BANK OF ENGLAND
RTGS

Record of Third Interoperability Working Group

10 January 2018

NPSO	Barclays
RBS	WorldPay
SWIFT	CLS
JP Morgan	FPS
Bacs	Bank of England
LLoyds	

Extended Character Sets

- The Bank presented a proposition for introducing extended character sets. This had been written following discussions at the second Interoperability Working Group.
- ISO 20022 supports UTF-8 character encoding enabling the use of the full range of Unicode characters. Given the technical feasibility, there are two questions to answer in relation to extending character sets: do we wish to introduce this functionality? If so, how do we implement it?
- The proposal states that the Bank believes there are benefits to end-users, payment institutions and to the central infrastructure from introducing extended character sets. The proposal sets out the conditions for implementation, the HVPS+ guidance and some principles that could guide the implementation.
- The group supported the proposal to introduce extended character sets, recognising that there are clear drivers, such as transparency obligations, for wanting this enhancement. The group discussed that extended characters could improve compliance while reducing risk. It would also help to facilitate efficiency improvements for payment processing by reducing manual intervention.
- In the proposal presented, the Bank focused on the international use case. Comments from the working group highlighted that there is also a domestic use case for extended character sets. For example, names requiring accents cannot be represented in the current limited character set.
- Regarding implementation, the group acknowledged the challenges which are as much to do with common consistent practices across the entire payment chain as they are to do with introducing technical capability in the renewed RTGS system. To ensure future proofing extended character sets ought to be supported. The group agreed with the proposal that the introduction of extended characters should be focused on specific data fields where these would prove to be most valuable, particularly name, address and remittance fields.

- Further consideration will be necessary on the transition and adoption of the extended character sets in the renewed RTGS system and the retail schemes. The wider industry consultation on ISO20022 planned for later this year could be used to validate the key drivers and timing toward the introduction of this functionality. It may be appropriate to consider a phased approach, being cognisant of the end-to-end processes, impacts and obligations to accept, process and pass on characters other than the current Latin character set.
- **Action: Bank to review the proposal following the Working Group's comments and finalise a draft proposal and supporting reasoning for use in the consultation document.**

Use Cases

- A sub-group were tasked with devising a number of use cases associated with clearing scheme payments in the UK.
- The sub-group set out a number of end-to-end payment use-cases that will form the basis for analysis of the requirements for an interoperable core credit message.
- **Action: the sub-group are to develop the most complex use case, including ISO 20022 terminology, to propose which ISO 20022 messages should be used end-to-end and to consider the roles of each actor in the transaction, including the end customers (initiating and receiving party).**

Routing Rules

- Following the discussion and progress made at the previous meeting, further work was being undertaken by retail payment schemes to comprehensively set out the structure of identifiers and the logic used for routing, clearing and settlement across each of these schemes, including CHAPS.
- This will enable the group to discuss possible efficiencies, the alignment of the business logic that surrounds routing and the determination of an ideal end-state, which will also need to align to the NPA design.
- **Action: NPSO to depict the current routing status quo in order to facilitate the forthcoming discussion.**
- It was recognised that the NPSO is not yet in a position to finalise an optimal routing solution. However, work is underway to determine the optimum for both current operations and strategic change to allow this to be incorporated into the RTGS ISO 20022 standards design. This has the potential to be a key enabler of interoperability.

The Business Application Header

- The Bank introduced the ISO 20022 Business Application Header (BAH) to the working group. This is the separate *head.001* message; part of the HVPS+ recommended implementation of ISO 20022.
- The group discussed the merits of introducing the BAH as a means of facilitating interoperability between the wholesale and retail implementations of ISO 20022, agreeing in principle that it should be adopted. It is anticipated that the BAH could contain scheme specific data as part of the implementation of ISO 20022, such as settlement cycle number, which could facilitate the design of interoperable messaging.
- Further investigation is required to determine the usage of the BAH (*head.001)* for bulk payments. Input from the HVPS+ group would be beneficial as would input from other FMIs utilising the BAH.
- The next step will be to determine what information should populate the BAH. Members of the working group will provide information at the next meeting to facilitate this.
- **Action: SWIFT to present the HVPS+ advice on the BAH.**
- **Action: NPSO to present advice from other jurisdictions' implementation of the BAH.**

A.O.B.

- The Chair informed the group that Payments UK were now formally part of the New Payment System Operator (NPSO) since the 2 January 2018.
- The Chair asked that the next meeting look at error codes in payment instructions and to begin to focus on the detailed structure of the message.
- The next meeting is Thursday 8 February (note: the meeting has subsequently changed to 26 February).

Appendix A: Reading Pack

Introduction

The Interoperability Working Group is advising the Bank of England on technical issues arising in the implementation of ISO 20022. The primary focus is on the core of the credit message for all UK payment systems and on developing a framework for developing and enhancing the messaging standard going forwards. The overall objective is improved interoperability between UK payment schemes whilst harmonising with international high value payment schemes.

At the second Interoperability Working Group we discussed principles for good interoperability, priorities for the working group and begun the discussion on routing rules and character sets. Further information is provided in the minutes, attached at the end of the reading pack.

Agenda for the third Meeting

The group will meet for the third time on **Wednesday 10th January 2018 at 12:00**. The session is scheduled to run until 17:00, but will hopefully conclude earlier. Lunch will be provided.

It is proposed that the following topics are discussed:

- **Review of the Bank's proposal on extended character sets**
  - Following the discussion at the previous meeting.
  - See Section A of the reading pack.

- **Use cases**
  - CHAPS, CLS and SWIFT will provide some material to consider exploring domestic and cross-border payments and the agents involved. This will inform the discussion on routing rules.

- **Routing rules**
  - The discussion at the last working group emphasised the challenges in this space, this next discussion will begin to set the foundations for aligned routing.

- **The Business Application Header**
  - To form a view on if/how to implement it in UK payments.
  - See Section B of the reading pack.

Section A: Proposition for the implementation of an extended character set in ISO 20022

At the second meeting of the Interoperability Working Group attendees were asked to discuss extended character sets for pacs.008 and pacs.009 messages. The Working Group advised the Bank of the strategic benefits of providing this flexibility. The Bank committed to developing a more detailed proposition for the implementation of an extended character set, and to bringing this back to a subsequent meeting for discussion.

Please view the following text as this proposition. If agreed, this will form the basis of the text for the relevant section of the forthcoming consultation document.

Proposal

The renewed RTGS service will support the ability to transmit an extended character set including non-Latin alphabets and a wider range of special characters. The implementation of this functionality would restrict using extended characters to certain key fields within the pacs.008 and pacs.009 messages.

It is envisaged that this will apply for across both the high value and the retail payment schemes for the purposes of interoperability.

Introduction

The SWIFT character set for the FIN network is limited to alphanumeric, i.e. upper & lower case Latin alphabetic characters and Arabic digits (more details in Annex 1). CHAPS currently uses this character set in MT messages on the FIN network.

The Bank will initially adopt ISO 20022 with an XML syntax on the InterAct network. InterAct and ISO 20022 can support the full range of Unicode characters.

ISO 20022 is restricted to only UTF-8 at present. UTF-8 can encode most of the characters in the SWIFT character set in one byte. More information on UTF-8 and UTF-16 can be found in Annex 2 and Table 1.

The HVPS+ taskforce has recommended that users of ISO 20022 restrict themselves to the Latin characters in use for MT messages described above (see also Annex 1). They further advise that only where bilaterally agreed within a community, could extended characters be used in the remittance information fields and non-financial institution identification fields. Thus our proposal is within the approach outlined by the HVPS+ taskforce.

Outline of proposition

In particular, attendees should consider the scope of this proposal and the conditions for implementation.

The ability to submit non-Latin characters would be constrained to a limited set of fields within the message, as indicated by the HVPS+ guidelines. This is explained in greater detail in ‘Principles

for use' on page 6. **Should direct participants be mandated to have the functionality to receive an extended character set?**

Those parts of the message being read by a computer should retain a restricted character set in accordance with global best practice of using Latin characters; those parts that are read by a human should have expanded set. The categories of field that we consider for expanded characters are:

- Name & Address for non-FI identification (except for certain elements, such as the drop-down ISO country code list comprising basic Latin characters such as 'GB')
- Remittance information and other related data elements (see exceptions below)

Out of scope are the following fields within the message:

- Fields containing the core settlement instructions, such as the account number
- Fields containing instructions about message handling in the payment scheme, such as the priority field in CHAPS messages
- Specific fields for the use of a given meta-data tag which contains only Latin characters.

The additional functionality would not be available for users from the go-live date. Extended character sets would become available later, **subject to the following conditions** being met[1]:

- Introducing this functionality before all direct participants have connected to InterAct is not an option. FIN cannot support extended character sets, thus the migration from MT messaging on the FIN network to ISO 20022 messaging on the InterAct network must be complete for CHAPS participants.
- Ideally, the HVPS+ guidelines would indicate that expanded character sets are expected to be available in remittance information and non-financial institution identification fields.
    - In order for use of expanded characters to be widespread there must be a range of communities that permit this facility owing to the cost of updating legacy systems in payment institutions and ordering institutions.

Rationale for expanding the character set

Allowing an extended character set provides advantages when users of CHAPS use characters outside of the Latin set to represent information such as their name, location or business activities.

End-users

- Ease of use and efficiency are improved if corporates/individuals/institutions using CHAPS can input names, addresses and remittance information using special characters and non-Latin alphabets.

[1] This does not preclude the Bank developing the functionality to transmit data extended characters in advance of this go-live. It may be that as part of the wider renewal of the RTGS service this functionality is added.

- It can reduce costs if corporate databases do not need to be translated into the current restricted character set.

Direct participants / payment institutions

- Compliance screening is enhanced as truncation or misrepresentation through translation is reduced.

Central Infrastructure

- Developing the functionality to transmit characters outside the current SWIFT character set reduces future development costs if the SWIFT characters are expanded or an alternate network connects to RTGS with this functionality.

- This ensures that the domestic infrastructure is well placed to serve international strategic objectives. A large proportion of CHAPS traffic is international and may benefit from wider character sets being available.

Principles for use

Use of extended characters may be optional. As such there must be provisions in place to ensure that where those using extended characters meet those not using them there is no loss of data. These form the principles for use.

The principles set out in the European Payments Council's SEPA requirements best practices document form an excellent basis for this work.[2]

Further to this:

- The decision on whether to send messages using this expanded character set may be left to direct participants of payment schemes. Following the SEPA principles and the Wolfsberg Group Payment Transparency Standards,[3] intermediary banks should forward on any text received and should not alter the text when doing so.

- Receiving institutions should use mapping tables where possible if they are not able to process the extended character set directly.

- Payments should only be returned insofar as the core settlement information contains non-Latin characters, or if the inclusion of non-Latin characters beyond any given mappings results in the payment not being capable of being processed within the receiving institution.

Technical implementation

- Non-Latin characters and additional Latin characters will be encoded in either the UTF-8 or the UTF-16 format. UTF-8 is the only encoding supported by SWIFT at present. Given the emphasis on Latin characters in payment instructions, UTF-8 is more efficient in terms of bits required for each encoding. See Annex 2 and Table 1 for more details.

[2] https://www.europeanpaymentscouncil.eu/sites/default/files/KB/files/EPC217-08%20Draft%20Best%20Practices%20SEPA%20Requirements%20for%20Character%20Set%20v1.1.pdf
[3] http://www.wolfsberg-principles.com/pdf/standards/Wolfsberg-Payment-Transparency-Standards-October-2017.pdf

- As part of the release of an extended character **set the Bank will collaborate with the NPSO to publish a translation guide which will define some mappings between the basic Latin character set and the extended character set.** For example, the German character 'ü' could be mapped as 'ue'. This allows those computer systems that do not accept extended characters to pass on the full information they receive. The SEPA conversion table provides a guide for how this could be implemented.

- **It is not proposed that non-Latin alphabets be translated at the centre by RTGS.** The liability for the correct translation of non-Latin text would rest solely with the receiving and sending institutions and not the Bank of England.

Section B: The Business Application Header

Background

The Business Application Header (BAH) is a separate ISO 20022 message (head.001) which may be combined with another ISO 20022 message to facilitate message processing. It is distinct from a batch, file or transport header.

The purpose of the BAH is to provide a consistent and predictable way for data to be conveyed with the message, regardless of implementation factors such as the choice of network. This doesn't prevent such data being conveyed either within the ISO 20022 message definition itself, or as part of a network header. A visualisation is provided on page 8.

The BAH can be used at the level of the entire implementation, groups of messages or at the individual message level. It typically contains information regarding:

- Sender
- Receiver
- Message definition identification
- Unique business message identifier
- Digital signature to confirm the message content is unchanged
- gpi reference

Payment messages were defined in ISO 20022 before the BAH and can be used without one. ISO 20022 messages contain much of the information in the BAH in fields already.

Recent versions of some ISO 20022 messages have had the fields that are included in the BAH removed from the message set.

Are BAHs implemented elsewhere?

SWIFT recommends use of BAH in MX:

- T2S, EBA Euro1 Clearing, Hong Kong, Australia NPP (retail), India, Brunei, Jordan & Bangladesh all use the BAH.
- Both Fedwire & CHIPS will use BAH.
- T2 will probably use the BAH, this is to be discussed in phase 2 of their project.
- TIPS and Canada will not use the BAH.

Further considerations regarding implementation

Use of the BAH in the V-shape topology means that the underlying pacs.008 can be digitally signed, thus showing it is unchanged; SWIFT recommend using the BAH for infrastructures maintaining a V-Shape network configuration.

Fedwire will be validating the pacs.008 content and will not be using signatures for their messages.

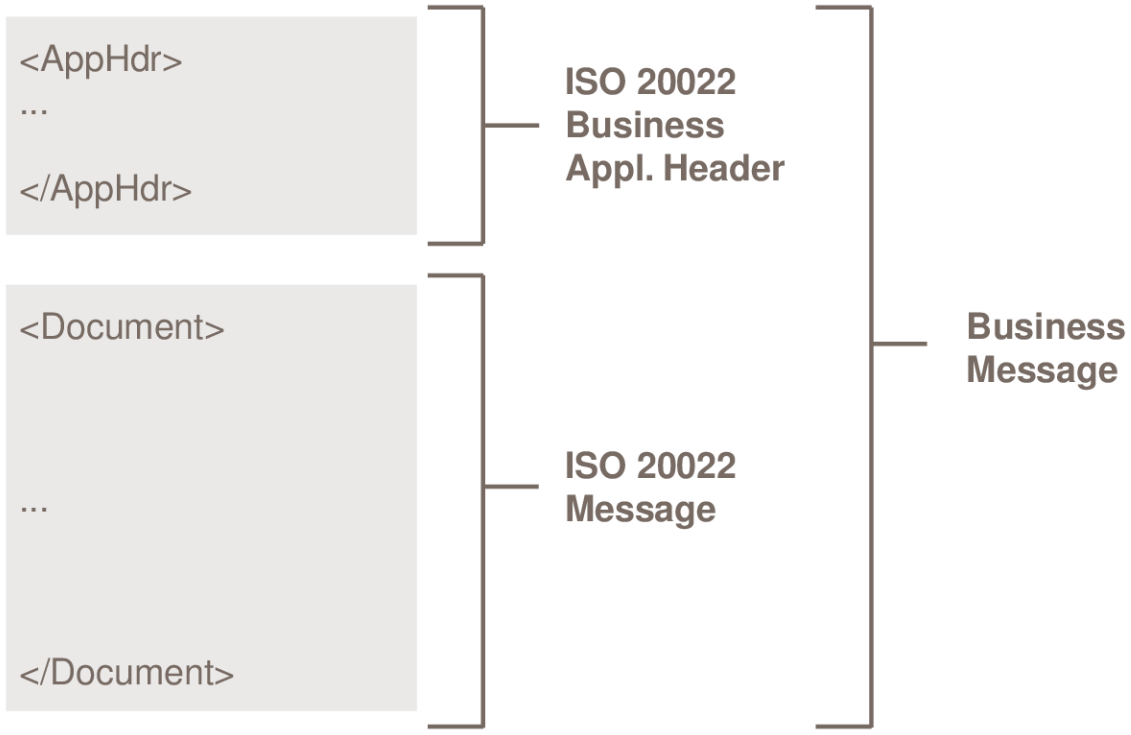
T2 will not have end to end signing, just point to point. It is possible that T2 being network agnostic may cause problems for the signature. If the MI needs to change the message content, then just a subset of the message could be signed.

The working group should consider the merits of the Business Application Header (BAH) in the context of security standards and security signatures for messages.

The group should consider the BAH in both the retail and high value context and how these separate implementations of ISO 20022 will need to interact.

Visualisation

The below is a diagrammatic explanation of the BAH taken from the ISO 20022 Business Application Header Message User Guide. It shows how the BAH interacts with the pacs.008 or pacs.009 message to form a complete Business Message.



Annex 1:

SWIFT Character Set:

- a b c d e f g h i j k l m n o p q r s t u v w x y z
- A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
- 0 1 2 3 4 5 6 7 8 9
- / - ? : ( ) . , ' +
- Space

Examples of non-Latin characters include:

- Chinese [你好]
- Arabic [مرحبا]
- Unicode emoji characters [☺]

Examples of special characters include:

- @
- €
- &

End-to-end

Annex 2:

UTF-8 and UTF-16

ISO 20022 uses XML 1.0. XML 1.0 supports UTF-8 and UTF-16.

UTF-8 is a methodology for encoding characters that can encode all 1,112,064 of the valid Unicode code points and all 136,690 characters. It encodes characters of variable length using one to four 8-bit bytes.

UTF-16 is an alternate character encoding that can also encode all 1,112,064 of the valid Unicode code points and all 136,690 characters. It encodes characters of variable length using two or four 8-bit bytes.

UTF-8 is more efficient at encoding the first 128 characters (US-ASCII) than UTF-16, as this encoding only requires one byte in the former and always require 2 bytes in the latter. This makes UTF-8 more efficient for encoding text in modern European languages than UTF-16.

There are a number of code points that would require 3 bytes in UTF-8 but only 2 bytes in UTF-16. Thus, on occasion, UTF-16 is more efficient than UTF-8. Text in Chinese, Japanese or Devanagari (used in India and Nepal) is thus more efficiently encoded using UTF-16.

If we envisage greater use of special characters and non-Latin alphabets then it may be more efficient to use UTF-16 rather than UTF-8 (see table 1). If it is expected that the extended character set's use will be limited then UTF-8 will remain the more efficient encoding.

As the SWIFT network presently only supports ISO 20022 in UTF-8, a decision to develop UTF-16 functionality should be viewed strategically in support of future needs via the SWIFT network or an alternate network provider.

Table 1: Comparison of UTF-8 and UTF-16

Characters Supported	Bytes Required		Unicode Code Point Range	Most Efficient Encoding	Number of Characters
	UTF-8	UTF-16			
Basic Latin (US ASCII)	1 byte	2 bytes	U+0000 → U+007F	UTF-8	128
Most remaining Latin-script alphabets, Greek, Cyrillic, Coptic, Armenian, Hebrew, Arabic, Syriac, Thaana, N'Ko	2 bytes		U+0080 → U+07FF		1,920
Remaining Basic Multilingual Plane (virtually all characters in common use including most Chinese, Japanese and Korean) and many symbols	3 bytes		U+0800 → U+FFFF	UTF-16	53,244
Less common Chinese, Japanese and Korean, historic, mathematical symbols, emojis	4 bytes	4 bytes	U+10000 → U+10FFFF	UTF-8/UTF-16	81,461