、採録活動が“採録”の専門家でなければ難しい。

各々の書式の長所短所を踏まえた上でどの様な書式を作るべきか

採録の書式は NRDF 的に行い、データの出力は EXFOR 的にしたい。

- 問題点は今までの NRDF の物理量の定義と EXFOR のコードが対応しない事にある。
- ← 一対一対応は難しいのでフィルターをかけていく要領で EXFOR コードを特定していく。
最終的に“数個程度のコード”まで絞る、これを“許容”する。

← フィルターのかけ方がそのままデータのラベル (書式) になる。

具体的に行う事

基本的に NRDF の書式形式を踏襲して進めたい。

1. 代表的な物理量をピックアップする: 5 ~10 個程度 (全断面積、微分断面積、Yield、etc.)
2. 1 で挙げた物理量と対応する、採録されている典型的な paper を各々で幾つか選択する。
(もっともシンプルなもの、条件が複数付属し複雑になっているもの)
3. データのラベルを選定していく。

このように纏めてみたが、いくつか留意しておくことがある。その1つは、検索やユーザーが必要なラベルが、上記までの NRDF の XML 書式で十分か、否か、まだ検討の余地なあること。例えば、書誌情報 (Title, Abst, Author, Affiliation, etc.) を細分化したり、キーワード (辞書) 化して検索しやすくすることが考えられる。更に、ImPACT の課題との関係で NRDF の DATA セクションの XML 書式の精密化と確定が急がれる。

参考文献

- [1] 能登宏, 「NRDF 作業報告」, 北海道大学原子核反応データベース研究開発センター年次報告 No.3, 25 (2013)
- [2] G. Powell, “Beginning XML Databases”, Wiley Publishing, Inc. (2007); Erik T. Ray, “Learning XML, Second ed.”, O’Reilly Media, Inc. (2003)
- [3] 椿原康介、松本琢磨、合川正幸、加藤幾芳, 「XML ベースの新フォーマットの提案」, 北海道大学原子核反応データベース研究開発センター年次報告 No.1, 23 (2011)
- [4] The Top Programming Languages by IEEE Spectrum’s 2014 Ranking, <http://spectrum.ieee.org/static/interactive-the-top-programming-languages>
- [5] A. Sarsembayeva, S. Imai, S. Ebata, K. Kato, M. Aikawa, M. Chiba, N. Otsuka, “Report on the status of IT enviroment”, JCPRG Annual Report No.4, 32 (2014)
- [6] 内閣府・革新的研究開発推進プログラム (ImPACT), <http://www8.cao.go.jp/cstp/sentan/about-kakushin.html>

付録 1

NRDF の XML 形式化 (原案・合川)

ファイル名 : xnrdf-sample.xml

<?xml version="1.0" encoding="UTF-8" ?>	</author>
<entry>	</authors>
<entryid>	<affiliations>
<nrdf>D0001</nrdf>	<affiliation no="1">
<exfor>E0001</exfor>	<institute>University</institute>
</entryid>	<division>Faculty</division>
	<code>2JPNJPN</code>
<bibliography>	</affiliation>
<references>	<affiliation no="2">
<reference no="1" main="1">	<institute>University</institute>
	<division>Faculty</division>
<title>Title</title>	<code>2JPNJPN</code>
	</affiliation>
<journal>	</affiliations>
<name>Physical Review C</name>	<purpose>Purpose</purpose>
<year>2013</year>	</reference>
<vol>87</vol>	<reference no="2">
<page>051001</page>	<title>Title</title>
<doi>10.1103/PhysRevC.87.051001</doi>	<comment>Secondary ref.</comment>
</journal>	</reference>
<authors>	</references>
<author no="1">	</bibliography>
<name>F.Author</name>	
<affiliation ref="1"/>	
<email>f.author@gmail.com</email>	
</author>	
<author no="2">	
<name>S.Author</name>	
<affiliation ref="2" />	

<pre> <compilation> <compilers> <compiler no="1">XXX</compiler> <compiler no="2">YYY</compiler> </compilers> <histories> <history no="1">XXX</history> </histories> </compilation> <experiment> <facilities> <facility no="1"> <institute>RIKEN</institute> </facility> <facility no="2"> <institute>CNS</institute> </facility> </facilities> <accelerators> <accelerator no="1"> <type>Cyclotron</type> <institute ref="1" /> </accelerator> </accelerators> <detectors> <detector no="1"> <type>NaI</type> <institute ref="1" /> </detector> </detectors> <methods> <method no="1"> <code>EDE</code> <detector ref="1" /> <detector ref="2" /> </pre>	<pre> </method> <method no="2"> <code>TOF</code> <detector ref="2" /> <detector ref="3" /> </method> </methods> </experiment> <reactions> <reaction no="1"> <equation>12C(p,gamma)13C</equation> <target> <nuclei>12C</nuclei> <thickness>12 mg/cm2</thickness> <comment>A natural carbon foil is used. Its thickness is 12 mg/cm2.</comment> </target> <projectile> <nuclei>Proton</nuclei> <accelerator ref="1" /> </projectile> <outgoingparticles> <outgoingparticle no="1"> <nuclei>gamma</nuclei> </outgoingparticle> <outgoingparticle no="2"> <nuclei>13C</nuclei> <detector ref="1">To identify particles</detector> <method ref="1">To detect energies</method> </outgoingparticle> </outgoingparticles> </reaction> </reactions> </pre>
---	---

<datasets>

<dataset>

<reaction ref="1" />

<status>Plotted in Fig.1</status>

<headings>

<heading>angle</heading>

<heading>d σ /d ω </heading>

<heading>error

d σ /d ω </heading>

</headings>

<units>

<unit>degree</unit>

<unit>mb/sr</unit>

<unit>mb/sr</unit>

</units>

of

<table>

<row>

<data>10</data>

<data>0.1</data>

<data>0.01</data>

</row>

<row>

<data>20</data>

<data>0.15</data>

<data>0.015</data>

</row>

</table>

</dataset>

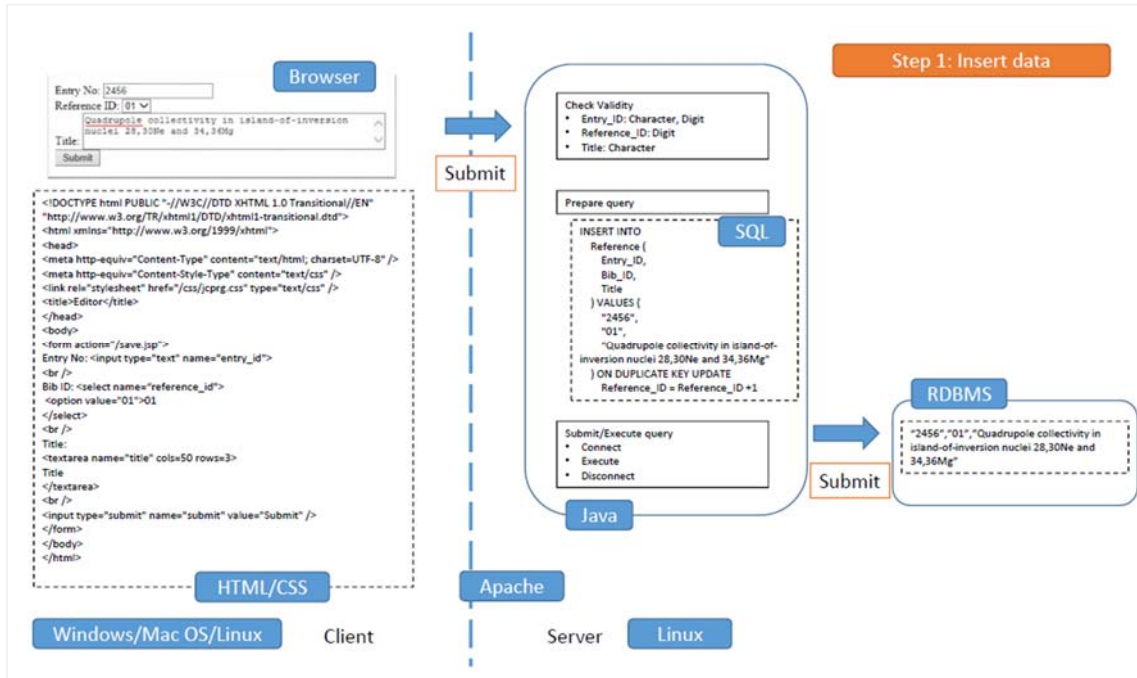
</datasets>

</entry>

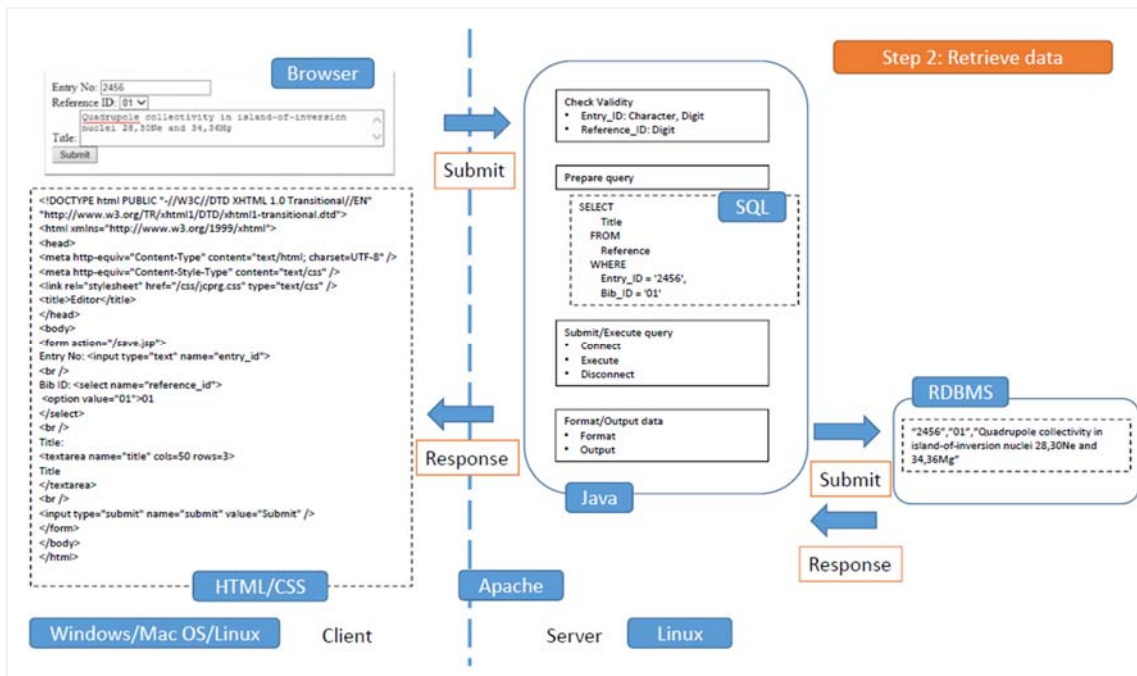
付録 2

NRDF エディタ・コーディングシステムの作業手順（原案・合川）

Step1: Insert data



Step2: Retrieval data



Step3: Output data

