Multimedia Electronic Mail with MIME: Overview and Prospects

Nathaniel S. Borenstein
<nsb@bellcore.com>

Bellcore (Bell Communications Research)
Morristown, New Jersey, USA

In November, 1992, I was met at the airport in Sydney by a colleague, Bob Kummerfeld.

Although we had never met in person, we were email friends.

Unlike most email friends, however, we immediately recognized each other by face and voice.

We were two early users of MIME.

[*** Insert picture of Kummerfeld mail here ****]

Why MIME?

Text mail, fax, voice mail very popular.

Text mail deficient for non-English speakers due to ASCII character set.

Modern computer hardware can handle images and audio, too.

Integrated multimedia mail, as in Andrew, Slate, NeXT, etc., has wide appeal, didn’t interoperate.

X.400, which purported to solve the problem, does not.

Why Is Data Interchange Hard?

Hundreds of Incompatible Formats & Standards

ODA -- re-invents world, still not enough!

X.400

- Re-invented the wrong things.
- Complex, Relatively little-used in practice
- Incomplete for Multimedia
- Requires software revolution, not evolution

Heterogeneous formats are here to stay.

Internet Mail Before MIME

Internet Mail: the *de facto* standard

Defined by RFC 821/822

Basis for most workstation & PC mail environments

Even more incomplete than X.400

Non-standard multimedia extensions (Andrew, Next, etc.)

The plan: Let email evolve gracefully

Need standardized Internet multimedia extensions

MIME History and Design Rationale

Designed by Internet Engineering Task Force Working Group on Email Extensions

Politics: remarkable alliance between X.400-lovers and X.400-haters

Working Group formed, Fall 1990.

First MIME draft, Spring 1991.

Proposed Internet Standard, June 1992.

Draft Internet Standard Status: March 1993 (a guess)

Primary design goals:

Text mail in any human language.

Non-text mail in any media type.

Complete compatibility with RFC 821/822.

Robustness over all email transport systems.

Openness to multiple well-known formats.

Easy Extension to new types and formats.

Formal mechanism for type registration.

Easy interoperation with X.400.

MIME Technical Overview

RFC 822 defines a message header (set of structured fields) & plain text body.

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com
Subject: Plain old email

This is a plain old email message.
It contains ASCII text, nothing more.

MIME: Technical Overview, continued

MIME allows extended bodies, with type information in header fields.

Content-type header field provides type/subtype, optional parameters

Seven types, many subtypes expected.

Content-Description gives a textual description of the body data

Content-ID gives unique ID for body parts.

Content-xxx are the only header fields that matter in body parts.

Content-Transfer-Encoding

Binary and long-line data does not survive email transport

Need to specify encoding algorithm for email transport

Two transfer-encoding algorithms defined

Base64

The densest simple (non-compressed) encoding possible for email.

33% data expansion

Uses only A-Z a-z 0-9 + / =

Content-Transfer-Encoding, continued

Quoted-printable

Maximizes readability of included ASCII

All ASCII characters unchanged except =

Other characters represented as =0A, etc.

Up to 200% expansion.

Encoding Design Rationale

uuencode -- insufficiently robust

Base64 based on PEM algorithm

Simple Example #1

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com
Content-type: text/plain
Content-Transfer-Encoding:
quoted-printable

This is text with a single non-ASCII character, =FF.

Simple Example #1

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com
Content-type: image/gif
Content-Transfer-Encoding: base64

R0lGODdhSgGgAfUAAENDQ01NTTw8PEVF...

The MIME Content-types

MIME defines seven content-types. Most extension via subtypes.

1. text.

Subtypes: plain, richtext. (More expected.)

Critical parameter: charset.

"text/plain; charset=iso-8859-1" permits French email.

text/richtext is an extremely simple "common denominator" markup language for enriched text.

The Text Character Sets

Character sets defined:

US-ASCII

ISO-8859-[1-9].

Deferred, but expected (politics...):

Asian languages
(ISO-2022-{JP,KR,etc})

ISO 10646 (UNICODE)

Formal registration of charset names with IANA.

A plain text message in US-ASCII:

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: Plain text mail
Content-type: text/plain; charset=us-ascii

This is plain text mail.

A plain text message in French (using ISO-8859-1, quoted-printable encoding)

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: French mail
Content-type: text/plain; charset=iso-8859-1
Content-transfer-encoding: quoted-printable

Le courrier =E9lectronique =E0 la fran=E7aise n=E9cessite quelques caract=E8res sp=E9ciaux pour faciliter la t=E2che du lecteur et =E9viter les ambigu=EFt=E9s

How the French Should Look

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: French mail

Le courrier électronique à la franc,aise nécessite quelques caractères spéciaux pour faciliter la tâche du lecteur et éviter les ambiguïtés

Why the Encoding?

Isn’t ’à’ better than ’=E0’?

-- Of course, but SMTP is 7-bit

-- Using so-called "8-bit clean" SMTP causes real problems

-- Can have "8-bit clean" networks with gateways to Internet/7-bit-SMTP

-- Mark internal 8-bit mail with:

```
Content-transfer-encoding: 8bit
```

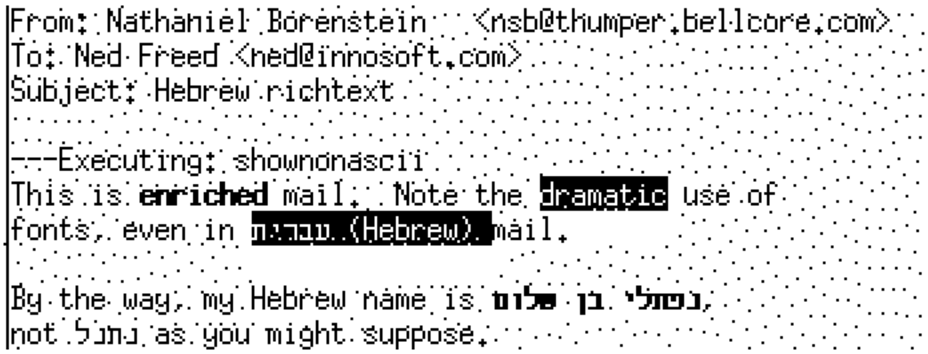
-- New ESMTP Protocol formalizes this

A Hebrew/English richtext message

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: Hebrew richtext
Content-type: text/richtext;
charset=iso-8859-8
Content-transfer-encoding: quoted-printable

This is <bold> enriched </bold> mail. Note the <italic> dramatic</italic> use of<nl>fonts, even in <italic> =FA=E9=F8=E1=F2 (Hebrew) </italic> mail. <nl><nl>By the way, my Hebrew name is <bold>=ED=E5=EC=F9 =EF=E1 =E9=EC=FA=F4=F0</bold>,<nl>not =EC=F0=FA=F0 as you might suppose.

How the Hebrew/English Richtext Looks



From: Nathaniel Borenstein <nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: Hebrew richtext

---Executing: shownonascii
This is **enriched** mail. Note the dramatic use of
fonts, even in עברית (Hebrew) mail.

By the way, my Hebrew name is **נפתלי בן שלום**,
not נתנל as you might suppose.

2. image.

Subtypes GIF, JPEG. Others expected.

3. audio.

Subtype "basic" for single-channel 8Khz u-law. Others expected.

4. video.

Subtype "mpeg". Others plausible.

5. multipart.

Allows multiple body parts of different types, each structured like a mini-message.

All multipart subtypes share a syntax, allowing future experimentation in structured types, etc.

Unique boundary parameter delimits part boundaries & end of message

Multipart Subtypes

Mixed: simple (serial) combinations.

Parallel: for parallel presentation *if possible*.

Alternative: multiple representation of the same data.

Digest: has special defaults for message digests.

```
From: Nathaniel Borenstein
     <nsb@bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: A multipart example
Content-Type: multipart/mixed;
     boundary=CUT_HERE

--CUT_HERE
Content-type: text/plain

Hey, Ned, look at this neat picture:
--CUT_HERE
Content-type: image/gif
Content-Transfer-Encoding: base64

5WVlZ6enqqqr....
--CUT_HERE
Content-type: text/plain

Wasn’t that neat?
--CUT_HERE--
```

```
From: Nathaniel Borenstein <nsb>
To: Ned Freed <ned@innosoft.com>
Subject: An alternative example
Content-Type: multipart/alternative;
     boundary=CUT_HERE

--CUT_HERE
Content-type: text/plain

Hey, Ned, Isn’t MIME great?
--CUT_HERE
Content-type: text/richtext

Hey, <bold> Ned</bold>, isn’t
MIME <italic>great</italic>?
--CUT_HERE
Content-type: application/ODA
Content-Transfer-Encoding: base64

5WVlZ6enqqqr....
--CUT_HERE--
```

6. message.

Subtypes: rfc822, partial, external-body

"message/rfc822" allows encapsulated message.

"message/partial" allows automatic fragmentation and reassembly.

"message/external-body" allows data to be passed by reference (hyperlinks).

Part One of a Two-part Message

From: Nathaniel Borenstein
<nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: A partial example
Content-Type: message/partial;
number=1; total=2;
id="unique-id"

Content-type: image/gif
Content-transfer-encoding: base64

5WVlZ6enqqqr...

Part Two of the Same Message

From: Nathaniel Borenstein
    <nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: A partial example
Content-Type: message/partial;
    number=2; total=2;
    id="unique-id"

Ozs3h4eIKCgo6Ojp...

```
From: Nathaniel Borenstein <nsb>
To: Ned Freed <ned@innosoft.com>
Subject: Some external references
Content-Type: multipart/alternative;
     boundary=42

--42
Content-Type: message/external-body;
     access-type=mail-server
     server="listserv@bogus.bitnet"

Content-type: application/postscript

get rfc-xxxx doc
--42
```

```
Content-Type: message/external-body;
     name="BodyFormats.ps";
     site="thumper.bellcore.com";
     access-type=ANON-FTP;
     directory="pub/nsb"

Content-type: application/postscript
--42
Content-Type: message/external-body;
     name="/u/nsb/BodyFormats.ps";
     site="thumper.bellcore.com";
     access-type=local-file

Content-type: application/postscript
--42--
```

7. application.

Subtypes: PostScript, ODA.

Catch-all. Most creative extensions expected here.

Possible examples: Acknowledgement-request, EDI, Interactive-survey

Already registered: Andrew-inset, ATOMICMAIL

A Complex MIME Message

```
From: Nathaniel Borenstein
     <nsb@thumper.bellcore.com>
To: Ned Freed <ned@innosoft.com>
Subject: A complex example
Content-Type: multipart/mixed;
     boundary=FOOBAR

--FOOBAR

PLAIN TEXT GOES HERE.
--FOOBAR
Content-Type: audio/basic
Content-Transfer-Encoding: base64

BASE64-ENCODED AUDIO DATA
```

--FOOBAR
Content-Type: image/gif
Content-Transfer-Encoding: Base64

BASE64-ENCODED IMAGE DATA GOES HERE
--FOOBAR
Content-type: text/richtext

This is <italic>richtext</italic>.
<nl><nl> Isn’t it <bigger>cool?</bigger>
--FOOBAR--

What the Complex Example Looks like with a standard mail reader

What the Complex Example Looks like with an integrated multimedia mail reader

Minimal MIME-Conformance

1. "MIME-Version: 1.0"

2. Content-Transfer-Encoding

3. text/plain in US-ASCII

4. ASCII portions of ISO-8859-*

5. Intelligent treatment of unrecognized types and character sets. (E.g. decode & write to file)

6. message/rfc822

7. multipart (mixed, alternative, *)

Non-ASCII Characters In Message Headers

MIME, as defined by RFC 1341, only addresses message bodies.

The need for non-ASCII header text is addressed by RFC 1342, a companion document to MIME.

Rationale: More complex, more ugly, less consensus

Non-ASCII headers are a special case.

- Only text (no images, etc.) is permitted

- Non-ASCII text is permitted only in certain very special locations

- All of this is complicated and made uglier by RFC 822 header syntax.

- Old parsers of RFC 822 headers should not fail with these extensions.

Non-ASCII data represented as "encoded words", details not given here.

Non-ASCII Data Are Permitted...

Only in fields intended for human reading. (Not automatic processing):

Subject: **NON-ASCII**
Comments: **NON-ASCII**
Content-Description: **ETC**

From: nsb (**NON-ASCII**)

From: **NON-ASCII** <x@y>

Examples of Non-ASCII Header Data

NOTE: Large presentaton fonts create erroneous line breaks in these examples.

From: nsb@bellcore.com (=?ISO-8859-8?Q?=ED=E5=EC=F9 =EF=E1 =E9=EC=E8=F4=F0?=)

From: nsb@bellcore.com (נפטלי בן שלום)

To: ysato@etl.go.jp (=?ISO-2022-JP?B?GyRAOjRGI0stGyhK?=)

To: ysato@etl.go.jp (佐藤豊)

Controversies and Problems

Banned/punted:

- Nested encodings
- compression
- uuencode
- Lots of top-level types

X.400 interoperation required:

- No use of preamble/epilogue
- Parts are NOT encapsulated messages.

Transport Issues:

- No dependence on line or byte counts.
- 8-bit transport independent
- Initial character sets are limited

Status Report

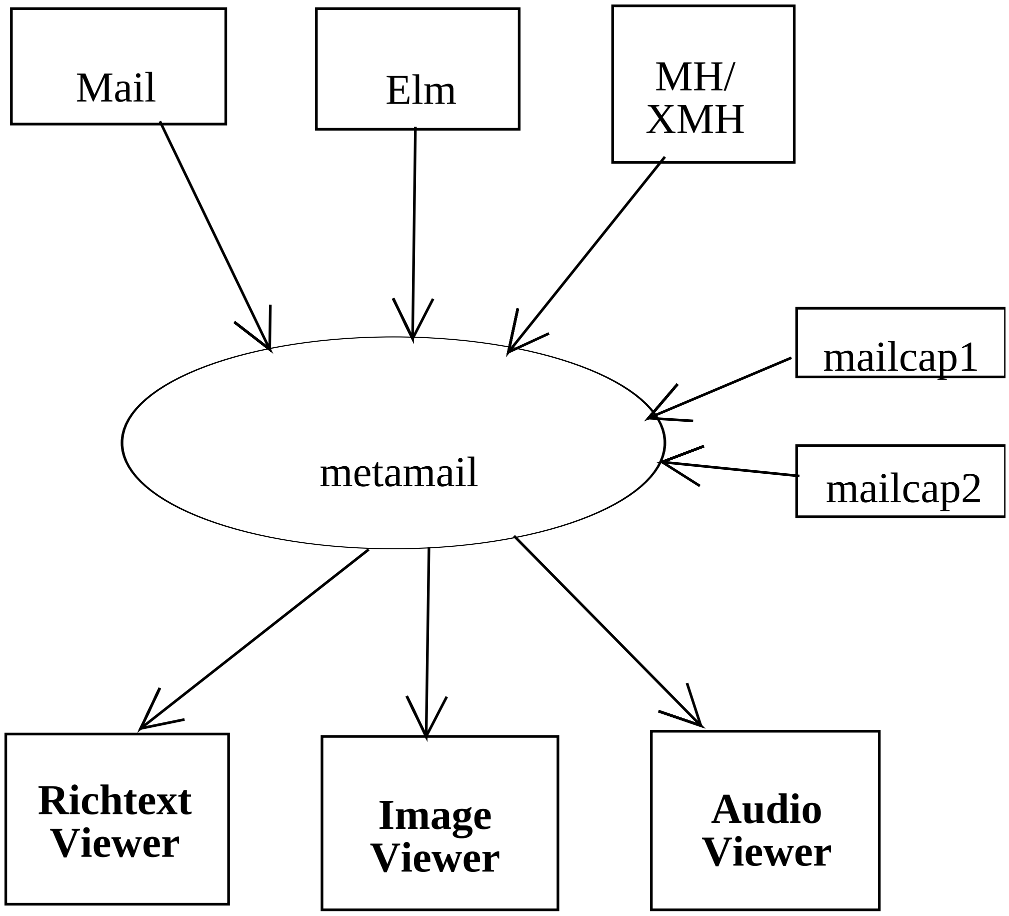
Dozens of implementations under way.

At least 3 public domain implementations available now.

One of these, metamail, is a MIME transition tool.

Metamail comes with patches for over a dozen mail readers on UNIX, DOS, and Amiga, and is already used at thousands of sites and is being incorporated into various products.

The Metamail Architecture



Other MIME Implementations

Currently available implementations include:

> PMDF, IMAP2, C-Client, MailManager, MH-MIME, Z-Mail, Andrew, Pine, Elm, Unix System 5 4.3, STI Document Browser, Servicemail, MIXMH

Many unannounced commercial implementations under way.

Private information suggests MIME support from key vendors on most platforms.

Glimpses of The Multimedia Future

Multimedia communication (e.g., MIME) opens up more opportunities than telephones and ASCII combined.

Two illustrations:

> ATOMICMAIL, a language for interactive (computational) email
>
> *The Electric Eclectic*, a magazine for the Internet era.

Computational Email

Computational ("active") email has a long history, at least back to 1975 (Rand).

Critical problems:

- Non-distributed paradigm
- Heterogeneous environments
- Lack of security
- Lack of interface portability

MIME solves the first two problems!

ATOMICMAIL tackles the other two.

Portability

Portability is achieved through "lowest common denominator" user interface.

Primitives: getstring, getmultiplechoice, 3D multiple choice with branching, etc.

ATOMICMAIL interpreters depend on environment, but not programs.

User interfaces:

Curses implementations for terminals

TK implementation for X11

Unsupported: Mac implementation

Security

Language must be constrained to do no harm.

No general access to operating system or CPU.

Constrained access to system resources.

Severe constraints on file access, sending messages, printing text, etc.

Applications of ATOMICMAIL

User surveys

Meeting scheduling

Document Distribution

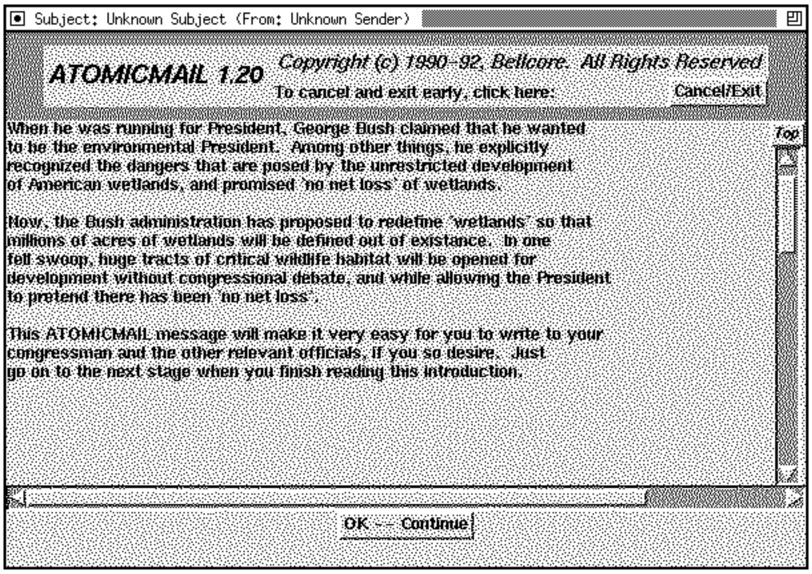
Information finder (organizational memory)

Mail-based game

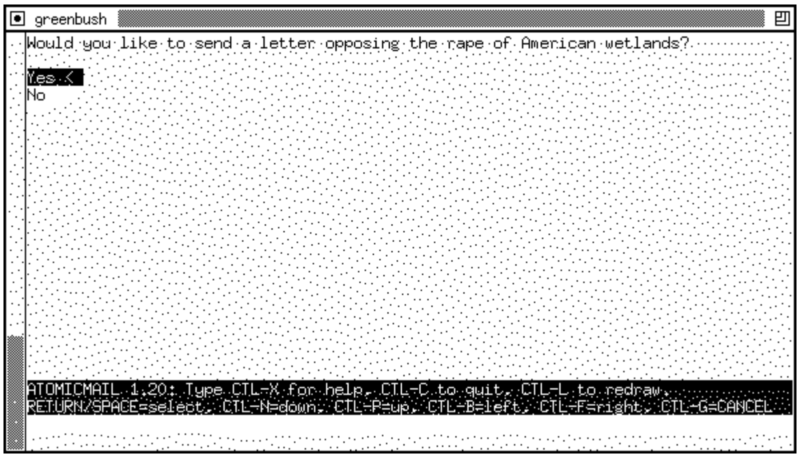
Employee time reporting

Activist alert

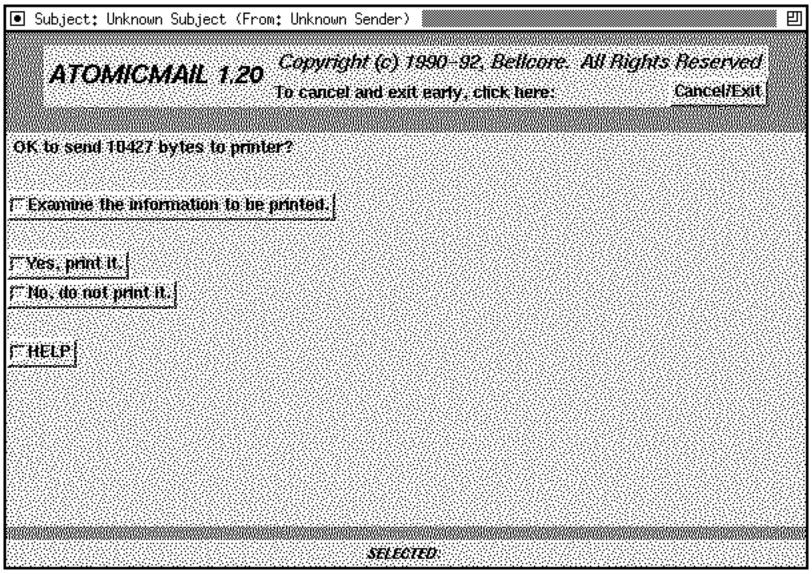
An Activist Alert, first screen, X11 Interface



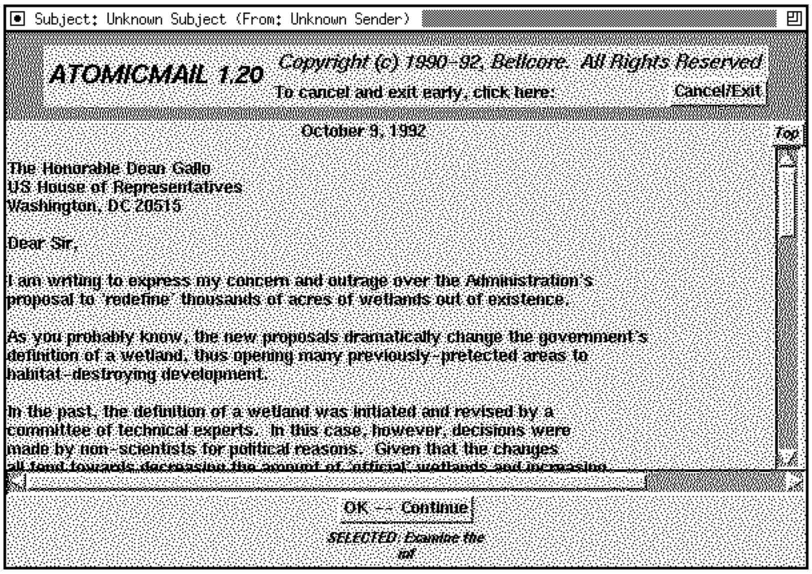
Activist Alert, second screen, Curses Interface



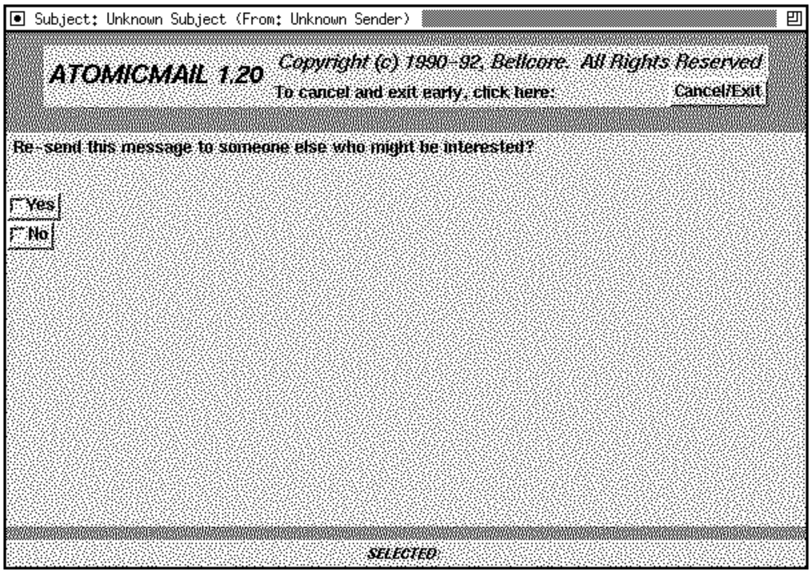
Activist alert, third screen, X11 Interface



Activist alert, fourth screen, X11 Interface



Activist alert, finall screen, X11 Interface



Computational Email As Network Infrastructure

Computational email could be key infrastructure for distributed applications

Example: Organizational Memory. Someone knows X, how do I find out?

Traditional approach: Ask a neighbor.

Netnews: Broadcast query to world

Innovation: Malone & Ackerman’s "Answer Garden" inhibited by software distribution & user buy-in.

Organizational Memory: An ATOMICMAIL approach

User sends query to "Mr. Wizard" server

Answers sent back to user with ATOMICMAIL for feedback.

If no good answers, ATOMICMAIL queries go to experts for answers or names of further experts

Simple queries answered from database

Moderate queries answered by local experts

Hard queries cross oceans to find experts

ATOMICMAIL Project Status

ATOMICMAIL is a research prototype, slightly flakey.

Free research licenses to universities, otherwise unavailable.

The Mr. Wizard project is on hold.

Forget the software, focus on the vision.

Gearing up: Efforts to define MIME standard type(s) for interactive email

One proposal: A "safe" Tcl/Tk.

The Electric Eclectic: A Multimedia Magazine

Researchers have prototyped multimedia magazine technology. Two missing pieces:

Standards for interoperation (MIME)

A reason for using it

Is multimedia useful? What for? Will people ever prefer it to paper?

Maybe best answered in practice...

...a multimedia Internet magazine, tentatively called *The Electric Eclectic*.

What is *the Electric Eclectic*?

Nothing yet. I need your help to refine its definition and make it real.

Current vision:

> A flexible magazine
>
> A multimedia magazine
>
> A magazine with high standards
>
> An evolving prototype for the future of publishing

The Electric Eclectic: A flexible magazine

Sub-magazines on topics of interest permit customized individual subscriptions.

Individual articles selectable using MIME multipart/digest mechanism.

Other customizations: variant formats, data inclusion versus external reference, etc.

If commercialized, custom advertising.

Consider electronic reader response cards.

The Electric Eclectic: A Multimedia Magazine

Multimedia will be used where appropriate. Where is that?

> Telecommunications sub-magazine: the audio of a Gore speech on NREN.
>
> Art history sub-magazine: pictures from the Louvre
>
> Medical technology sub-magazine: video of new MRI techniques

MIME external-reference makes retrieval of large objects optional

The Electric Eclectic: A Magazine with High standards

NOT netnews or mailing list. Submissions reviewed, revised, edited.

My goal: the worst article is better than 999 out of 1000 netnews messages.

Articles must be submitted to and approved by an editorial board.

Flexible publication schedule (no printer!) helps preserve quality over quantity

The Electric Eclectic: An Evolving Prototype for the Future

Editorial & organizational policies, tools, & MIME types will need to evolve.

Initially, a free volunteer-based Internet service, but the future may be commercial.

Subscriber-paid magazines could use customization and filtering for added value.

Magazines with advertising could remain free, permit copying if ads are retained.

Volunteers are wanted: send me email!

nsb@bellcore.com

Remaining Technical Hurdles (Opportunities for Developers)

More MIME implementations

Wider Deployment

Better MIME-Generating Tools

More MIME-Based Applications

X.400 Gateways

Remaining Standardization Hurdles (Opportunities for MIME Extensions)

Constraining & defining the set of subtypes

Standard formats for tight coupling of separate objects.

Standardized control structures for interaction (e.g. surveys, forms, EDI)

Richer support for audio data

Character Set Chaos (ISO 10646?)

New media types (e.g. smell, virtual worlds)

All of this seems to be MIME-able.

Implications of MIME for Email and Other Multimedia Applications

Whither X.400? Choose one:

MIME is the death of X.400

MIME will help X.400 a lot.

I don’t believe either, actually.

Nothing ever dies. FORTRAN, COBOL...

It doesn’t matter. Gateway document will define mutual encapsulation, users will never again need to care.

Whither Fax?

Probably no effect.

Might facilitate future merging with email

Whither voice mail?

May help with interoperation.

MIME is being considered for digital AMIS standard.

MIME in non-mail applications

Paranoid robustness can’t hurt elsewhere.

It’s nice to know your data format is mind-bogglingly robust.

Open architecture invites re-use.

Metamail software already useful in non-email applications. (Superbook, WWW, Gopher, WAIS, multimedia databases)

De Facto standards happen bottom-up.

Access to MIME

Anonymous ftp (thumper.bellcore.com, "cd pub/nsb")

Mail to mailserver@thumper.bellcore.com, subject "Help".

Nathaniel S. Borenstein
<nsb@bellcore.com>