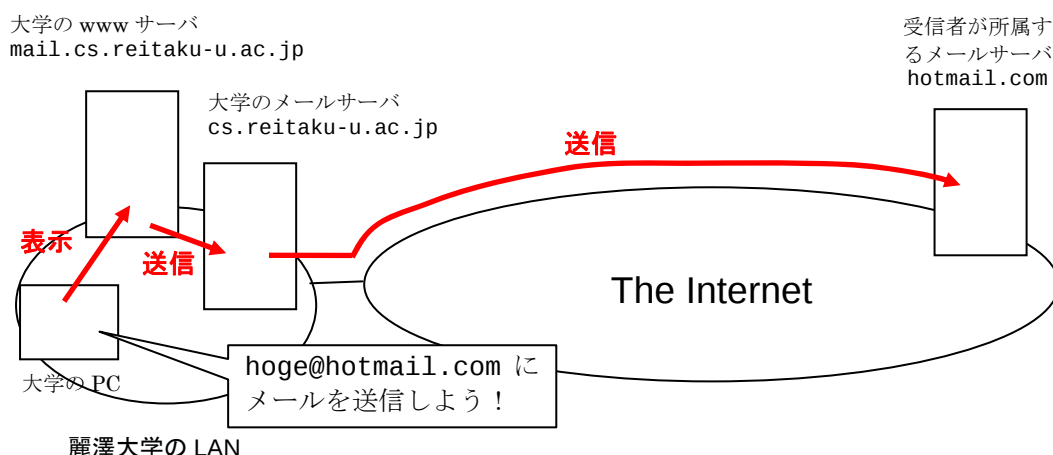


第 8 回：外国語で電子メール

1. 電子メールの仕組み

1.1. メールの送信とエンコード方式

インターネット上でやりとりされるメールの送信には、SMTP (simple mail transfer protocol) という種類のネットワークサービス (通信プロトコル protocol) を使う。送信されたメールは、SMTP を使い、インターネットを経由してあて先のメールサーバに届けられる¹。



もともと、SMTP は 7bit でしか文字を扱うことができない。7bit, つまり 7 桁の 0 と 1 の組み合わせで表現できる文字の種類は、最大 $2^7 = 128$ 文字である (第 7 回資料参照)。アメリカ国内で使われている ASCII は 7bit でテキストを送信可能であるが、ヨーロッパでは 8bit, 日本を含む漢字圏は漢字などを 2 つの 8bit (=2 バイト byte) で表現, というふうに、エンコード方式によっては SMTP でメールをやりとりすることができない。

そこで、メール送信用に特別な 7bit の文字エンコード方式を開発し、SMTP でテキストを送信可能にしたのが日本語である。日本語の場合、電子メールは通常の 8bit の文字エンコード方式 (Windows, Macintosh では Shift JIS) から 7bit の ISO-2022-JP というエンコード方式に自動的に変換されて送られる。この方法であれば、SMTP を使ったメールサーバでも、日本語のメールを安心してやりとりできる。

その後、8bit の文字コードをサポートする ESMTP (SMTP Service Extensions) が開発され、7bit でないその他の文字エンコード方式のテキストでも送受信が可能になった。しかし、日本語のメールは現在でも 7bit に変換して処理されるのが普通である。

メールサーバの多くは、現在 ESMTP を使っているが、SMTP を使うメールサーバも依然

¹ サーバから PC にメールを取り込む際には、POP3 (「ぽっぷ」 post office protocol version3), IMAP4 (「あいまっぷ」 internet message access protocol 4) などのメール受信用プロトコルを使う。送信 (SMTP) と受信 (POP, IMAP) は、システム的设计によって異なるコンピュータで行われる場合もある。

としてインターネット上に存在する。ESMTP を使って送信されたメールは、SMTP サーバでも正しく処理できるように工夫されており、ESMTP サーバから送信すれば、さまざまな言語のメールをインターネット上で無理なくやり取りできるようになっている。

1.2. 電子メールのヘッダ情報

電子メールのテキストに使われるエンコード情報は、電子メールの先頭にある **ヘッダ header** という情報の中に記述されている (Active!Mail では、メール閲覧画面では隠されているが、「sources」ボタン (右図) で表示できる。他の多くのアプリケーションでもヘッダの内容を表示する機能がある)。以下は、大学のメールサーバから送った日本語のメールのヘッダの例である：



```

Return-Path: schiba@reitaku-u.ac.jp
Date: Thu, 23 Aug 2001 17:41:49 +0900
From: schiba@reitaku-u.ac.jp
To: [REDACTED]@nifty.com
Subject: test
Received: from mail158.nifty.com (mail158.nifty.com [202.246.37.130])
    by ums511.nifty.ne.jp (8.9.3+3.2W/3.7W000628) with ESMTP id RAA13886
    for <[REDACTED]@nifty.ne.jp>; Thu, 23 Aug 2001 17:42:04 +0900 (JST)
Received: from poplar.cs.reitaku-u.ac.jp
    by mail158.nifty.com (8.11.4+3.4W/3.7W-10/13/99) with ESMTP id f7N8fpe18025
    for <[REDACTED]@nifty.com>; Thu, 23 Aug 2001 17:41:51 +0900
Received: from reitaku-u.ac.jp (localhost [127.0.0.1])
    by poplar.cs.reitaku-u.ac.jp (8.9.3/3.7W-POPLAR) with ESMTP id RAA16068
    for <[REDACTED]@nifty.com>; Thu, 23 Aug 2001 17:41:50 +0900 (JST)
Received: from [REDACTED] (keyaki [157.17.10.20])
    by reitaku-u.ac.jp (8.9.1/3.7Wpl2-KEYAKI) with SMTP id RAA01591
    for <[REDACTED]@nifty.com>; Thu, 23 Aug 2001 17:41:49 +0900 (JST)
Message-Id: <[REDACTED].RAA01591@reitaku-u.ac.jp>
MIME-Version: 1.0
Content-Type: text/plain; charset=iso-2022-jp
Content-Transfer-Encoding: 7bit
X-Mailer: impression office
X-UIDL: 3b84c1f0044TIACW

テストメールです。

千葉 庄寿
=====
CHIBA Shoju
e-mail: schiba@reitaku-u.ac.jp
phone: +81-(0)471-73-3019 (office)

```

送信日時 (Date), 送信元アドレス (From), 宛先アドレス (To), 表題 (Subject) などの情報

文字コードなど、データの内容に関する情報 (Content-Type と Content-Transfer-Encoding): 内容はテキストで、文字コード体系は iso-2022-jp、データのビット長は 7bit

メール本文は、ヘッダの後

1.3. 外国語のメール送受信

外国語でメールを送受信する場合には、使用するアプリケーション (メールクライアントソフト) が外国語でのメール送受信に対応していることが必要である。つまり：

- (1) 適切なフォントでテキストを表示し、
- (2) 適切なエンコードでテキストを編集し、
- (3) 適切なエンコードのヘッダ情報を加え、
- (4) 必要ならエンコードの変換を行う

ことが必要である。このうち (3) と (4) についてはメールを送信するアプリケーションが自動的に処理をおこなうので、ユーザは (1) と (2) に気をつけることになる。なお、大学の電子メールシステム Active!Mail (現在のバージョン) では、送信の際 (2) に気をつけることを除けば、ほぼ自動で処理が行われる。

2. Active!Mail を使ったメールの送受信

一昨年 11 月から学部生全員が利用している電子メールソフト「Active!Mail 2003」は WWW ブラウザ上で動作するタイプのメールソフト (「Web メール」と呼ばれる) であり、ブラウザを使い、大学のメールサーバとオンラインで接続してメールを送受信する²。

Active!Mail URL: <https://mail.cs.reitaku-u.ac.jp/active-mail/>

※ https は「セキュア secure http」と呼ばれ、PC と WWW サーバの間でやりとりされる送信データが暗号化される。

これに対し、プロバイダなどと契約してメールを送信する場合、多くはメール送受信のための独立したソフトウェア (「メーラー」とも言う。例えば Outlook Express) を使う。この場合、PC からサーバへのメールの送信自体に SMTP が使われる。これに対し大学の Active!Mail では、メールはブラウザから WWW サーバを介し、メールサーバへ渡される。

Active!Mail は、日本語の他に英語、中国語 (簡体字)、韓国語のインターフェースを備えた手軽な多言語メール環境であり、日本語など数言語のインターフェースで多言語の送受信ができる。

実習の準備: Active!Mail にログオンし、日本語メニューで自分あてにテストメールを送りなさい。受け取ったら、さらに返信を送ってみなさい。Active!Mail に慣れていない人はこの機会に情報システムセンターのマニュアル³を参考に操作に慣れること。

なお、Internet Explorer など、Active!Mail が利用する WWW ブラウザは、言語ごとにフォントの設定をすることができ、§ 1.3. で述べた「(1) 適切なフォントでテキストを表示」という条件がクリアできる。一方、§ 1.3. で述べた「(2) 適切なエンコードでテキストを編集」については、受信メールの表示の際は Active!Mail 側で自動的に行われるが、送信の際にはユーザが正しく設定をおこなう必要がある。

2.1. インターフェースの変更

メニューの言語は、英語、日本語、韓国語、中国語 (簡体字) から選ぶことができる。ログオン時のほか、ログオン後でも、「オプション option」から「表示と編集」を選び、言語を変更して「OK」ボタンを押すとメニューの言語を変更できる。



² かつて大学のメールソフトとして使われていた Impression Office は「グループウェア」と呼ばれるアプリケーションで、PC からサーバにオンラインで接続した状態でメールの送受信を含むさまざまな作業をおこなう。Impression Office は古い SMTP でメールを送信する日本語のメール専用ソフトであったため、ASCII の基本アルファベット以外の文字を用いて外国語のメールをやりとりすることはできなかった。

³ 電子メールの利用方法 URL: <http://www.reitaku-u.ac.jp/risc/pdf/activemail-2006-2.pdf>

各言語のインターフェースのメニュー構成はほとんど同じである。中国語、韓国語を学習している人は、この機会にその言語で使うメールの専門用語を覚えてはどうだろうか。

2.2. 各言語でのメール送受信

昨年度までは、日本語、韓国語、中国語(簡体字) のインターフェースでは、原則として ASCII とそれぞれの言語で書かれたメールしか送受信できなかった。4 月にバージョンアップがおこなわれた結果、「多言語を同時に表示」オプションをチェックすることで、全ての言語のインターフェースで多言語電子メールの送受信ができるようになった。

実習：中国語、または韓国語のメニューでログオンし、日本語のメールが正しく表示されていることを確認しなさい。

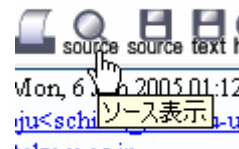
2.2.1 メール受信画面のエンコード設定

新しい Active!Mail⁴ は、メールのヘッダで指定されている文字エンコード方式に従い、メールをユニコード (UTF-8) に変換して表示させている。そのため、さまざまな外国語のテキストを同時に表示することができる。



※ メールによっては、ヘッダに文字エンコードの指定がない場合がある。現在の Active!Mail のシステムでは、このようなメールのエンコードの処理がうまく行かず、文字化けがおこってしまうことがあるようだ。

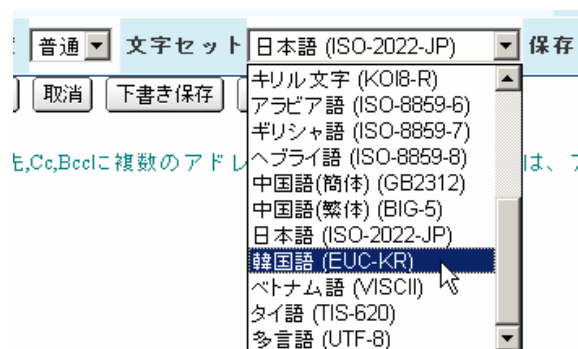
実習：Active!Mail の「メール受信」を開き、任意のメールを選択して表示し、メニューの「ソース表示」(右図) を使ってメールのオリジナルのヘッダ情報を調べ、§ 1.2. で確認した、Content-Type ヘッダの charset の指定を確認しなさい。



2.2.2 メール作成画面のエンコード設定

「メール作成」画面にある「文字セット」を言語に合わせて変更する(下図)。

⁴ なお、大学の Active!Mail は来年 4 月をめどに Active!Mail 2005 にバージョンアップする予定である。



「文字セット」では、一部、前回紹介した「標準的な文字エンコード方式」と異なるものがサポートされている（西ヨーロッパ言語の文字エンコード方式 ISO-8859-15 の特徴については補足資料を参照）。

言語	Active!Mail の「文字セット」の名称	前回紹介した「標準的な」文字エンコード方式の名称
韓国語	韓国語 (EUC-KR)	EUC-KR
西ヨーロッパ言語	欧米 (ISO-8859-15)	ISO-8859-1
中国語 (台湾)	中国語(繁体) (BIG-5)	big5
中国語 (大陸)	中国語(簡体) (GB2312)	GB2312
日本語	日本語 (ISO-2022-JP)	Shift JIS
タイ語	タイ語 (TIS-620)	TIS-620
多言語	多言語 (UTF-8)	UTF-8

実習：日本語メニューから外国語で書いたメールを自分あてに出し、受信して表示結果を調べなさい。送信の際、エンコードを正しく指定していることを確認すること。また、「件名 (Subject)」は全て ASCII の半角英数字を使うこと。

次に、英語、中国語、または韓国語のメニューから、自分あてに日本語でメールを出し、送信されたメールが正しく表示されることを調べなさい。エンコードは ISO-2022-JP を正しく指定すること。

さらに、日本語メニューから、多言語で書いたメールを自分あてに送信し、受信して表示結果を調べなさい。エンコードは UTF-8 を指定すること。

電子メールで使用できる文字は、その言語の文字エンコード方式で表現できる文字に制限される。文字エンコード方式を正しく設定せずに送信すると、指定したエンコードに含まれない文字は「文字化け」を起こしてしまう。

例えば、日本語で表せない文字を含むテキストをうっかり日本語として送信すると、文字の脱落が起こってしまう。⁵

日本語メニュー上で打ったテキスト：

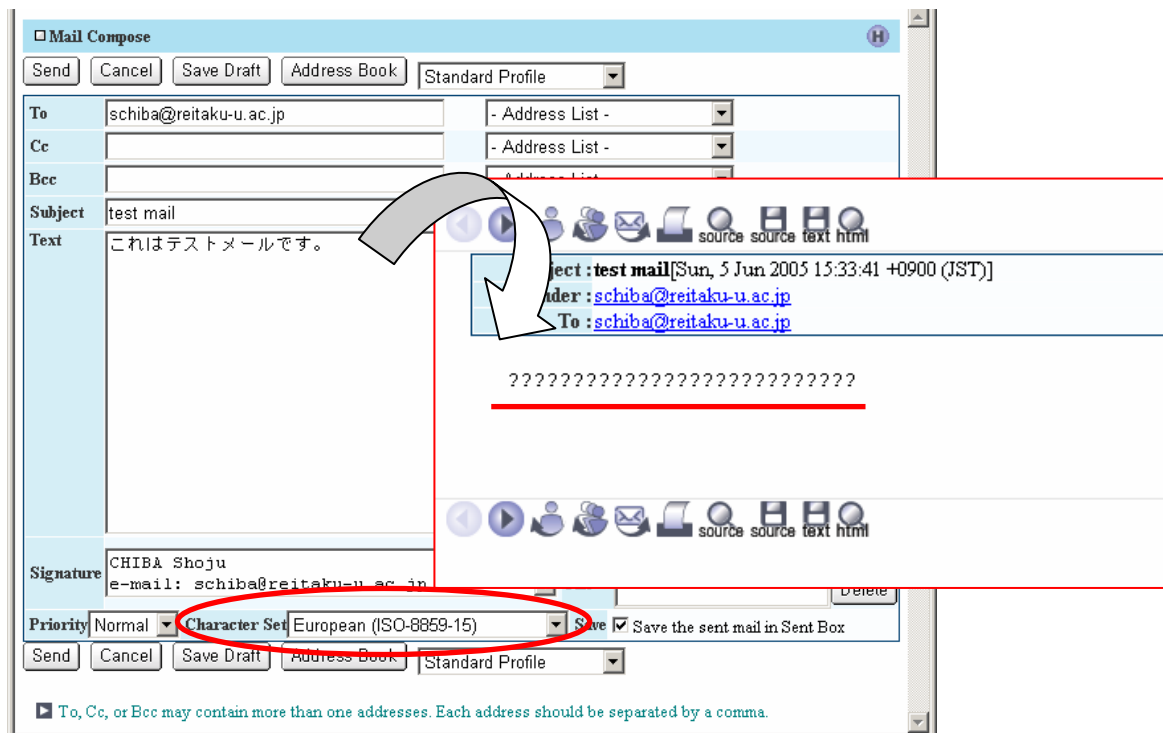
⁵ Active!Mail 以外のメールソフトには、指定したエンコード方式に含まれない文字がある場合に、その文字を HTML の「文字参照」形式（後日授業で説明する）で表現するものもある。この場合でも、Active!Mail を含め、メールソフトの多くはこのような参照文字をそのまま表示してしまうため、その文字をすぐに読むことはできない。

Hei, tässä on pieni viesti sinulle.
 Uusi sähköpostiosoite on nyt käytössä.
 Terv. Tsibale

実際に受け取ったテキスト :

Hei, tss on pieni viesti sinulle.
 Uusi shkpostiosoite on nyt kytss.
 Terv. Tsibale

また、日本語のテキストを西ヨーロッパ言語のエンコード方式 (ISO-8859-15) で送信すると、以下のように日本語が「?」に置き換えられてしまう。



送信時の設定ミスによるこのような文字化け(文字欠け)メールは修復が難しく、相手に迷惑をかけてしまうことになるので注意しよう。

実習：以下のように、わざと間違った文字エンコード方式を指定したメールを自分あてに送り、受信して表示結果を調べなさい。第 5 回授業で作成した自己紹介のテキスト (日本語と、日本語以外の言語) や挨拶表現とその日本語訳 (多言語) など、ある程度の長さのあるテキストを送ると実際におこる文字化けを発見しやすいだろう。

1. 外国語で書いたメールを日本語 (ISO-2022-JP) の文字セットで送信した場合。
2. 日本語で書いたメールを韓国語 (EUC-KR) の文字セットで送信した場合。
3. 多言語で書いたメールを、入力言語の 1 つだけを文字セットで指定し送信した場合。

なお、「件名 (Subject)」は全て ASCII の半角英数字を使うこと。

2.3. 多言語でメールを送受信する際のヒント

その 1：メールの件名 (subject) は ASCII で表記する

Active!mail では、送信時、件名 (subject, 表題) を指定された文字エンコード方式に変換して送信する。従って、日本語以外の言語で書かれたメールに日本語で件名をつけることや、その逆はできないので注意しよう。

そもそも、相手のメールソフトが多言語にどの程度対応しているかによって、外国語のテキストで書かれた件名を正しく表示できるかどうかが変わってくる。受け取り手がどのようなメール環境にいる人なのか (どの言語のメールを頻繁に受け取る人か etc.) を考え、メールの件名をどのような文字で書くかを決めるとよい。確実に件名を読んでもらうためには、外国語で書くメールについては件名を ASCII の基本英数字・記号のみで書くよう、癖をつけるとよい。

その 2 : UTF-8 による多言語混在メールはなるべく送らない

Active!Mail は UTF-8 を使い、Unicode テキストを含む多言語メールを送受信できる。ただし、一般的なメールソフトの Unicode 対応はさまざまであり、また、Unicode に対応したメールソフトでも、届いたメールを Unicode で表示しなおす方法を知らなければ正しくメールを表示できない場合がある。従って、現状では、多言語混在メールはなるべく送らず、単一の言語とその文字エンコード方式を使って送るよう、心がけたほうがよい。

その 3 : 日本語で書かれた名前や署名は使わない



Active!Mail では、「オプション」メニューの「プロフィール」欄で自分の名前や署名をあらかじめ指定しておくことができる。便利な機能であるが、日本語で作成された名前や署名は日本語が表示できないメールソフトでは正しく読めないばかりか、文字エンコードの変換の際に文字化けが発生してしまう原因となる。日本語が入った署名が登録されている場合は、外国語メールの送信前に自分で削除すること。(Active!Mail では、プロフィールを言語ごとに個別に指定できるので、外国語用のプロフィールを個別に用意するとよい。)

3. Active!Mail 利用上のヒント : 外国語のメールの保存と活用

メールの文面を、Word などに貼り付けて活用したいことがある。Active!Mail では、ブラウザ上にメールを表示するので、メールの必要箇所をコピーし、他のソフトに貼り付けることができる。

そのほかに、Active!Mail の「Save Text」機能 (右図) を使うと、以下の文字エンコード方式で、メールをファイルに保存することができる。



英語	UTF-8
その他の言語	その言語の標準的なエンコード方式 (日本語 : Shift JIS; 中国語 : GB2312; 韓国語 : EUC-KR)

課題 (来週授業開始時まで印刷のうえ提出のこと) :

Active!Mail で日本語以外のメールを送受信する手順を確認し、自分の選択する言語ごとに 2 人 1 組になって何度かメールを交換してみなさい。外国語メールの送信方法を確認したら、相手に自己紹介のメール (日本語以外の言語で作成したもの) を作成し、送信しなさい。送信済みのテキストを表示して印刷しなさい (メールは「送信箱」に保存されている。印刷の際は第 3 回授業 § 4 で説明した「印刷に関する注意」を確認すること)。印刷した結果に学籍番号、名前を書いて提出しなさい。

次回は外国語テキスト作成の応用例として、多言語 HTML 文書を作成してみる。

補足：西ヨーロッパ言語の文字エンコード方式 ISO-8859-15

Active!Mail では、授業で西ヨーロッパ言語の「標準的な」文字エンコード方式として紹介した ISO-8859-1 の代わりに、ISO-8859-15 が西ヨーロッパ言語の送信時の文字エンコード方式として登録されている。ISO-8859-1 と 15 で多くの文字は同一のコードポイントをもつが、ISO-8859-15 は比較的新しい文字エンコード方式で、ISO-8859-1 とは異なり、ISO-8859-1 のあまり使われない記号類をユーロ記号など最近の記号に置き換えてある。

ISO-8859 が規定するさまざまな文字エンコード方式については、清水哲郎(2001)『図解で分かる文字コードのすべて』(日本実業出版社, pp.23—26)を参照。以下は Roman Czyborra 氏の「ISO 8859 Alphabet Soup」(URL: <http://czyborra.com/charsets/iso8859.html>) から引用した表である。文字の対応を確認してみよう。

ISO-8859-1

A0	A1	A2	A3	A4	A5	A6	A7	A8	A9	AA	AB	AC	AD	AE	AF
	í	í	í	í	í	í	í	í	í	í	í	í	í	í	í
B0	±	±	±	±	±	±	±	±	±	±	±	±	±	±	±
C0	À	À	À	À	À	À	À	À	À	À	À	À	À	À	À
D0	Đ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ
E0	à	á	á	á	á	á	á	á	á	á	á	á	á	á	á
F0	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä

ISO-8859-15

A0	A1	A2	A3	A4	A5	A6	A7	A8	A9	AA	AB	AC	AD	AE	AF
	í	í	í	í	í	í	í	í	í	í	í	í	í	í	í
B0	±	±	±	±	±	±	±	±	±	±	±	±	±	±	±
C0	À	À	À	À	À	À	À	À	À	À	À	À	À	À	À
D0	Đ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ	Ñ
E0	à	á	á	á	á	á	á	á	á	á	á	á	á	á	á
F0	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä	ä

なお、ISO-8859-15 の文字の配置は Windows の西ヨーロッパ言語のコードページ (1252) とは異なるので注意。後者は Microsoft 社が ISO-8859-1 をベースにユーロなど一部の文字を独自の方法で追加したもので、Windows でのみ利用可能である (第 7 回資料 § 1.5.を参照)。

【重要】Active!Mail が採用している ISO-8859-15 は、比較的新しい規格であり、多くのメールソフトが「標準」としている西ヨーロッパ言語のテキストの文字エンコード方式は、現在でも ISO-8859-1 であると考えられる。従って、ISO-8859-15 で導入されたユーロ記号などの文字を使ったメールが、相手のメールソフトによっては正しく処理できないことがあると思われる。第 7 回授業資料 § 1.5.で言及したように、ユーロ記号など、文字化けすると困る重要な記号は、「EUR」などの代替表記を使い確実に相手に伝わるよう配慮したほうがよい。

なお、ISO-8859-15 は、ISO-8859-1 の後継として今後普及する可能性がもっとも高い西ヨーロッパ言語の文字エンコード方式と考えられている (Active!Mail が ISO-8859-1 の代わりに採用したことも、一つの証左だろう)。