

# 荷電粒子核反応データ採録活動における メールアーカイブシステム “Stock”

## Mail Archive System “Stock” in Charged-Particle Reaction Data File Compiling Activities

北海道大学大学院理学研究院<sup>1</sup>

鈴木 隆介

SUZUKI Ryusuke

Faculty of Science, Hokkaido University

### Abstract

A mail archive system “Stock” was developed for assisting in the NRDF (Nuclear Reaction Data File) compilation. This system started to be used in 2004 and over than 12,000 e-mails have been stored in “Stock” system. By using the system, we can track the coding processes and the discussions talked about in the NRDF compilation of each paper published in the journal on the charged particle nuclear reaction experiment, and share the relevant knowledge and details of our compilation. In this report, the overview of the “Stock” system is presented and its usages are explained.

## 1 はじめに

この報告では最近新たに開発された「電子メールアーカイブシステム”Stock”」についてその概要と使用法について説明する。

原子核反応データ研究開発センター (JCPRG) では、20 年以上に亘り日本の加速器で生産された荷電粒子核反応データの NRDF 形式による採録活動を継続し、これまでに蓄積されたデータ量は 3 万レコード、90MB にのぼる [1]。NRDF の採録の現場においては、このような成果物であるコーディングファイルの他にすべての論文に対して紙媒体の資料がファイリングされ保存されている。紙媒体の資料とは、コーディングシートや採録の際に用いられた論文のことで、採録者によるコメントやメモ等も記載されている。

他方、最近の NRDF 採録作業はインターネット上で遂行されるのが標準的となっており、採録の際に発生する疑問点・問題点や確認事項について採録者間を中心になされる議論や連絡も電子メールによって遣り取りされることが一般的である。この報告で説明する Stock システムは、前述のような紙媒体ではなく、電子媒体上で展開されている採録の作業工程、採録について採録者間を中心に行き取りされた議論、および、採録の元となったデータなどの情報を系統的に保存し、採録作業過程全般に関わる重要な了解事項や議論点を JCPRG で共有し活用し供して行くことを目的としたシステムである。

第 2 章では Stock システム作成の背景と動機、第 3 章では Stock システムの利用方法について説明する。第 4 章では Stock システムについての設計の詳細と管理法について説明する。第 5 章ではまとめを述べる。

<sup>1</sup>平成 19 年 12 月、北海道大学医学研究科放射線分野へ異動。平成 20 年 4 月より北海道大学病院分子追跡放射線医療寄附研究部門に所属。

## 2 Stock システム作成の背景と動機

どのようなデータベースであれ、比較的規模の大きなデータベースを構築するためには、データベースの作成、維持、管理に多くのエネルギーや時間、労力が費やされる。NRDF データベースは絶えず発展しつつある基礎科学の原子核実験データを対象としているため、データベースの構築がさらに難しく、千葉氏の言葉を借りれば [2]、

一度データ収集のフォーマットを決めれば、その後のデータ収集は自動的に進むというものではない。新しい実験による新しいデータが次々と生産されている。NRDF フォーマットはそのことを考慮して考え出されたとはいえ、一貫性をもってデータベースを編成していくには「どこを保存しどこを変えていくか」の的確な判断と創造性が絶えず求められる。

と説明される。事実 NRDF 形式が開発され 20 年以上たった現在でも NRDF 形式の採録方法や採録基準についての議論が辞書作業部会や管理運営委員会等の場で活発に行われている [3]。NRDF 形式は、千葉氏が述べているように、新しい実験やデータに対応する目的や、多種多様な利用への対応のために柔軟な設計がなされており、採録方法や採録基準については必ずしも厳密な基準がはじめから与えられているわけではない。NRDF の採録方法の代表的なコーディングマニュアル [4, 5] や、最近の採録において利用される採録エディタ HENDEL のマニュアル [6] においても、細かな採録基準は掲載されていない。NRDF の採録では、採録に直接関わる現場の作業者が一番最初に論文を理解し、判断を求められる立場にある。このため、作業者レベルでの採録方法の理解が大変重要であり、作業者全体としての共通認識が必要である。

これまで JCPRG で残されてきたコーディングファイルや紙媒体の資料だけでは、これまでにどのような理解や判断により採録が行われていたかを十分知ることはできない。過去に行われた採録の見直しや再採録の際に、現在残されている資料だけでは、なぜそのようなコーディングが行われたかが不明瞭なコーディングファイルに出会うことがある。

著者は当時から既に主流となっていた、採録にまつわるさまざまな議論を電子メールで遣り取りする点に目をつけ、そのメールを系統的に収集・保存することで、放っておくと埋もれてしまうその時々における採録に関する経験や知識を JCPRG で共有し、今後への情報提供を目的とする Stock システムを開発することにした。

Stock システムは、2004 年に開発され、現在 12,000 を越える数のメールが貯められている。また、JCPRG のメンバーからのメールだけではなく、論文の著者から直接 Stock システムに送られてくるメールも増えている。このシステムはコーディングファイルなどの紙媒体の資料同様に、NRDF データファイルそのものとしては表には出て来ないが、NRDF の採録活動を影で支えるインフラの一つとなっている。

## 3 Stock システムの利用

### 3.1 採録の現場

Stock システムでは、stock@jcprg.org のアドレスに送られてくる NRDF の採録に関するメールを対象とする。北海道大学原子核理論研究室では、stock というエイリアスが作成されているので、研究室のメールサーバを使う場合は stock 宛のメールも対象となる。メールを書く際には、タイトルに論文の D 番号 (例えば “D1900” または “d1900” といった文字列、または EXFOR の番号である



図 1: Stock Viewer 起動時の画面



図 2: 論文に関連したメール一覧の表示 (図 1 から d1801-d1900 を選択した画面)

“E1900”, “e1900” といった文字列) が入っていることが必要となるが、タイトルの中であればどこでもかまわない。もし 2 つの論文に関連する場合は、2 つの論文の D 番号が入っていることが必要である。送られたデータは Stock システムに D 番号ごとに分類され保存される。

### 3.2 情報の閲覧

Stock システムに貯められたメールは Stock Viewer と呼ばれるシステムを使って閲覧することができる。Web ブラウザで <http://www.jcprg.org/stock/> にアクセスし、認証のためにユーザ名、パスワードを入力すると、図 1 のような画面が表示される。

画面構成は上から順にヘッダ部、ブックマーク部、メイン部となっている。ヘッダ部は、タイトル、stock view と呼ばれるフォーム、検索フォームから構成される。ブックマーク部は、採録に関連したページへのリンクから構成される。メイン部では、論文が D 番号で 100 ごとに分類される。メイン部の一番右側の列は、採録する論文とは直接関係のない特別な分類の一覧である。現在、ntx-wg, jcprg-sc, jcprg-ac, coder, coe, stock, gsys, nrdc, other の 9 個の分類項目がメール蓄積のために用いられている。ntx-wg, jcprg-sc, jcprg-ac, coder, coe には JCPRG のメーリングリスト、stock, gsys には、それぞれ、Stock システム、グラフ数値読み取りシステム GSYS [7] に関しての議論内容が集められ、nrdc には、核反応データセンターネットワーク (Nuclear Reaction Data Center network) に関する情報が集められている。stock 宛のメールのタイトルにこれらの単語が入っていた場合は D 番号が記述されているメールと同様に分類される。タイトルにキーワードが入っておらず、分類ができないメールは、other に集められる。

メイン部から希望の論文が含まれている分類を選択すると、図 2 のような画面が表示される。もし、図 1 において “all” (本体部分、2 列目一番下) を選ぶと、Stock システムに貯めてあるすべての論文の一覧が表示される。

一覧からある論文 (d\*\*\*\*) をクリックするとその論文に関連したメールの一覧が図 3 のようにメイ

ン部に表示される。ヘッダ部には、D 番号の他に論文へのリンク、採録者、チェック者、査読者の情報、進行状況等が表示される。この情報は、JCPRG が管理している表から作成されているので、常に最新の情報が表示される。ブックマーク部には、その論文に対する採録エディタ (HENDEL)、NRDF、EXFOR のファイルへのリンクが追加される。



図 3: メール一覧の表示 (図 2 から d1900 を選択した画面)

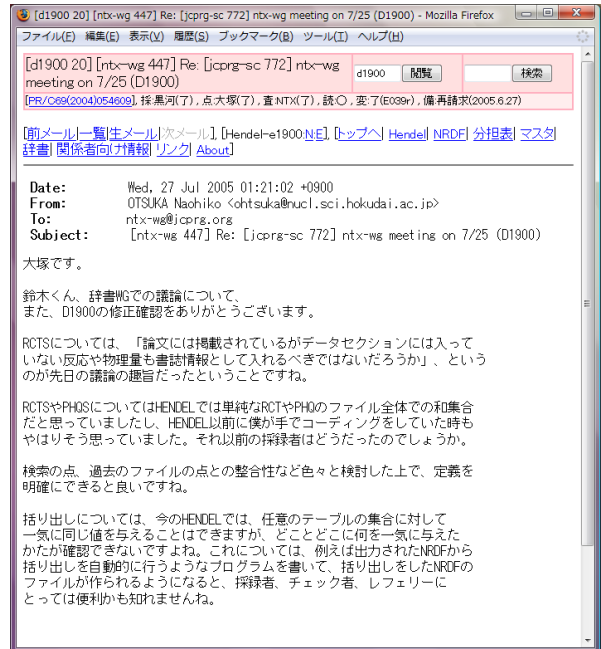


図 4: メール本文の表示 (図 3 から [d1900 20]... を選択した画面)

このメール一覧から、見たいメールを選択すると、図 4 のようにメールの本文が表示される。ヘッダ部にはメールのタイトルが表示され、ブックマーク部には現在表示されているメールの前後のメール、その論文に関するメール一覧、および、生メールへのリンクが追加される。

添付ファイルがある場合は、図 5 のようにメール本文の下に添付ファイルへのリンクが表示される。

現在では、800 近い論文に関するデータが Stock システムに保存されている。JCPRG は現在までに 1800 近くの論文を NRDF 形式に採録しているので、半分近くの論文に関する採録情報がこのシステムに保存されていることになる。

### 3.3 stock view および URL からの利用

ヘッダ部の右側にあるフォームを stock view と呼ぶ。このフォームに d1900 といった番号を入れ、**閲覧** ボタンを押すとその D 番号に対応するメール一覧が表示される。また、d1900\_1 という形で入力すると、メール本文を表示できる (この例では d1900 の論文の 1 番目のメールが表示される)。d1900\_1-10 と入力すると、1 番目から 10 番目までのメールの一覧、d0001-d9999 と入力すると D 番号の一覧が表示される。

URL を直接入力してアクセスする場合は以下のように利用する。

**【例 1】** <http://www.jcprg.org/stock/stview.cgi?d0001-d9999>

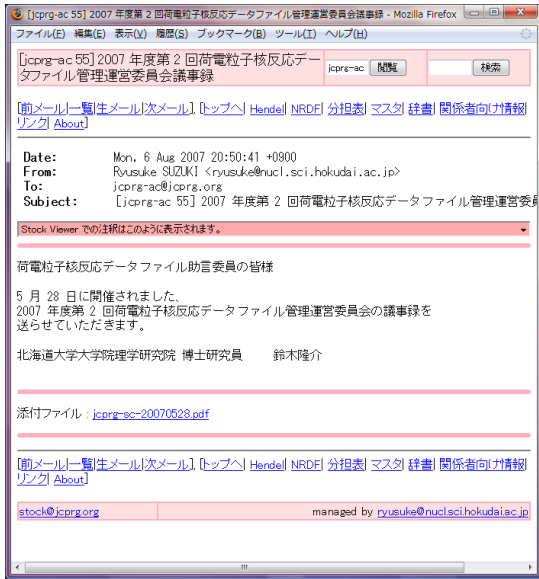


図 5: 添付ファイルの取扱いについて

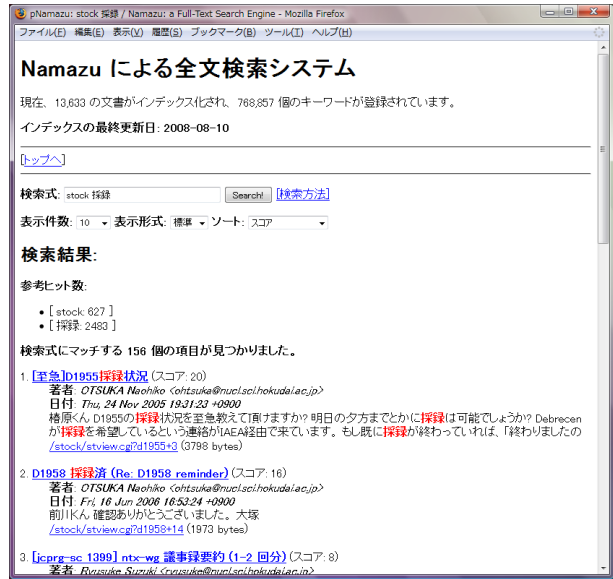


図 6: 全文検索システム NAMAZU による検索結果

【例 2】 <http://www.jcprrg.org/stock/stview.cgi?d1900>

【例 3】 <http://www.jcprrg.org/stock/stview.cgi?d1900+1-10>

【例 4】 <http://www.jcprrg.org/stock/stview.cgi?d1900+1>

【例 1】では D 番号の一覧が表示される。【例 2】の例では、D1900 番の論文に関するメールの一覧が表示される。【例 3】では D1900 番の論文に関する 1 番目から 10 番目までのメール一覧が表示される。【例 4】では D1900 番の論文に関する 1 番目のメールの本文を見ることができる。

### 3.4 情報の検索

Stock システムでは貯められたメールの情報を検索することが可能である。Stock システムでは NAMAZU (および、perl 版の pnamazu) [8] と呼ばれる検索システムを採用している。図 6 が検索結果である。検索対象については、本文だけではなく、一度閲覧したあとの添付ファイルであれば、pdf の内容についても検索が可能である [次節 4.1 の tmp ディレクトリ配下のファイルの展開を参照]。

## 4 Stock システムの詳細と管理法

### 4.1 Stock システムの構成

Stock システムはサーバに届いたメールを振り分ける 150 行程度の perl スクリプト (stock.pl) および、システムに貯められたメールを閲覧する際に用いる Stock Viewer から構成される。Stock システムの概略図を図 7 に示す。

Stock Viewer は stview.cgi と呼ばれる perl スクリプトおよび system サブディレクトリ配下に置かれる設定ファイルから構成される。このプログラムの実体は 1,500 行程度の perl で書かれた CGI である。

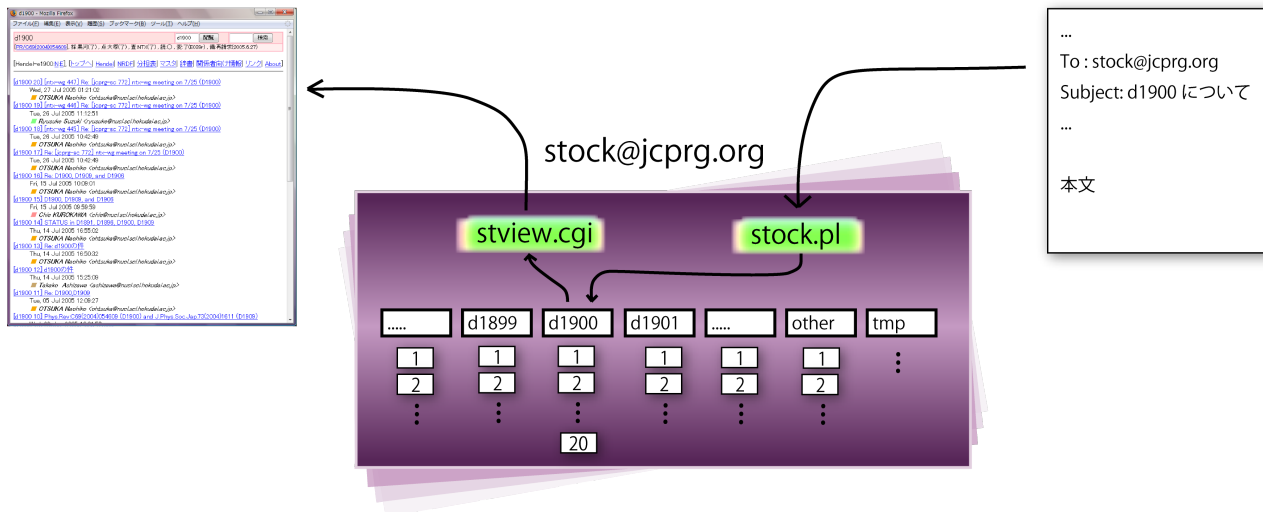


図 7: Stock システムの概略図

メールが stock@jcprg.org に送られると、図 7 のように stock.pl スクリプトが、メールの Subject からのディレクトリにメールを振り分けるかを判断し、そのディレクトリにメールを一つのファイルとして保存する。図 7 の例で言うとメールの件名に d1900 という文字列が入っているので、stock.pl は、d1900 ディレクトリ配下にメールを保存する。d1900 ディレクトリにはすでに 20 通のメールが保存されているので、このメールは 21 という名前で作成される。

D 番号のつかない特別なサブディレクトリも同様に扱われる。

また、tmp 配下のファイルは、メールの添付ファイルの展開用に作成されるものである。検索時には、このディレクトリ配下のファイルを添付ファイルとして検索する。

逆にブラウザからの要求に対しては、stview.cgi (Stock Viewer の実体) が www サーバから呼び出され、要求に対応する。

## 4.2 Stock システムの管理

### 4.2.1 設定の変更

システムの設定は system サブフォルダ配下のファイルにより設定される。

#### メール送信者の色を変更

図 3 で表示されているメール送信者の色については、color.user ファイルを書き換えることで変更ができる。

#### 特別な分類の設定

Stock システムでは D 番号の他に特別な分類を作成することができる。Stock Viewer では、このような情報を list ファイルで扱う。ただし、タイトルから適切にメールが振り分けられるように stock.pl を設定する必要がある。

## ブックマークの設定

ブックマーク部に表示されるリンク先は bookmark ファイルにより設定される。

## メールに注釈を付ける

stview では、図 5 (メール本文中のタイトル以下に表示) に示されるように、メールに注釈を付けることができる。comment ファイルにてコメントの追加、削除が可能である。

## リンク

この報告では説明をしていないが、Stock Viewer では、ブックマーク部とは別に、採録に関連するリンカー一覧のページがある。このリンカー一覧は link ファイルにて設定される。

### 4.2.2 その他

Stock システムでは、検索システムとして、NAMAZU を採用している。NAMAZU を利用するには、検索用のインデックスファイルの作成が必要であり、現在では、一日一回 cron を用いて、インデックスファイルを更新している。

また、メールは MH 形式 (1 メール 1 ファイル) で貯められているので、もしメールの分類が間違っていたり、必要のないファイルがあった場合には、MH のコマンドを使ってメールの管理を行うことができる。

## 5 おわりに

前章までで Stock システムの利用についての簡単な説明を終える。Stock システムの作成当時、著者は採録者として活動していたが、採録方法の情報や自分の担当する論文以外についての採録過程や議論などの情報を得ることが大変難しいと感じ、採録の議論などを共有できるような仕組みを作りたいと考えようになった。情報収集システムを採用するプロジェクトは世の中に数多くあるが、情報が更新されなくなったシステムをかかえたプロジェクトはアクティビティの低下という負のイメージを持たれる危険性がある。Stock システムにおいては、採録作業全般に関する重要な了解事項や議論点、問題点等の情報(「採録情報」と呼ぼう)をあらためて新規に生成するのではなく、現在も継続的に進行している採録者及び採録者間で展開されている採録活動の実態そのものを、採録情報としてそのままアーカイブ化することによって、持続可能なシステムとして存立することを目指した。JCPRG からの指示はなかったと思うが、当時から採録の現場においては遣り取りがメールで行われていたので、当時行われていた採録の遣り取りのメールをタイトルで分類し収集する方法を考えた。この方法であれば、採録が行われれば大まかな分類で必ず情報が集まり、採録情報を貯めるという目的のために採録者が採録以外に余分に行う作業は少なくすむ。Stock システムの導入後、採録者にとって増えた作業はタイトルに D 番号を入れるということと、宛先に stock と書くことだけであった。著者は 2005 年後半ぐらいから、採録の直接の現場から離れることになったが、現在でも採録に関わるデータを集めることに成功している。集められたメールの数を単純に平日の日数で平均すると、1 日あたり約 10 通のメールを誰かがどこかで作成していることになり、これは JCPRG のアクティビティの高さを表していると考えられる。

現在では、大変ありがたいことにメールを書く際に stock 宛に送ることが採録者に受け入れられているように感じられるが、採録者には Stock システムにメールを残すことは意識せず、採録に集中し

ていただくことを期待している。他方、採録関係者の参照に供したい、或は、何か残したいと思われるものがあれば、採録に関わるものは何でもよいのでとりあえずは stock 宛に送って頂きたい。単なる報告のメールでも採録の検証の際に助けとなると考えている。情報の価値は時代とともに変わるので、残せる情報は少しでも残しておけば将来の助けとなる可能性がある。NRDF の採録の品質向上に Stock システムが少しでも役に立てば幸いである。

## 謝辞

Stock システムを使っていたいただいた皆様に感謝しています。また、“stock” の名付け親であり、初代 stock.pl 実装者の合川正幸さんに感謝します。当時は海のものとも山のものともわからない stock のアイデアを実装してくださいました。

Stock システムを褒めてくださった方々に感謝します。特に辞書作業部会で能登先生が「おもしろい」と褒めてくださったこと、メール一覧で送信者に対し色を導入した際、黒河さんが褒めてくださったことはとてもよく覚えています。

最後に、このシステムに密接に関わって頂くことになった大塚直彦さん、吉田ひとみさんに感謝します。stock 宛のすべてのメール (12,000 以上ものメール!) は、JCPRG のサーバに送られると同時に彼らにも送られていました。現在でも Stock システムが使われているのは彼らのおかげです。

## 参考文献

- [1] 吉田 ひとみ, 加藤 幾芳「資料：2006 年度入力データ」(荷電粒子核反応データファイル年次報告 2006 年第 20 巻 [1996 年 3 月]) p.108.
- [2] 千葉 正喜「はじめに-若い力に期待-」(荷電粒子核反応データファイル年次報告 2002 年第 16 巻 [2003 年 3 月]) p.1.
- [3] 管理運営委員会の議事録は、各年度の荷電粒子核反応データファイル年次報告を参照。辞書作業部会の内容については、荷電粒子核反応データファイル年次報告 1999 年第 13 巻以降を参照。
- [4] 手塚 洋一「NRDF コーディングマニュアル (ver.1.1)」(荷電粒子核反応データファイル年次報告 1995 年第 9 巻 [1996 年 3 月]) p.71.
- [5] 向井 重雄, 吉田 ひとみ「コーディングに対する補足と注意 (1995 年改訂版)」(荷電粒子核反応データファイル年次報告 1995 年第 9 巻 [1996 年 3 月]) p.86.
- [6] 大塚 直彦「ウェヴエディタ "HENDEL" を用いた核データ採録入門」(荷電粒子核反応データファイル年次報告 2001 年第 15 巻 [2002 年 3 月]) p.12.
- [7] 新井好司、蓑口あゆみ、大塚直彦、内藤謙一「GSYS : グラフ数値化システムの開発とその利用法」(荷電粒子核反応データファイル年次報告 2004 年第 18 号 [2005 年 3 月]) p. 78.  
伊藤 慎他、鈴木 隆介「グラフ数値読み取りシステム (Gsys2.2) 利用の手引き」(荷電粒子核反応データファイル年次報告 2006 年第 20 巻 [2007 年 3 月]) p. 39.
- [8] 全文検索システム Namazu (<http://www.namazu.org/>)  
全文検索システム Namazu の perl 版検索プログラム pnamazu (<http://www01.tcp-ip.or.jp/~furukawa/pnamazu/>)