

## **[Draft] Minutes of the Teleconference of**

### **ISO/IEC JTC 1/SC 34/WG4, 2014-08-21**

**Rex Jaeschke ([rex@RexJaeschke.com](mailto:rex@RexJaeschke.com))**

**2014-08-21**

#### **1. Opening remarks**

The meeting started at 13:05 GMT. The convener, Murata-san, welcomed everyone to the 61<sup>st</sup> teleconference of WG4.

#### **2. Roll call of delegates**

The following members were present during part or all of the meeting:

<b>Name</b>	<b>Affiliation</b>	<b>Employer/Sponsor</b>
Makoto Murata	WG4 Convener, JP	International University of Japan
Rex Jaeschke	Ecma, Project Editor	Consultant
Caroline Arms	Ecma	Library of Congress
John Haug	Ecma, US	Microsoft
Chris Rae	Ecma	Microsoft
Francis Cave	GB	Francis Cave Digital Publishing

Present were 6 people from 3 NBs, and 1 liaison.

#### **3. Adoption of the agenda**

The agenda (SC 34 N 2079) was adopted as published.

## 4. Administration

### Approval of Previous Meeting Minutes [WG4 N 0284]

The draft minutes were approved, as circulated.

### Outstanding Action Items

- Caroline will send email to the list on her thoughts re “content type” and “media type”. **Done** (see below)

### Report from the WG4 Secretariat

The following NBs and liaisons have registered delegates to WG4: BR, CA, CH, CI, CN, CZ, DE, DK, Ecma, FI, FR, GB, IN, IT, JP, KR, NL, NO, OASIS, PL, US, W3C, XML Guild, and ZA. All requests for additions, deletions, and changes to the delegate list should be sent to the WG4 Secretariat ([rex@RexJaeschke.com](mailto:rex@RexJaeschke.com)).

The WG4 email list is [e-SC34-WG4@ecma-international.org](mailto:e-SC34-WG4@ecma-international.org). The document repository is now at <http://isotc.iso.org/livelink/livelink?func=ll&objid=8912947&objaction=ndocslst>.

## 5. Revising Part 2 (Open Packaging Conventions)

### Review of a Personal Draft from Murata

#### 2014-08-13 Caroline Arms:

As promised, I have been reading the June MM0619 draft by Murata-san that is on the Assembla site.

Since most of the changes relate to moving most of Annex A into clause 8, I will send notes for clause 8 and Annex A separately -- not immediately. Below (and attached as a text file) are notes on the other clauses.

#### 1. Scope

General editorial comments:

(a) This clause shows several significant deletions. Somewhat equivalent text is in §7, “Overview”. We should check how things will look in the ISO preview (online browsing platform). Scope will show but I'm not sure that clause 7 would. We should check whether any text deleted from Scope should be recovered to get relevant words into either Scope or Overview.

(b) The ... bulleted list seems to be temporary. It definitely needs work for grammatical consistency.

Typo: "dissiminated" should be "disseminated"

## 2. Conformance

General editorial comment:

(a) Could do with improved distinction between statement about conformance for OPC (this part) and conformance in the multi-part standard.

## 3. Normative References

Presumably should update the Augmented BNF RFC 4234 to RFC 5234, since 4234 is obsolete. [Note: New text from Murata-san refers to RFC 2234 (an even older obsolete version) and uses EBNF rather than ABNF. Presumably, all references should use RFC 5234 and "ABNF."]

## 4. Terms and Definitions

I notice that some ISO standards have Notes associated with definitions. If this is acceptable practice we might consider whether any of the terms warrant a note as clarification.

Clearly this a clause to revisit when the rest of the document is done. The first term confused me but I assume we will make a holistic pass through Terms and Definitions later and haven't reviewed this clause in detail.

### **access style**

style in which local access or networked access is conducted

Proposed replacement: style in which local or networked access to a package is conducted

[Aside: Personally, I might use "mode" rather than "style", but I assume we are only addressing DRs and significant shortcomings in the document.]

Murata-san introduced the terms NFC and NFD for the two normalization approaches applied commonly for UNICODE. We should either spell out in full and prefix with "Unicode" (my preference) or add to Terms and Definitions.

## 5. Notational Conventions

Trusting this to Rex.

## 6. General Description

Will need to look back at this. Murata-san's changes have made Annex A informative, for example.

## 7. Overview

See my notes on 1, "Scope", above.

## 8. Package Model

Will deal with in separate set of notes apart from a few small details.

There are still several internal references and example captions that need A to be changed to 8.

I assume "EBNF" should be "ABNF". Latest RFC for ABNF is RFC 5234. 2234 and 4234 are obsolete.

## 9. Physical Package

## 10. Core Properties

## 11. Thumbnails

No changes have been made. Will need to check Annex references.

## 12. Digital Signatures

Ignoring for now.

Annexes

Ignoring from now. Will make notes on Annex A in conjunction with clause 8.

### **2014-08-15 Murata-san:**

> 1. Scope

> General editorial comments:

> (a) This clause shows several significant deletions. Somewhat equivalent text is in 7. Overview. We should check how things will look in the ISO preview (online browsing platform). Scope will show but I'm not sure that clause 7 would. We should check whether any text deleted from Scope should be recovered to get relevant words into either Scope or Overview.

The scope is not for expressing requirements. It is for describing what the standard does. I think that the package model, physical model, and the use of MCE have to be eliminated from the scope clause. All of them are details.

I agree that some of the deleted text might be useful in Overview.

> (b) The ... bulleted list seems to be temporary. It definitely needs work for grammatical consistency.

True. I'm not the right person for providing good English. Help!

> Typo: "dissiminated" should be "disseminated"

> 2. Conformance

> General editorial comment:

> (a) Could do with improved distinction between statement about conformance for OPC (this part) and conformance in the multi-part standard.

I'm wondering if this clause should be simply deleted. I will write more about this in another e-mail.

> 3. Normative References

> Presumably should update the Augmented BNF RFC 4234 to RFC 5234, since

> 4234 is obsolete. [Note: New text from Murata-san refers to RFC 2234

> (an even older obsolete version) and uses EBNF rather than ABNF.

> Presumably all references should use RFC 5234 and "ABNF."]

Yes, you are right. Thanks!

> 4. Terms and Definitions

> I notice that some ISO standards have Notes associated with definitions. If this is acceptable practice we might consider whether any of the terms warrant a note as clarification.

> Clearly this a clause to revisit when the rest of the document is done. The first term confused me but I assume we will make a holistic pass through Terms and Definitions later and haven't reviewed this clause in detail.

We have done some changes to this clause already, but I also think that we have to revisit it when the rest is done.

**> access style**

> style in which local access or networked access is conducted

> Proposed replacement: style in which local or networked access to a package is conducted

> [Aside: Personally, I might use "mode" rather than "style", but I assume we are only addressing DRs and significant shortcomings in the document.]

> Murata-san introduced the terms NFC and NFD for the two normalization approaches applied commonly for UNICODE. We should either spell out in full and prefix with "Unicode" (my preference) or add to Terms and Definitions.

Agreed.

**2014-08-14 Caroline Arms:**

Understanding clause 8 has been a real struggle. Below, and attached as a text file, are notes that I have accumulated over several readings, re-readings, referring to cited RFCs, and trying to see what has been lost from the 3rd edition and whether the deletion was deliberate. What I am hoping is that my struggle (as reflected in the notes) can be a useful guide as to how to make the document more comprehensible. It would be useful to get into the draft (perhaps as drafting notes) a list of the objectives for the revised standard, distinguishing between changes to the technical specification and changes to the presentation of the specification.

The notes are not necessarily complete, but I have no more time to devote to this task today (and may be not tomorrow either). I wanted to get something out in plenty of time for people to read it before the call next

Thursday and possibly provide answers to questions I raise here about things I don't understand in time for me to take another look before the call.

To be continued.... Caroline

===

I started to read clause 8 and found the section very hard to understand in a holistic way, either by reading in order or through the headings in the table of contents. The clause reads like chunks of un-related text. The forest is obscured by the trees (or weeds).

I believe we need to:

A. do a much better job of explaining things at a higher level (not only "what" but "why") and providing "sign-posts" in the narrative that guide the reader; B. consider whether a different organization into clauses and sub-clauses is warranted, particularly related to where the Pack URI stuff goes.

B really needs to be done before A. I don't have a particular recommendation to make on re-organization (at least not at this stage) but believe there should be a discussion on this before long. Here are some questions/comments I have:

Clauses 8.2.4 (Growth Hint), and 8.2.5 (XML Usage) seem completely overwhelmed by the detailed content of 8.2.2 (Part Names) and 8.3 (Pack URI scheme) which appear to be (and indeed are) related. The flow between 8.2.5 and 8.3 is particularly jolting.

Is the Pack URI scheme really part of the Package Model (the top heading for clause 8)? I believe the objective of the pack: scheme is to permit references to parts (and part fragments) from OUTSIDE the package. To me, it is important to understand the key internal structures first, including Relationships. For one thing, I see relationships parts as soon as I look inside a package. I don't see Pack URIs, only path segments (or isegments).

Treating Pack URI in an annex did have the advantage of being able to ignore it until you had grasped the essential elements of the package model. And the headings for clauses and subclauses in clause 8 of the previous edition did a somewhat better job of suggesting a logical structure that helped in understanding the model as a whole and knowing where to find what. However, the function of Pack URI was obscured because its function was inadequately introduced in the main text. I realize that I don't have a good sense of when it is used in practice.

Did you have a particular reason for locating Pack URI as 8.3? Might it fit better after 8.6 Relationships or in a new top level clause between 8 (Package Model) and 9 (Physical Package).

I'm not certain whether 8.4 Base IRIs and 8.5 Resolving Relative References (non-normative) apply only to the use of pack: or also in other contexts. That would determine where those clauses should go. I do find myself wondering whether we should revive Part Addressing as a clause or subclause heading and put Pack URI scheme, Base IRIs, and Resolving Relative References as subsidiary clauses.

===== now for some notes on some particular subclauses=====

#### 8.2.2 Part Names

\*\*\*Related terms and definitions

Text in 8.2.2.1 doesn't seem compatible with the trimmed definition in Terms and Definitions. 8.2.2.1 is better.

old version in Terms and Defs:

4.27

part name

The path component of a pack URI. Part names are used to refer to a part in the context of a package, typically as part of a URI.

Rex trimmed to path component of a pack URI

Suggested replacement: Unicode string that identifies a part in the context of a package

I think the fact that a path component of a pack URI is a part name belongs in the text not a definition.

The trimmed definition of Pack URI (URI scheme that allows URIs to be used as a uniform mechanism for addressing parts within a package) also needs work. As we use the term, a "Pack URI" is not a scheme, but an instance compliant with the scheme syntax and rules. Do we need a clause 3 entry for Pack IRI too?

\*\*\* References to RFC 3986 and RFC 3987 -- and URI/IRI confusion \*\*\* This is really several issues.

1. I think we need a sentence or two to introduce URI and IRI, particularly given that you have added a section defining a pack\_IRI as distinct from a pack\_URI. I need to fully understand the intent of the changes wrt URI and



IRI in order to come up with suggested wording and look for inconsistencies. I believe that one important modification you are looking for to Part 3 is to let part names contain non-ASCII characters without percent-encoding. Is that the only change to the model?

2. I think we need to introduce RFC 3986 and RFC 3987 briefly in the text and relate them to URI and IRI. This could be combined with addressing point 1, possibly near beginning of 8.2.2. Also applies to 8.3 (Pack URI scheme), 8.4 (Base IRIs/URIs), and 8.5 (Resolving relative references). I believe another of your objectives was to avoid duplication of content from the RFCs but to rely on referring to them. Am I right? Would it work to use an Annex to introduce URIs, IRIs, RFC 3986, RFC 3987 and how OPC uses them (and varies from them)? Should such an Annex be normative or informative?

3. There are various references to using terms as defined in the two RFCs. I think we need a consistent approach to listing these terms. This could be in the text near each ABNF chunk, making sure that all relevant terms are covered. Alternatively, it could be by putting all the terms into Terms and Definitions, with associated notes that indicate which RFC they are defined in.

4. In 8.3.2 we should probably say explicitly that the syntax of the Pack URI scheme can be extended to IRIs. I haven't chased down whether a pack\_IRI as you have added it in 8.3.2 is still a valid pack URI according to the provisional scheme registered with IANA or not. But clarifying text might be different depending on the answer.

5. May need to change "URI" to "URI or IRI" or to "IRI" in more places. I'm very confused on this.

6. 8.4 is entitled Base IRIs but mostly discusses URIs and 3986.

\*\*\* Base URIs/IRIs

8.4

For "This subclause defines" to be true, I think the first Note needs to be regular text, i.e., normative rather than informative.

Section 5.1 of RFC is NOT a four step model, but four ways (cases) of determining a Base URI in order of precedence. So I believe both notes to be incorrect, the first in detail, the second in substance.

First Note: Use "case" (or some equivalent word) rather than "step."

Second Note: Case 3 and 4 will only come into play if case 2 (being described for OPC in this subclause) is not applicable. The relationship between what is "defined" here and the other cases in 3986 needs to be clarified.

\*\*\*Resolving relative references \*\*\*

## 8.5

One underlying issue is now clear to me: there may be references that are intended to function like "relative references" as defined in RFC 3986, but are not technically valid as defined there or in 3987. For example, they may have backward slashes that should be interpreted as forward slashes. I will be happy to try and produce some clearer wording when I understand more fully. Does the resolution process as described in 8.5 come into play only within a pack URI? All the examples are with pack URIs. However, in the published Part 2, annex A is not tied to Pack URIs, but relates to URIs in general.

### 2014-08-14 Murata-san:

> I started to read clause 8 and found the section very hard to understand in a holistic way, either by reading in order or through the headings in the table of contents. The clause reads like chunks of un-related text. The forest is obscured by the trees (or weeds).

> I believe we need to:

> A. do a much better job of explaining things at a higher level (not only "what" but "why") and providing "sign-posts" in the narrative that guide the reader;

I hope that SC34-WG4\_N0279 (esp. "My understanding of Design Goals of the pack URI scheme" in page 7) can provide a basis of such an overview. At least, it provides an answer to one of your comments (see below).

> B. consider whether a different organization into clauses and sub-clauses is warranted, particularly related to where the Pack URI stuff goes.

> B really needs to be done before A. I don't have a particular recommendation to make on re-organization (at least not at this stage) but believe there should be a discussion on this before long. Here are some questions/comments I have:

> Clauses 8.2.4 (Growth Hint), and 8.2.5 (XML Usage) seem completely overwhelmed by the detailed content of 8.2.2 (Part Names) and 8.3 (Pack URI scheme) which appear to be (and indeed are) related. The flow between 8.2.5 and 8.3 is particularly jolting.

This is true. From the beginning, 8.2.4 and 8.2.5 are not related with the rest of 8.2.

> Is the Pack URI scheme really part of the Package Model (the top heading for clause 8)? I believe the objective of the pack: scheme is to permit references to parts (and part fragments) from OUTSIDE the package. To me, it is important to understand the key internal structures first, including Relationships. For one thing, I see relationships parts as soon as I look inside a package. I don't see Pack URIs, only path segments (or isegments).

SC34-WG4\_N0279 (esp. "My understanding of Design Goals of the pack URI scheme" in page 7) provides four goals.

One of them is:

C) The same mechanism should be used for referencing from inside and outside.

I strongly believe that Pack URI schemes are for both for INSIDE and OUTSIDE references. This is the key reason why I put pack URIs in 8.3.

To me, part names, pack URIs, base URIs, and relative references are first class citizens. They can possibly be extended for other package formats. Relationships look second-class citizens.

They are specific to OPC , and is intended to make the migration from 8.3 file names easier.

INSIDE references to parts are relative references.

Resolution of relative references is not understandable unless you understand what is a base URI. Pack URIs provide base URIs (8.4 now says so).

> Treating Pack URI in an annex did have the advantage of being able to ignore it until you had grasped the essential elements of the package model. And the headings for clauses and subclauses in clause 8 of the previous edition did a somewhat better job of suggesting a logical structure that helped in understanding the model as a whole and knowing where to find what. However, the function of Pack URI was obscured because its function was inadequately introduced in the main text. I realize that I don't have a good sense of when it is used in practice.

I think that treating Pack URIs in an annex significantly hampers understanding of the package model. Without understanding pack URIs, we cannot understand the OPC package model.

> Did you have a particular reason for locating Pack URI as 8.3? Might it fit better after 8.6 Relationships or in a new top level clause between 8 (Package Model) and 9 (Physical Package).

See above.

> I'm not certain whether 8.4 Base IRIs and 8.5 Resolving Relative References (non-normative) apply only to the use of pack: or also in other contexts. That would determine where those clauses should go. I do find myself wondering whether we should revive Part Addressing as a clause or subclause heading and put Pack URI scheme, Base IRIs, and Resolving Relative References as subsidiary clauses.

Base URIs and resolution of relative references are introduced in RFCs 3986 and 3987. 8.4 and 8.5 explain how generic rules in the two RFCs are adjusted for OPC.

> ===== now for some notes on some particular subclauses=====

> 8.2.2 Part Names

> \*\*\*Related terms and definitions

> Text in 8.2.2.1 doesn't seem compatible with the trimmed definition in

> Terms and Definitions. 8.2.2.1 is better.

> old version in Terms and Defs:

> 4.27

> part name

> The path component of a pack URI. Part names are used to refer to a

> part in the context of a package, typically as part of a URI.

> Rex trimmed to path component of a pack URI

> Suggested replacement: Unicode string that identifies a part in the context of a package

Your suggestion looks nice to me.

> I think the fact that a path component of a pack URI is a part name belongs in the text not a definition.

> The trimmed definition of Pack URI (URI scheme that allows URIs to be used as a uniform mechanism for addressing parts within a package) also needs work. As we use the term, a "Pack URI" is not a scheme, but an instance compliant with the scheme syntax and rules. Do we need a clause 3 entry for Pack IRI too?

We should replace "pack URI" by "pack URI scheme" in Clause 2.

"Pack URI" is a casual term for "URIs of the pack URI scheme".

> \*\*\* References to RFC 3986 and RFC 3987 -- and URI/IRI confusion \*\*\*

> This is really several issues.

> 1. I think we need a sentence or two to introduce URI and IRI, particularly given that you have added a section defining a pack\_IRI as distinct from a pack\_URI. I need to fully understand the intent of the changes wrt URI and IRI in order to come up with suggested wording and look for inconsistencies. I believe that one important modification you are looking for to Part 3 is to let part names contain non-ASCII characters without percent-encoding. Is that the only change to the model?

Historically, OPC part names were restricted to ASCII names.

This means that non-ASCII characters have to be percent encoded.

However, Annex A appear to allow verbatim non-ASCII as part of relative references.

Because of the Japanese comment, verbatim non-ASCII was allowed as part of OPC part names. (EPUB 3 does so too.) But what is published is only halfly cooked.

Ideally, we should always use IRIs and avoid URIs. But the way RFCs 3986 and 3987 are written does not allow such a simple approach. We have to define both URIs and IRIs of the pack scheme.

> 2. I think we need to introduce RFC 3986 and RFC 3987 briefly in the text and relate them to URI and IRI.

Agreed.

>This could be combined with addressing point 1, possibly near beginning of 8.2.2. Also applies to 8.3 (Pack URI scheme), 8.4 (Base IRIs/URIs), and 8.5 (Resolving relative references). I believe another of your objectives was to avoid duplication of content from the RFCs but to rely on referring to them. Am I right?

Yes, you are.

>Would it work to use an Annex to introduce URIs, IRIs, RFC 3986, RFC 3987 and how OPC uses them (and varies from them)? Should such an Annex be normative or informative?

I do not think an Annex would work. These terms are so crucial for the package model.

> 3. There are various references to using terms as defined in the two RFCs. I think we need a consistent approach to listing these terms. This could be in the text near each ABNF chunk, making sure that all relevant terms are covered. Alternatively, it could be by putting all the terms into Terms and Definitions, with associated notes that indicate which RFC they are defined in.

I think that putting all such terms into Terms and Defs is a good idea.

> 4. In 8.3.2 we should probably say explicitly that the syntax of the Pack URI scheme can be extended to IRIs. I haven't chased down whether a pack\_IRI as you have added it in 8.3.2 is still a valid pack URI according to the provisional scheme registered with IANA or not. But clarifying text might be different depending on the answer.

The two sets of BNFs in 8.3.2 differ in two main points:

iunreserved and ipchar rather than unreserved and pchar.

Adding a note about this would help.

> 5. May need to change "URI" to "URI or IRI" or to "IRI" in more places. I'm very confused on this.

This is confusing, but unavoidable as long as we use RFCs 3986 and 3987.

> 6. 8.4 is entitled Base IRIs but mostly discusses URIs and 3986.

True. This is because RFC 3987 simply relies on RFC 3986 for the resolution of relative references. See

<http://tools.ietf.org/html/rfc3987#section-6.5>

> \*\*\* Base URIs/IRIs

> 8.4

> For "This subclause defines" to be true, I think the first Note needs to be regular text, i.e., normative rather than informative.

The case statement is the core of 8.4, and is normative.

We have to make this clearer.

> Section 5.1 of RFC is NOT a four step model, but four ways (cases) of determining a Base URI in order of precedence. So I believe both notes to be incorrect, the first in detail, the second in substance.

I do not see any differences between

(1) steps and (2) cases having precedence.

Should we say "ways", since it is the term used in RFC 3986?

> First Note: Use "case" (or some equivalent word) rather than "step."

I do not buy this.

> Second Note: Case 3 and 4 will only come into play if case 2 (being described for OPC in this subclause) is not applicable. The relationship between what is "defined" here and the other cases in 3986 needs to be clarified.

This is what I tried in the two notes in 8.4. Change proposals are very welcome.

> \*\*\*\*Resolving relative references \*\*\*

> 8.5

> One underlying issue is now clear to me: there may be references that are intended to function like "relative references" as defined in RFC 3986, but are not technically valid as defined there or in 3987. For example, they may have backward slashes that should be interpreted as forward slashes. I will be happy to try and produce some clearer wording when I understand more fully. Does the resolution process as described in 8.5 come into play only within a pack URI? All the examples are with pack URIs. However, in the published Part 2, annex A is not tied to Pack URIs, but relates to URIs in general.

This what we extensively discussed in Prague. See "Relative references (A.3)" from John on August 9th, 2014.

**2014-08-17 Murata-san:**

> I hope that SC34-WG4\_N0279 (esp. "My understanding of Design Goals of the pack URI scheme" in page 7) can provide a basis of such an overview. At least, it provides an answer to one of your comments (see below).

Here are a few paragraphs created using N0279 as a basis.

Each part has a unique name. Within a package, a part name uniquely identifies a part.

This part of ISO/IEC 29500 provides the pack URI/IRI scheme for this construction. Given a part name and a URI/IRI of a package, it is possible to construct a pack URI/IRI that identifies the part. Such an URI/IRI allows OUTSIDE references to parts.

Note: A fragment of a part can be referenced by appending a fragment identifier to a pack URI/IRI for the part.

To resolve relative references within a package, a base URI/IRI is needed. This part of ISO/IEC 29500 specifies how a pack URI/IRI is constructed as a base URI/IRI.

Resolution of relative references against such a base URI/IRI provides another pack URI/IRI, which identifies another part. Thus, relative references within a package allow INSIDE references to parts.

Hopefully, these paragraphs, current "Overview", and the text I deleted from "Scope" provides a good basis for new "Overview". How do people feel?

We discussed many of the points raised in these mails. Murata will factor in this feedback in his next draft.

Long-Term Digital Signature

Murata-san asked delegates to be prepared to discuss this in Kyoto and possibly make a decision at that meeting.

Metadata

**2014-08-17 Murata-san wrote:**

Metadata extensions for OPC have been occasionally discussed in this WG. (Links to e-mails are shown below.) Should we do some extensions as part of the ongoing revision?



Core properties of OPC are very restricted. In the past, nobody has proposed significant extensions to them except my own proposal for adding phonetics. (I'm not hoping to do that anymore.)

As far as I know, users do not extensively specify core property metadata. I also know that EPUB metadata has not been heavily used.

EPUB allows any foreign metadata object to be embedded within an EPUB publication. I do not know if this feature is heavily used, but it certainly provides a nice extension point.

Do we have to do something for OPC? OPC already allows any metadata to be embedded as a part.

Moreover, we can specify a relationship type from the package to such embedded data. For example, we can specify <http://ns.editeur.org/onix/3.0/short> as a relationship type (ONIX 3). I think that OPC is already as powerful as EPUB in allowing foreign metadata to be embedded. Do we have to do something?

There was a short discussion.

"Content type" and "media type"

**2014-07-17 Caroline Arms:**

Here is the message, I promised to send to the list on today's call.

After the earlier call where John reported on his action item asking the folks who designed and documented OPC about their use of content type and media type, I was motivated to look at RFC 2616 and Part 2 in a bit more detail, in part to try and understand Murata-san's underlying concern.

The explanation John reported is consistent with how I have interpreted the phrase in Part 2 (OPC). John indicated that the OPC designers saw content type as an abstract concept. To me that makes sense, since there are lots of typologies for "content" used in different contexts. In OPC, "media-type" as defined in the RFC is the particular syntactic representation used to identify types of content. I see a value in having two terms, one for the concept and the other for the particular syntax and instances thereof. I do not believe that Part 2 gets the usage of the terms quite straight and do think that the definition of "content type" needs adjustment, particularly after the WDO adjustment in line with ISO guidelines. [Aside: I don't think RFC 2616 gets its term usage quite consistent either.]

Comparing OPC and RFC 2616:

Each has a data structure with a name based on the words "content" and "type." Both structures incorporate values that follow the media-type syntax.

OPC has the Content Types stream (defined in 9.2.3.3 in Rex's base WD0 doc) which makes use of a ContentType attribute which takes values that follow the media-type syntax.

RFC 2616 defines a Content-Type entity-header field by:

Content-Type = "Content-Type" ":" media-type

=== My personal views ===

1. We need the two terms but need to review Part 2 carefully for appropriate usage of the terms. For example, "A content type as defined in RFC 2616." in annotations in the tables in 9.2.3.3.3 and 9.2.3.3.4 seems wrong, since we do not use the Content-Type structure defined in RFC 2616 (see above).
2. Personally, I have always found the use of "media" in media-type a perplexing use of the word. I would never use the word "media" to replace "content." They are not synonyms.
3. We should avoid major changes in terminology because several formats based on OPC have documentation that makes substantial use of "content type" as a phrase. Indeed, ECMA 388 (XPS spec.) does not appear to use "media-type" or "media type," only "content type."

Other examples from documentation for OPC-based formats:

"The part's content type property uses a MIME-style media type to describe part content." "Content Type Component: XML markup (stored in a ZIP item) that identifies the MIME Media Type of each Part in the Package."

4. The fact that OPC allows "parameters" in the media-type syntax to be disallowed (and XPS does disallow them) seems like another reason to keep the terms distinct.

= = =

If others agree in general, I am prepared (at some convenient future point) to do a careful read of Part 2 with this particular issue in mind and propose changes.

**2014-08-08 John Haug:**

I think Caroline raises well-considered points and I tend to agree with her assessment.

I don't think we discussed this topic on the previous call, so I may not yet have shared this info. Previously, someone asked why OPC's content type allows an empty string (RFC 2616 does not allow this for media-type). Some package parts are files with known types, others may be serialized bits with no or no known types. The designer told me it was a tradeoff decision to avoid having to guess or fake a type.

**2014-08-08 Murata-san:**

Agreed [with John]. Let's discuss in the next teleconf.

The same thing applies to web servers. They use application/octet-stream rather than the empty string.

**2014-08-08 Murata-san:**

Caroline, Thank you for your analysis. This is very helpful.

RFC 2616 defines HTTP 1.1, but there are two important RFCs. They are:

[RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

[RFC2046] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", RFC 2046, November 1996.

I believe that they are more authoritative about media types.

[snip]

> 4. The fact that OPC allows "parameters" in the media-type syntax to be disallowed (and XPS does disallow > them) seems like another reason to keep the terms distinct.

OPC now allows parameters. An ugly regular expression in opc-contentTypes.xsd does allow parameters. I do not think that parameters are useful but should we drop them from OPC?

## The Status of the Use of “IRI”

**2014-08-21 Caroline Arms:**

I ran into the following just now. Have they been discussed in relation to the Part 2 revision?

1. <https://www.w3.org/International/wiki/IRIStatus>

Particularly:

At some point it was recognized that "Web addresses" and "IRIs" and "URIs" were all trying to accomplish effectively the same thing. With the demise of the IRI working group at IETF, the WHATWG "URL" spec is currently the focus of development effort. Development of this document appears to have better implementer support.

...

The IETF Working Group for IRI was shut down in January 2014 in recognition of the lack of progress or potential progress on a revision of the RFC 3987 document and URL is now the main vehicle for providing a solid reference that specifications (such as HTML) and implementations (such as browsers) can be based on. Similar to the HTML5 effort, URL's editors mainly focus on documenting what browsers actually do and, where interoperability is not currently present, working with implementers on resolving differences.

This means that the issues with IRI are unresolved. The URI/IRI transformation is not fully described and it is unclear if the URL spec will resolve all or only a part of the issues being worked on in the context of IRIs.

2. <http://url.spec.whatwg.org/>

This was news to members. It sounds controversial. In any event, even if this gained traction, it's too far out for us to let it affect our current revision.

## **6. Defect Reports**

The public, online DR log is now at

<https://onedrive.live.com/?cid=c8ba0861dc5e4adc&sc=documents&sa=501765342&id=C8BA0861DC5E4ADC%21105>. Access individual DRs via the hyperlinks contained within the spreadsheet's left-most column.

**DR 10-0048 “OPC: Processing model for handling ZIP encryption”**

There was a brief discussion.

**DR 13-0013 “WML: omissions and inconsistencies in the specification of attributes”**

Chris is making good progress on the various pieces of this. (See recent mails titled ‘DR 13:0013: Proposed solution to "the [yellow|blue] ones"'.) He will discuss this offline with Francis.

**DR 14-0002 “WML: use of if this element is omitted in the specification of attribute”**

There was a brief discussion.

**DR 14-0006 “General: Values that exceed the Specification”**

There was a brief discussion. Chris will revise his proposal to say something like “For values outside this list the behavior is implementation-specific.”

## 7. Other Business

Interim DCORs for Parts 1 and 4 [N 0285 and N 0286]

Rex reported that should we agree to freeze the contents of these at the end of Kyoto meeting, he is ready to produce the final drafts for circulation and subsequent balloting.

Thanking Host

We thanked Microsoft for hosting the teleconference.

## 8. Future meetings

### Face-to-Face Meetings:

- 2014-09-22/26, Kyoto, JP (with other WGs, and Opening/Closing Plenaries)

Mon: 10:00 SC 34 opening plenary; Afternoon: 14:00 WG1, 16:00 WG5

Tuesday: 09:00-17:00 WG4; 17:30 WG 6 teleconference

Wednesday: 09:00-17:00 WG4

Thursday: 09:00-17:00 WG4

Friday: 10:00 SC 34 closing plenary

- Tentatively 2015-02-23/27 or 2015-03-02/06, location to be determined (with other WGs)
- 2015-06-15/17, BSI, London, UK (possibly with other WGs)
- 2015-09-21/25, Beijing, CN (with other WGs, and Opening/Closing Plenaries)

**Teleconferences:**

- None scheduled

## **9. Adjournment**

Adjourned by unanimous consent at 15:00.