



よしだともこの Linux 事始めの書

第6回 UNIXコマンド使いへの道 その4 sed徹底活用

大人になると「あれから1年なんて早いなあ……」と思うことが多いですね。1年前の私は、京都で開催された「Linux Conference '98」の運営スタッフとして、最高に盛り上がっていました。その時の出会いや感動を胸に、この1年、走り続けられたような気がします。

よしだともこ <http://www.tomo.gr.jp/>

My "Happy イベント 積極的参加 Life"

先日、大阪で開催された、来場者500名規模の「'99インターネットと教育フォーラム¹」で、司会をしました。このフォーラムは、「情報教育の新展開～インターネットと情報倫理～」をテーマに掲げ、インターネット利用の「影の部分」ともいえる、セキュリティー、モラル、プライバシーなどに関わる情報倫理について、講演や具体的な教育実践の紹介を通じて積極的に検討しようという、日本初の試みでした。

主催は「インターネットと教育フォーラム実行委員会」で、共催は、インターネットと教育に関係している4つの団体²。そして後援には、文部省、通商産業省、郵政省、大阪府教育委員会、大阪市教育委員会、(社)日本教育工学振興会、情報処理振興事業協会、全国国立大学附属学校連盟・同PTA連合会、マルチメディアエデュケーションフォーラムというように、非常に多くの分野に渡る団体が名前を連ねていました。ただ名前を連ねているだけではなく、日曜日だというのに、ちゃ～んと(笑)お役所の方々が来てくださり、最前列に用意した来賓席に座っておられました。

実は今回のこのフォーラム、開催のアナウンス後たった1週間で申込者が会場定員数に達して、その後の受け付けをお断りしたという盛況ぶり。全国各地から(北海道からも沖縄からも)、小中高校の先生方を中心とした関係者が多数参加しておられ、「小中高校、特殊教育諸学校でのインターネット教育の注目度のすごさ」を感じました。そういう背景もあり、2000年3月11日(土)に、東京の早稲田大学を会場に、来場者1000名規

模のフォーラムを追加開催することが決まっています。こちらについて詳しくは、「K12³インターネットと教育」研究協議会のページ(<http://www.k12.gr.jp/>)を参照してください。

さて、このフォーラムのプログラム構成は、(1)情報科学研究者および情報倫理学研究者による講演、(2)小中高校・特殊教育諸学校の教員による教育実践報告、(3)インターネットと教育に関連する集中ディスカッション「児童・生徒全員に電子メールアドレスを発行するべきか否か?」でした。

今回の各種講演、発表などの中で、特に私が印象に残ったフレーズは、次の2つでした。

- ・早稲田大学の後藤滋樹先生の「萌芽的研究は、周囲から適切な反応が与えられないと、研究者自身、重要性を正しく認識できない」
- ・大阪市立瀬田小学校の石原一彦先生の「今後は、“How to use”ではなく“How to think”を学べる情報教育にすることが課題」

ちなみにこれらは、私の耳に残ったフレーズで、これらの先生方の講演や発表の概要ではありませんので、お間違いないように。で、これらの点について、自分が関係することから改めて考えてみたいと思います。

萌芽的研究に周囲から適切な反応を得るためには……

早稲田大学の後藤滋樹先生は「日本のインターネットの歴史と教訓」というタイトルの講演の中で、「日本は、米国からのインターネットの拡大を後追いしたように書かれることがあ

¹ 大阪で開催された「'99インターネットと教育フォーラム」については、<http://forum.k12.gr.jp/>ページ参照。また、<http://www.ascii.co.jp/ascii24/issue/1999/1130/topi02.html>、<http://www.ascii.co.jp/ascii24/issue/1999/1130/topi03.html>には、当日の内容が詳しく報告されている。

² インターネットと教育に関係している4つの団体のURLは、次の通り。<http://www.k12.gr.jp/>(K12³インターネットと教育研究協議会)、<http://www.fine.chiba-u.ac.jp/>(FINE)、<http://www.cec.or.jp/>(財団法人コンピュータ教育開発センター)、<http://www.jeric.gr.jp/>(JERIC)

るが、事実は多少異なる。WWWの先駆となった『gopher』³に類似したシステムや、Javaに先行するような遠隔言語の研究は日本でも行われていた。しかし、日本では大規模に展開されなかった」と述べられました。

確かに日本でも、同じ時期に同じようなシステムが開発されていました。例えばJavaが出たときに、私は身近な人々と「これって画期的だと言われているけど、Wnnとペアで開発されたGMW Ver.3(GSB)のG言語と同じじゃん。G言語は1989年に形になってたというのが、世の中がついてくるには早過ぎたのね……」なんて言ったものでした。

こういう現象について後藤先生は、「研究者自身、周囲から適切な反応が与えられないと、自分の研究の重要性を正しく認識できないからだろう」と分析されていました。本当にその通りだと思います。「開発されるものが先駆的過ぎて、日本でその価値がいまいち理解されずにいるうちに、外国で同じようなものが開発されて世界的に普及し、結局、日本人はそれを後追いしてる……」なんて、悲しい構図過ぎますよね……。

で、私なりに考えた改善提案は、「日本における開発も、最初から英語で世界に向けてガンガン発表し、世界からその価値を認めさせる」ということ、さらに、「日本の研究者の先駆的な研究成果に適切な反応が出せるように、評価する側の目を養ったり、環境を整えたりする」ということです。というのは、開発現場と評価する側の環境が違うと、例えば、専用線を大前提に開発されたものをダイアルアップで評価したりすると、残念ながら正当に評価できないんですよ。

また、特にオープンソースの場合、「Wnnってすごいよね、Namazuってすごいよね、Rubyってすごいよね」などと、無邪気にはしゃぐような(笑)の存在や、「すごいと報道されていることを、右に習えと、すごいと報道する」マスコミの存在も、場合によっては、うまく利用できるでしょう。その結果、期待度がより高まったり、ユーザーからの反応が多くなったりして、その相乗効果で、もの自体がより良くなるのが理想ですね。最初は、たとえそれほどすごくなくても、みんなが「すごい、すごい」と言っていると、本当にすごくなるという構図(-)。

「最初から英語で世界に向けてガンガン発表し……」は、私を含む平均的日本人にとっては、その重要性は分かりつつも、後回しにしてしまいがちなことのナンバーワンではないでしょうか。でも、Linus氏がLinuxを発表する投稿ニュースを、最初から英語で書いたことや、英語を書いたり話したりすることを(たぶん)苦にしていなかったことは、Linus氏が、専用線が使える環境にいたことと同様に、Linuxの発展において必須だっ

たと、私には思えるんですよ。それに、Linuxだって、ひょっとしたら世界中の人が「すごい、すごい」と言っているうちに、相乗効果で本当にすごくなった例かもしれないし……。

ちょっと余談になりますが、英語、日本語間の翻訳に関しては、Linux環境に強力な助っ人が登場しています。本誌の前号の特集で紹介された、「翻訳の王様(日本IBM社)」と「翻訳魂(オムロンソフトウェア社)です。諸事情から(謎の笑)私は「翻訳魂」の方を使い始めました。そこで嬉しがって、「オープンソースと手作りぬいぐるみ」概要⁴の重要なフレーズを、「翻訳魂」で翻訳させてみました。

[日]しかし、通常、ソースコードなどを目にしない多くの人々にとって、オープンソースの世界は理解しにくいように思えます。

[英]However, it seems to be hard to understand and can consider the world of open source for a lot of people who do not see source code usually.

[日]もっと一般的な例で、オープンソースの概念や楽しさを紹介したい」と思って始めたのが、Open Source Toys Projectです。

[英]It is Open Source Toys Project that I think that "a more general example wants to introduce a general idea and pleasure of open source", and began.

いい線、いってますよね。ま、私の書く日本語は、よく「英語を日本語訳したような文だなあ……」と言われるので、こういときには都合が良いと思います。この雑誌が書店に並ぶころには、「オープンソースと手作りぬいぐるみ」の概要(英語版)⁵が公開できていると思いますので、お楽しみに……(と書いておけば、ちゃんと英語にするだろう>自分)

ちなみに「翻訳魂」はパッケージに「対応するOSは、LASER5 Linux 6.0、Vine Linux 1.1、TurboLinux 4.0、FreeBSD 2.2.8、FreeBSD 3.2」と書かれています。つまり、LinuxはRed Hat系に対応しているということです。私が普段、原稿書きに使っているのはSlackware系なので、LAN上のTurboLinux日本語版4.0にインストールして、Slackware系からそちらにインストールされた翻訳サーバにつないで使うことにしています。が、Plamo Linux上に、手動でインストールして使えるようにした方々もいらっしゃいます。まずは、Vine Linuxに自動インストールして、どこにどのようなファイルが作られるかを参考に、Plamo Linuxに手動でインストールされたそうです。尊敬……。

³ gopherの開発元であるミネソタ大学を訪ね、gopherの開発者にその背景を聞いた記事が、「よしだともこのルート訪問記書籍版(ソフトバンク発行、ISBN 4-7973-0888-5)に掲載されている。gopherについて、詳しくはこちらをどうぞ(宣伝モード)

⁴ http://www.tomo.gr.jp/ost/1c99_llug.html

⁵ http://www.tomo.gr.jp/ost/1c99_llug_english.html

✂ “How to think”を学ぶ 情報教育にはLinuxを……

次に、天津市立瀨田小学校の石原一彦先生が「インターネットと情報倫理 小学校 / 盲聾養護学校の部」のまとめにおいて、「学校のネットワーク接続を前に、まず倫理的な判断力や、科学的な思考力をつけさせ、“How to use”ではなく“How to think”を学べる情報教育にしていくことが課題であろう」と言われたことについて、コメントしてみたいと思います。

現在、多くの小中高校、そして大学のクライアント環境としては、Windowsパソコンが導入されています。コンピュータ環境の整備にしっかりとお金が使える一部の大学や一部の小中高校を除いて、ほとんどの場合、少ない予算の中で、環境が構築されています。構築担当者や運用担当者が現場の先生自身であることも、小中高校の場合、当たり前です。そして、環境が導入された後は、バージョンアップの予算が見つからないことがほとんどですから、一度、構築された環境を、何年間も使い続けなければいけません。

そういう状況には、UNIX系のOSが向いてると私は思うんですね。ユーザー権限で書き込める場所は限られているから、学生に好きなように触らせても変になりにくいし、小手先の使い方よりも前に、ディレクトリの概念のような基本的なことも教えられるし……。Windowsの使い方はすでに詳しい学生は、年々飛躍的に増え、上手に使えることが万能だと考える学生(や教師)が多いですね。だからこそ、UNIX系のOSを使って、情報検索や情報発信の概念や文書ファイルとしてのプレインテキストの重要性を教えることに意味があると思うんだけどな。

ま、一般的な企業や文系の研究論文の作成には、MS WordやExcelを使わなくてはいけない状況があるというのも否定できませんから、大学の文系学部ではWordやExcelの使い方も教えられる環境は必要でしょう。しかし、そういう必要性がなくて、少ない予算で環境を整えなければいけない小中高校においては“How to use”ではなく“How to think”を学ばせていく方向なら、UNIX系OS、特に、中古パソコンにLinuxなどを入れたクライアント環境などが都合が良いのではないのでしょうか。今すぐは無理だとしても、近い将来の標準環境としては……。

フォーラムの舞台裏で石原先生の発表を聞きながら、Real中継のスタッフや発表を終えられた先生方とそういう話をしていたところ、発表を終えられた石原先生がそこに加わられて、同じことをおっしゃっていました。

また、別のところで(具体的には、HSL本読者の会メーリングリストなんですが)、「お年寄りが初めて使う環境にも、

UNIX系OSが向いているのでは？」という理由が書かれているメールも見かけました。

別にExcelやWordのデータなど扱わないなら、Windowsである必要はない。

GNOMEやKDEのようなグラフィカルなソフトウェアもある。Windowsだってマニュアルを読んでいる限りは難しい。

これを見た私は、「お年寄りが初めて使う環境にUNIX系OSが向いている理由」として、以下のものを書き加えました。

好まないなら、グラフィカルなソフトウェアを使う必要がない。コマンド入力で、最低限のことができる(ダブルクリックが苦手なお年寄りにはメリット)。

最初は、誰かに手取り足取り教えてもらうのなら、ボランティアのお助け隊のメンバーには、UNIX系OSのサポートを好む人が多い(少なくとも私たちのまわりには……)。

例え変な操作をしても、ユーザー権限での変な操作は大して害がない。

メーカーのバージョンアップに振り回されない。

でもね、「その初心者を知り合いたち(特にメール仲間たち)が、Windowsパソコン環境を、標準環境だと思って使っている場合には、上記のメリットは、成立しなくなってしまうんですね。

知り合いたちが、Windows独自フォーマットのファイルをメールに添付してくるため、Linux環境を使っている初心者は、どう対処していいか、チンプンカンプン。相手に聞いても、その状態は理解してもらえない。初心者は、みんなと同じ環境を使っているのが無難……。

こういう状況への対処は、的確な指導ができる先生がいる小中高校では、ちょうどいい教材になるんですけど、そういう先生の絶対数はまだまだ少ないしなあ……。とにかく、初心者がLinuxをクライアント環境として使う時代までには、まだまだクリアしないといけないハードルは多そうですね。私としては、その可能性に期待しているんですけどね。

✂ UNIXコマンド使いへの道 ～ sed 続編 ～

余談でページを使ってしまい、本文を書く場所が少なくなってしまい恐縮ですが、今回の「UNIXコマンド使いへの道」は、“sed”の続編です。

前回の最後に、sedのオプションとして、編集コマンドが記述されたファイルを実行させる場合には、“-f”オプションの後に、ファイル名を記述すると書きました。

そして、次のような実行例を紹介しました。

```
$ cat sedcommand
1i\
この本の2章の登場人物は、以下の方々です。
s/(.*):\(.*\):\(.*\):\(.*\)/<li>4さん(3)。記事本文は<a
href="\1\2.html">\1\2.html</a>/

$ fgrep Chap.2 list | cut -d: -f2,3,4,5 | sed -f sedcommand
この本の2章の登場人物は、以下の方々です。
<li>有賀妙子さん(KAIC)。記事本文は<a href="9605.html">9605.html</a>
<li>大垣斉さん(大阪産大)。記事本文は<a href="9607.html">9607.html</a>
<li>金山典世さん(稚内北星)。記事本文は<a href="9611.html">9611.html</a></li>
<li>戸田洋三さん(千葉大)。記事本文は<a href="9706.html">9706.html</a>
<li>馬場肇さん(京大)。記事本文は<a href="9710.html">9710.html</a>
```

そこで今回は、

```
fgrep Chap.2 list | cut -d: -f2,3,4,5 | sed -f sedcommand
```

というコマンド行について、詳しく説明します。

まず最初の `fgrep` というコマンドは、検索文字列を探し出してくれる `grep` コマンドの親戚で、書かれたままの(正規表現を解釈しない)文字列を探し出すものです。つまり、`"Chap.2"` という文字列を含んでいる行を、探し出してくれるのです。実行してみましょう。

```
$ cat list
:
16:96:05:KAIC:有賀妙子:ありが たえこ:Ariga Taeko:Chap.2
18:96:07:大阪産大:大垣斉:おおがき ひとし:Ohgaki Hitoshi:Chap.2
19:96:08:和歌山工技:井口信和:いぐち のぶかず:Iguchi Nobukazu:Chap.5
21:96:10:OSK:山下康成:やました やすなり:Yamashita Yasunari:Chap.5
22:96:11:稚内北星:金山典世:かなやま のりよ:Kanayama Noriyo:Chap.2
:
$ fgrep Chap.2 list
16:96:05:KAIC:有賀妙子:ありが たえこ:Ariga Taeko:Chap.2
18:96:07:大阪産大:大垣斉:おおがき ひとし:Ohgaki Hitoshi:Chap.2
22:96:11:稚内北星:金山典世:かなやま のりよ:Kanayama Noriyo:Chap.2
29:97:06:千葉大:戸田洋三:とだ ようぞう:Toda Yoza:Chap.2
33:97:10:京大:馬場肇:ばば はじめ:Baba Hajime:Chap.2
```

次に、`cut -d: -f2,3,4,5` というコマンドについて、説明します。まず、`cut` コマンドというのは、標準入力からデータを読み込み、必要なカラム(`-c` オプション)やフィールド(`-f` オプション)を切り出すコマンドです。フィールドを切り出す場合に使う、フィールドとフィールドの区切り記号は、`"-d"` オプションで指定します。

`"-d:"` というのは、区切り記号を `:` (コロン) にするという意味で、`"-f2,3,4,5"` というのは、第2、第3、第4、第5フィールドを切り出すという意味です。では、実行してみます。

```
$ fgrep Chap.2 list | cut -d: -f2,3,4,5
96:05:KAIC:有賀妙子
96:07:大阪産大:大垣斉
96:11:稚内北星:金山典世
97:06:千葉大:戸田洋三
97:10:京大:馬場肇
```

そして、最後の `sed -f sedcommand` というコマンドは、`sedcommand` ファイルに記述された内容を、`"sed"` コマンドに渡すという意味です。

`sedcommand` ファイルの内容のうちの、1行目と2行目の、

```
1i\
この本の2章の登場人物は、以下の方々です。
```

というのは、1行目にこういう行を挿入するという意味です。`"1i"` というのは、1行目の `"1"` と `insert` の `"i"` ですね。そして、次の行は、`sed` でよく使われる、

```
s/検索文字列/置換文字列/
```

という文型の応用です。まず検索文字列の、

```
\(.*\):\(.*\):\(.*\):\(.*\)
```

という部分のうち、`"\"` は、特殊文字の解除のために使われているので、とりあえず取り除くと、以下ようになります。

```
(.):(.):(.):(.)
```

`"*"` はそれぞれ正規表現で、`"."` は任意の1文字に合致し、`"*"` は直前文字の0個以上の繰り返しで、それを括弧でくくることで、それぞれのフィールドに、番号を打つことができます。番号は、前から、`"\1 \2 \3....."` となります。例えば、`"96:05:KAIC:有賀妙子"` を合致させた場合は、`"\1"` は `"96"`、`"\2"` は `"05"`、`"\3"` は `"KAIC"`、`"\4"` は `"有賀妙子"` となります。そのために、置換文字列として、

```
<li>4さん(3)。記事本文は<a href="\1\2.html">\1\2.html</a>
```

と指定すると、

```
<li>有賀妙子さん(KAIC)。記事本文は<a href="9605.html">9605.html</a>
```

と置換されるというわけです。

なんだか、手品のようなですね。これはなかなか便利です。例えば、学校の先生なら、多くの学生のログイン名が含まれる `/etc/passwd` ファイル記述から、各学生の Web ページへのリンク集なんかを、「ひよひよいのひよい!」と作ることができるのですから。

今回は、これでおしまいです。では、また。