



Standards Board

ELECTRONIC EXCHANGE GROUP

Lisbon, 11 September 2003

Remarks and proposals for new EDI messages for mail (PTC/EMS Cooperative)

(Agenda item 5)

Document by the International Bureau

I. Introduction

1 The UPU PTC and EMS Unit conducted a joint review of the new EDI messages. The outcome of this review is included in a joint working document provided in annex 1. This document raises a number of issues and offers a number of proposals for discussion by the EXG.

II. Scope of the review

The main area covered by the document include:

- EVTRPT: transition period, volume of messages and potential usage,
- ITMATT: Scope, design and usage.

III. Outcome

5 The EXG is invited to review the 13 proposals included in the review and comment on the approach submitted by the EMS Unit and the PTC. Finally, the EXG is asked to decide on a way to progress with the proposals.

**Remarks and proposals for
new EDI messages for mail**

PTC: Jean-Marc Coeffic

EMS cooperative: Marilia Carvalho
David Chamberlin

Table of contents

INTRODUCTION	3
1 GENERAL REMARKS: SCOPE OF THE NEW MESSAGES	4
1.1 PARTIES INVOLVED	4
1.2 AN OPEN STANDARD	4
2 EVTRPT: REMARKS AND PROPOSALS	6
2.1 INTRODUCTION	6
2.2 TRANSITION PERIOD.....	6
2.2.1 Message conversions	6
2.2.2 Distributed conversions or centralized conversions?.....	6
2.3 SIZE OF EVTRPT MESSAGES.....	7
2.3.1 Description of the test performed	7
2.3.2 Initial results	8
2.3.3 Proposed optimisations.....	8
2.3.4 Summary of tests results.....	11
2.4 EVTRPT IN REPLACEMENT OF RESPONSE MESSAGES (RESDES, RESCON, RESDIT).....	12
2.4.1 Transport information	12
3 ITMATT: REMARKS AND PROPOSALS	14
3.1 SCOPE	14
3.2 DESIGN/MAINTENANCE OF THE STANDARD.....	14
3.3 USAGE.....	14

Introduction

This document describes a study performed by the EMS Unit and the PTC (Postal Technology Centre) of the UPU on the new EDI messages.

It is meant as a working document for the next meeting of the EXG in Lisbon, with proposals and questions.

We will continue our work on the new EDI messages after the Lisbon meeting.

In order to be easily referenced, each proposal has a unique number within the document.

1 General remarks: scope of the new messages

The new EDI messages have been designed to cover all possible usages by any postal parties involved in the process of international mail items.

1.1 Parties involved

The EDI networks (GXS / POST*Net) are well adapted when each partner exchanges the same type of information with many partners worldwide.

This is the case with exchanges between postal administrations about international mail.

Questions and remarks:

Is it going to be the same situation between a postal administration and parties others than postal administrations?

In case of bilateral exchanges (say between a postal administration and a delivery agent within the same country), is it worth going through an EDI network?

1.2 An open standard

The new EDI standards have been designed to be open: each new message type offers many possibilities, so that groups of partners can decide what they want to exchange among what is available. Most likely, each group of partners will exchange a (small) subset of what is offered by each message type; but different groups of partners may not exchange the same subset.

In other words, the new standards are the union of all possibilities, as opposed to the intersection. The message definitions do not tell each party what they should exchange, but what they can exchange.

This is particularly true with the messages ITMATT and AGGATT.

Example:

Posts A, B and C want to exchange info 1 and info 2.

Posts D and E want to exchange info 2 and info 3.

⇒ The new standards provide the possibility to exchange info 1, info 2 and info 3.

In order to use the messages efficiently, postal parties need to agree between themselves.

Questions and remarks:

How to use the standards? Should the party sending a message provide all possible information (all the information they have?), even if it has no interest for the receiving party?

Or should both parties agree on a subset?

In case exchanges are based on agreements, who will decide how to use the standards?

Should the decision be left to each postal administration, with the risk of having many different agreements between posts? Or should the standards group or another entity propose a subset of information to be used between postal administrations at international level?

The definition of each piece of information to be transported should be as precise as possible, to avoid multiple interpretations. The name given to a field should clearly reflect its content.

We also recommend avoiding free text fields as much as possible, because they cannot be analyzed by computer systems. A code should be preferred whenever possible.

2 EVTRPT: remarks and proposals

2.1 Introduction

For EVTRPT, we focussed on the following topics:

- Transition period: possibilities of conversions from EMSEVT/TRAKIT v2
- Volume of EVTRPT messages and comparison with TRAKIT v2
- Possibilities of usage in replacement of RESDES, RESCON and RESDIT

2.2 Transition period

In the new set of messages, EVTRPT is the only new message coming in replacement of existing ones: EMSEVT/TRAKIT and all the RESxxx messages (RESDIT, RESDES and RESCON).

What is going to happen with item event messaging during the transition period, when some posts start handling EVTRPT and others continue handling EMSEVT/TRAKIT only?

2.2.1 Message conversions

EVTRPT provides 2 event types corresponding to posting:

- XP-10: posting
- PP-10: acceptance for processing

If both codes are possible, both must be handled by applications.

It does not seem necessary to have 2 event types with the same meaning in EVTRPT.

Proposal n°1:

In order to ease message reading and conversion programs, it is proposed to restrict the list of valid combinations event type – location type used in EVTRPT so that each event has a unique representation. For example, PP-10 should not be a valid combination, as it is a duplicate of XP-10 (or the other way around).

There might be other examples (to be studied).

Proposal n°2:

In the PIA segment, the event.associated-entity.type is conditional.

What happens if it is not provided? Does inheritance apply? What if it is not provided in the first occurrence of a PIA segment within a set?

Proposal: give event.associated-entity.type the same behaviour as event.entity.type.

(Note: this is not directly related to message conversion)

2.2.2 Distributed conversions or centralized conversions?

We consider that the new messages will be exchanged on the existing networks: GXS and POST*Net.

On each of these 2 networks, there are 2 possibilities for handling the transition period:

1. Centralized conversion: the network provider performs conversions centrally, so that each post connected to this network can choose to handle TRAKIT/EMSEVT (depending on the network) or EVTRPT for all their partners
2. Distributed conversion: the network provider does not provide any conversion, forcing each post to exchange the correct message type depending on the partner, or to use a conversion program locally

The table below lists the advantages and inconveniences for postal administrations if **centralized conversion is provided** by the network:

Advantages	Inconveniences
Easier transition: can continue using a system handling the previous standards only, but communicate with partners handling the new set; when switching to the new standards, no need to continue handling the old set	Risk to send information that is lost during the conversion, without knowing it
Easier management of EDI exchanges: no need to know what exact type of message is handled by each partner; no management action required when a partner switches from the old set to the new set	If centralized conversion is not optimised, then conversion of TRAKIT/EMSEVT to EVTRPT will increase size (see next section)

So it seems better for all postal administrations if centralized conversion is provided.

We do not have any special proposal here. Any reaction from the working group is welcome.

2.3 Size of EVTRPT messages

The size of a message, in other words, the amount of bytes needed to store the information in the message, is important for posts and network providers, as there are costs involved. The evolution in the past few years was that transport costs are less and less based on volume, as the internet is more widely used to transport the messages.

Nevertheless, there are network costs that depend on volumes. This may be transparent for postal administrations if their network provider proposes a fixed price that includes these costs.

But the costs are there, still, and therefore, it is important to check the impact of the new messages in terms of volume.

2.3.1 Description of the test performed

The test consisted in converting a large number of TRAKIT v2 messages into EVTRPT messages and checking the size differences.

Conversion program:

A specific computer program was written to perform the conversion automatically. In order to get objective results, the program sorts the TRAKIT events (by event type, date-time, etc.) so that the conversion takes full advantage of the inheritance possibility in EVTRPT.

Limitations of the program:

- The content of the BGM segment (document reference number) is set to a static and meaningless value
- The conversion program only accepts TRAKITv2 messages;
- The conversion program generates a carriage return at the end of each segment; this makes it easier to read the message, but it generates extra space (but all the messages used for the test had carriage return in their original format)
- If the last digit or last 2 digits of the time are '0', it would be possible to drop them to save space (example: '1950' -> '195', '2000'-> '20'), but this could be considered as an abuse of the format definition, as such a usage is meant when no more precision is known (for example, '195' should mean between 19:50 and 20:00)
This "abusive" optimisation was not implemented
- Location codes: when the location code in TRAKIT does not start with the country code, it is kept as it is in EVTRPT, which means that the country code is not added as a prefix, although it

should, in order to be compliant with S25; this might have an impact on conversion rates (EVTRPT would be slightly bigger)

For information, the size of a TRAKIT message is approximately 25% bigger than the EMSEVT for the same contents.

So if we want EVTRPT to be more compact than an EMSEVT, it should be at least 25% smaller than a TRAKIT.

Data used:

The data used consisted of all TRAKIT messages that transited on POST*Net on 22 July 2003 (date chosen arbitrarily, just because its data was easily accessible by the tester, in the PTC).

All TRAKIT messages generated on that date, from 0:00 to 23:59 were converted to EVTRPT with the conversion program described above.

There were 4,789 TRAKIT messages, with a total size of 5,263,887 bytes.

Note: all the TRAKIT messages read contained a carriage return at the end of each segment.

As the conversion program also generates this carriage return, we considered that the extra volume before and after the conversion could be ignored.

2.3.2 Initial results

EVTRPT size: **-1.2% compared to TRAKIT**

Test with precision of 10 minutes in EVTRPT: **-6.0%** compared to TRAKIT

Note: in the rest of the tests, the precision is kept to 1 minute.

A summary of all test results is provided at the end of the document, to ease comparisons.

2.3.3 Proposed optimisations

All the proposals below are intended to optimise the size of EVTRPT.

For each proposal, the impact was tested on the same set of data as described in the previous section (the conversion program was modified accordingly for each proposal). The results of the associated test are provided.

Note that, unless otherwise specified, the changes to the program were made individually, not combined.

Proposal n°3:

LIN, Seg 5 and Seg 17

Item identifier: for 13 character identifiers (items) or receptacle identifiers (29 char), is the prefix ('J10A') compulsory?

In order to ease the transition from EMSEVT/TRAKIT and reduce size, it is proposed to tolerate identifiers without this prefix.

The rule could be for example that if the first character of the ID is 'J' and the second is a digit, then the prefix is considered present, otherwise it is not.

Test result: **-8.3%** compared to TRAKIT

Note: an alternative approach would be to add an attribute 'item.ID-type' to indicate whether the ID is a UPU standard (13 characters), a license plate, etc.

Proposal n°4:

LIN, Seg 5 and Seg 17

LINE ITEM IDENTIFIER: this field is a number incremented in each LIN, it does not contain any postal information; EVTRPT can still be constructed and is still readable without it.

The purpose of this field appears to be to underline the importance of the order of fields within the message, since information is not repeated for each event when not necessary (inheritance).

But this field makes the construction of EVTRPT more complicated and increases the volume of the message.

The risk of having messages badly built because of dropping this LINE ITEM IDENTIFIER seems very small: messages are built by automated systems that can easily order information.

Let us imagine that this field is kept and a message is generated with the wrong order (say LIN+2 appearing before LIN+1): is it going to be checked by the network provider and/or the receiving party? If it is and if a problem is detected, would the whole message be rejected?

In order to simplify the process, both for the sending and receiving party, and to minimize the size of EVTRPT, it is proposed to remove the line item identifier

Test result: **-4.9%** compared to TRAKIT

Proposal n°5:

UPU standards glossary - DTM, Seg 23

In many cases, events reported in a message happen on the same day, but it may be at different times of the day; it would be nice (size reduction) if a specific date format could be provided to only contain the time; in EVTRPT, the rule would be that if only the time is provided in Seg 23, it means that the date is the same as the one provided in the last DTM Seg 23 and Seg 11 that contained the full date.

In order to guaranty inheritance, we could decide that in DTM Seg 11, the full date must be provided.

Format proposal: 'Dhhmm...', 'D' indicating that the date is not provided

Example: 'D1355': 13h55 on the same day as the previous DTM segment containing a day

Test result: **-5.2%** compared to TRAKIT

Proposal n°6:

UPU standards glossary - DTM, Seg 11 and 23

In the UPU standards glossary, the date-time format now has a 4 digit format for the date, including century and millennium, as opposed to 2 digits previously.

It was indeed a problem to have only 2 digits around year 2000, but the problem will not appear again before year 2099-2100.

Do we expect the new standard to last 97 years without any possibility to change it?

The extra information increases the size without any benefit.

Proposal: change back the year format to 2 digits, keeping in mind that it will need to be changed to 4 digits by 2090.

Note: this change would also apply to the other messages

Test result: **-2.6%** compared to TRAKIT

With proposals 3 to 6 combined, the test result is: **-19.2%** compared to TRAKIT

Proposal n°7:

DTM, Seg 23

The new date format gives the possibility to include an offset with UTC in DTM, Seg 11 and Seg 23. Exchanging this information can be very valuable for the receiving party and also the network provider (for calculations).

But if this offset is provided in each DTM segment, it also generates additional volume.

It is proposed to add the following rule to the EVTRPT construction: if an offset with UTC is provided in a DTM Seg 11/Seg 23, it is valid for subsequent DTM Seg11/Seg 23 until another offset is provided or until the next set (GRU+EVT).

This is actually an extension of the inheritance process already in place.

Impact with test: not applicable, since this UTC offset cannot be transported in TRAKIT

Proposal n°8:

In EVTRPT, an event is defined by the combination event type + location type. Currently, these two pieces of information are in different segments: event type in STS and location type in LOC.

It would seem more logical to include the location type in the 'STS' segment.

Here is the proposed modified STS segment:

STS	STATUS	C		
C601	STATUS CATEGORY	C		
9015	Status category code	M	an..3	event.status
C555	STATUS	C		
4405	Status description code	M	an..3	event.code
C556	STATUS REASON	C		
9013	Status reason description code	M	an..3	event.location-function
C556	STATUS REASON	C		
9013	Status reason description code	M	an..3	event.reason

The change cannot be performed without a change of the LOC segment: this segment contains a mandatory field of 3 characters.

No solution was studied for this. Instead, the next proposal below is to reorganize other segments, solving the problem raised here.

In other words, the solution proposed at the end of the next proposal also includes the change in this proposal.

Proposal n°9:

The information stored for most events is:

- object ID
- object type
- date-time
- location code

The current structure stores this information over several segments.

It is proposed to reorganise the structure of the message in order to compact the storage of these pieces of information.

Proposal:

LIN and LOC segments can be replaced by a PCI segment (then the PCI segment already existing in EVTRPT will be replaced by a LIN segment!)

PCI in replacement of LIN and LOC:

PCI	PACKAGE IDENTIFICATION	C		
4233	MARKING INSTRUCTIONS CODE	C	an..3	event.entity.type
C210	MARKS & LABELS	C		
7102	Item identifier	M	an..35	event.entity.ID
7102	Event location code	C	an..35	event.location.code
7102	Event sub-location	C	an..35	event.sub-location
7102	Event detector	C	an..35	event.detector

LIN in replacement of PCI:

LIN	LINE ITEM	C		Event capture method and status
1082	LINE ITEM IDENTIFIER	M	an..6 (an..5)	event.entity-ID-capture-method
1229	ACTION REQUEST/NOTIFICATION DESCRIPTION CODE	M	an..3 (n..2)	Event.entity-ID-capture-status

Note that with this major change, the inheritance process is slightly more difficult to handle.

Example:

```
PCI++EU446692255US:D1305::PLWAWA'
PCI++EU844084301US:D1335'
```

This means that for the second item, the location is also 'PLWAWA'. Now, if we want to record an event that is not associated with any office, we must do the following:

```
PCI++EU446692255US:D1305::PLWAWA'
GRU+EVT'
PCI++EU844084301US:0308141335'
```

Impact with test (combined with proposals 3 to 6): **-24.0%** compared to TRAKIT

Proposal n°10:

Provided the proposal above is accepted, it seems possible to also merge the DTM segments within the PCI segment, to compact even more.

It is also proposed to re-introduce an *event.signatory* element, much easier to use than the *postal-address.delivery-point-specification.free-text.line1*.

Note that the *postal-address.delivery-point-specification.free-text.line1* does not clearly specifies that it must contain the signatory name, so its usage might lead to confusions and problems between parties.

This new element *event.signatory* is also part of the PCI segment.

Resulting PCI segment:

PCI	PACKAGE IDENTIFICATION	C		
4233	MARKING INSTRUCTIONS CODE	C	an..3	event.entity.type
C210	MARKS & LABELS	C		
7102	Item identifier	M	an..35	event.entity.ID
7102	Event latest date-time	C	an..35	event.latest-date
7102	Event earliers date-time	C	an..35	event.earliest-date
7102	Event location code	C	an..35	event.location.code
7102	Event sub-location	C	an..35	event.sub-location
7102	Event detector	C	an..35	event.detector
7102	Event signatory	C	An..35	event.signatory

Impact with test (all optimisations combined, time precision kept to 1 minute): **-32.8%** compared to TRAKIT

2.3.4 Summary of tests results

Test #	Test performed	Result (compared to TRAKIT)
1	Initial conversion	-1.2%
2	Precision of 10 minutes only	-6.0%
3	Proposal n°3: no 'J10A' tolerated for item ID	-8.3%
4	Proposal n°4: no line count in LIN segment	-4.9%
5	Proposal n°5: date format with time only	-5.2%
6	Proposal n°6: date format 'yy' instead of 'yyyy'	-2.6%
7	Proposals 3 to 6 combined:	-19.2%
8	Proposals 3 to 6 combined + 8 and 9 (STS changed, PCI iso LIN)	-24.0%
9	Same as above, + date-time information in PCI iso separate DTM	-32.8%

Details per message size group:

Message size (bytes)	Number of msg	Test 1	Test 2	Test 7	Test 8	Test 9
<500	3,056	19.3%	17.5%	10.2%	5.4%	-3.3%
500< <1,000	785	9.6%	5.5%	-7.3%	-12.2%	-23.6%
1,000< <2,000	454	5.0%	-0.4%	-15.3%	-20.4%	-32.1%
2,000< <5,000	340	-3.8%	-10.0%	-25.2%	-30.4%	-39.4%
5,000 < < 10,000	96	-10.7%	-17.4%	-32.8%	-36.9%	-44.5%
> 10,000	58	-13.1%	-17.8%	-34.0%	-38.5%	-45.2%
Totals	4,789	-1.2%	-6.0%	-19.2%	-24.0%	-32.8%

Remarks:

- For very small messages, except for the final test, EVTRPT is always bigger than TRAKIT.
- The bigger the TRAKIT message, the better the conversion ratio

2.4 EVTRPT in replacement of response messages (RESDES, RESCON, RESDIT)

It seems important to consider EVTRPT in replacement of the response messages together with the other usage of EVTRPT (replacement of EMSEVT/TRAKIT) because there might be impacts in the message structure and also because posts migrating to the new standards may want to use EVTRPT in replacement of response messages as well, in order to simplify their system (1 message type instead of 4).

2.4.1 Transport information

Transport information is a very important part of the current response messages.

Note: it is also used in EMSEVT, but not in TRAKIT.

EVTRPT does not contain fields that refer directly to transport information.

It is possible to store a flight number in segment PIA, field 'event.associated-entity.id', associated with field 'event.associated-entity.type' containing value 'T' (transport).

Example:

PIA+1+SN432:T'

(the flight number is SN432)

M35 describes the concepts and representation of transport information in the new messages, but it is not used by EVTRPT.

The format used in segment PIA is very free, leading to potential problems:

- There is no information on the transport type; partners must agree to always use flight numbers (for instance), otherwise there is a risk of misinterpreting the data received.
Note: in RESDES, for instance, the field 'mode-of-transport' clarifies this
- Carrier code and flight number are merged, without any defined format; each post may choose a different format, leading to problems for the receiving party
Example: flight Qantas n°123 can be coded 'QF123', 'QF 0123', 'QF 123', etc.
The format provided would even allow sending '123' only, or 'QF' only or 'Qantas 123'.

Proposal n°11:

Segment 'PIA' does not seem very adapted to the storage of transport information.

Would it be possible to consider the addition of another segment for transport? This segment would clearly define the transport elements (to be added to the list of data elements in EVTRPT) and would add rules to guaranty that all necessary information is put in the message.

The information to be sent is a subset of M35.

3 ITMATT: remarks and proposals

3.1 Scope

ITMATT is meant to be exchanged between postal handling organisations.

ITMATT is designed to allow the transmission of all possible mail item information; in other words, the list of attributes is meant to be exhaustive, so that each postal handling party can pick the information they are interested in exchanging with their partner(s).

3.2 Design/maintenance of the standard

We believe that the list of attributes provided in ITMATT can never be exhaustive: a post might start providing an added value product requiring new attributes that are not in ITMATT.

Example: we do not see the possibility to store the sender's email address, although we could imagine cases in which the sending party might want to send this information to the receiving party or an intermediate party (or a centralized system such as CAPE), so that an email can be automatically sent by this party to the sender when the item is delivered.

Second example: insured value must be expressed in SDR (unlike other monetary amounts in the messages). 2 partners might want to exchange it in another currency, thus not respecting the standard as it is.

Note that these are just examples, not a request to modify the standard.

We believe that maintaining ITMATT up to date will require a lot of work: whenever a post asks for a new attribute (and we believe it would happen often if ITMATT was used as it is planned), the topic will have to be discussed by the working group, then a new version of the message will have to be created.

3.3 Usage

The fact that ITMATT will require many regular updates in order to be as exhaustive as possible will cause problems for the users, who will have to keep modifying their system to keep up to date with the standards.

There are many ways of using ITMATT, as most of the information is conditional.

This means that for a post using ITMATT, it may be necessary to send a certain subset of information to a first partner, another subset to a second partner, etc.

Maintaining this is very troublesome and complicated, therefore costly, with the risk that only a few postal handling organisations will do it correctly.

At international level, it is indeed interesting to exchange some item information in order to avoid re-capturing it: the sender and addressee information for example.

Nevertheless, there is an important issue that needs to be addressed: languages.

With the EDIFACT constraints, the set of characters allowed is very limited. No accents are permitted and the alphabet is Western European.

This means that information must be recorded in the Western European alphabet in order to be transmitted.

Let us imagine exchanges between Russia and Ukraine: it would make sense to exchange sender and addressee information in Cyrillic alphabet, since both countries use this alphabet. This is not possible with the EDIFACT standard.

-> As it is, the standard is limited to countries using the Western European alphabet.

This is a problem for sender and address information, in an international standard.

Proposal n°12:

The language issue needs to be addressed

Proposal n°13:

Study the possibility to provide a limited version of ITMATT; this limited version would fit the needs of postal organisations at international level: the post in the sending country would use this limited version to send sender/addressee (and maybe more) information to the post in the final destination country and the receiving postal organisation would use it to inform the sending post on additional item attributes.

It might be possible to put in place and maintain such a limited version, without modifying the standard:

- Provision could be made in the item.attributes-version-number field (Seg 1), for example by reserving values 0 to 999 for the limited version
- The list of authorized fields would be listed; this list would be a sub-set of ITMATT.

Note: if members of the EXG find this approach interesting, we can propose a subset in a future meeting. We could not do it this time, due to the fact that many attributes of ITMATT are still unclear to us, so we need to investigate further.

In fact, we might end up having 2 subsets: one subset for the postal administration sending the mail, the other subset for the postal administration receiving the mail.