

# IRI BOF -- Presentation

IETF 76, Hiroshima Japan

11/9/2009

Larry Masinter

# History of prior IRI specifications

(10 min)

- 1993: discussions of “internationalized URLs”
- 1997-2001: draft-masinter-url-i18n-00 to -08

This memo defines two kinds of URLs: internationalized but hex-encoded URLs (compatible with [RFC-URL-SYNTAX](#)), and 8-bit URLs. For traditional URLs, characters should be translated into UTF8 and then any octet not allowed in RFC-URL-SYNTAX should be hex-encoded. (with Dürst)

- 2002-2005: draft-duerst-iri-00 to -11  
2005: RFC 3987
- 2007-2009: draft-duerst-iri-bis-00 to -06
- Topics: translation IRI -> URI, URI -> IRI, characters to avoid, comparison, BIDI

# Review of current documents

(15 min)

- draft-duerst-iri-bis-07  
From -06: IRI always processed, not translated
- draft-duerst-mailto-bis-07  
Update in progress
- RFC 4395 (“Guidelines ... new schemes”)

## Major issues

- IRI as protocol element vs IRI -> URI
- Normative reference to IDNA
- Different levels of “liberal processing”

# Other documents & committees

(15 min)

- **HTML5: WhatWG and W3C HTML WG**

Depends on this work

Affects browser and OS implementations

Want “liberal processing rules” in IRI specs

- **IETF IDNABIS WG**

Consistency with IRI on host names

- **IETF EAI WG**

Consistency with mailto: scheme definition

## **And please try not to break:**

HTTPBIS (http URI scheme), XML Core (LEIRI),  
ICANN (I18N TLDs), URI itself, Unicode

# Charter review

(20 min)

- **Scope: focus on IRI, mailto:**  
No other URI “fixes” or “tweaks”  
Update BCP for I18N but no other requirements
  - **Goals**  
general principles (needed to be explicit)
  - **Splitting the documents (tactic for speed)**
  - **Explicit Liaisons**  
HTML5, IDNA, EAI
  - **Schedule**  
Aggressive timeline  
splitting documents helps  
finish IRI update before EAI
  - **Other documents for review**
- To hum or run, that is the question...**