
WOULD YOU
LIKE FRIES
WITH THAT?

html5

a security-minded reader's guide

1999-12-24

html 4.01

document format

2000-01-26

xhtml 1.0

2001-05-31 xhtml 1.1 module-based xhtml

2008-05-21 xhtml 2.0 working draft

meanwhile...

IE 8

FF 3

Opera 9.50

Safari 3.1

Safari/iPhone

Nokia S60

Opera Mini 4

HTML 5

A vocabulary and associated APIs for HTML and XHTML

W3C Working Draft 16 May 2008

Latest Published Version:
http://www.w3.org/TR/html5/
Latest Editor's Draft:
http://www.w3.org/html/wg/html5/
Editors:
Ian Hickson, Google, Inc.
David Hyatt, Apple, Inc.

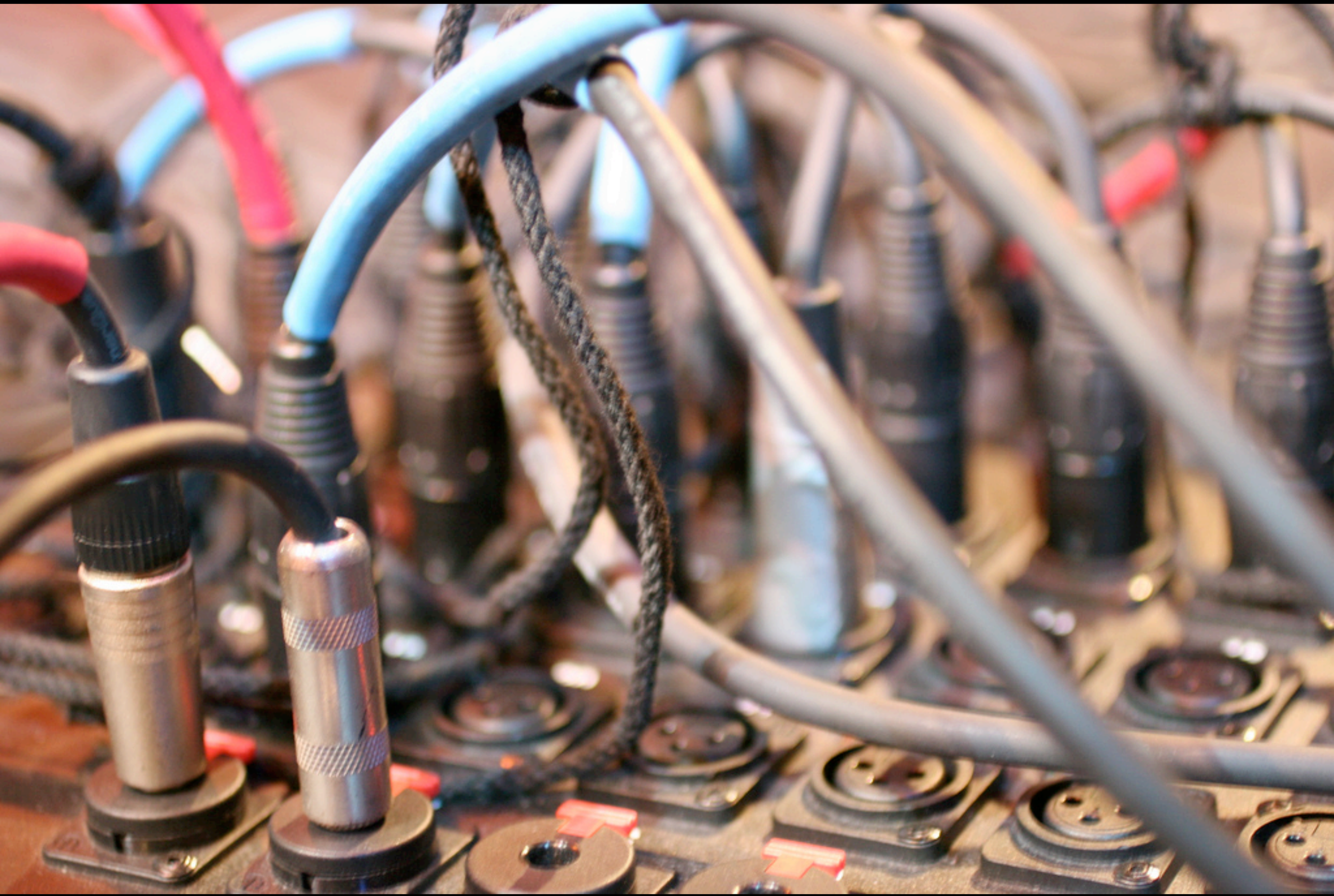
Abstract

This specification defines the 5th major revision of the core language of the World Wide Web: the Hypertext Markup Language (HTML). In this version, new features are introduced to help Web application authors, new elements are introduced based on research into prevailing authoring practices, and special attention has been given to defining clear conformance criteria for user agents in an effort to improve interoperability.

Status of this document

This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the most recently formally published revision of this technical report can be found in the W3C

“a vocabulary and associated apis”



http://www.flickr.com/photos/85182154@N00/296492538/

Casas de Marcos Briones y Miramontes

Fundada

Moraga y Quirós

San Franco de Asis

Muralla

Ruinas

Muralla

Ruinas

Iglesia

Casa del Comte

Officiales

Esta era la fachada principal del Presidio, antes que fuese arruinada la iglesia por los temblores de 1812.

Fundado el 17 de Setc. 1776 por Moraga, Quirós y Fr. Palou.

PRESIDIO de San Francisco 1820

Hasta Bandera

Guardia Principal

Porte de la Compa

Cuartel

Calabozo

Ruinas

Muralla

Muralla.

Ruinas.

Muralla

Ruinas

Muralla

Vista á la bahia, entrada del Puerto, Alcatraz, Angeles, &c. &c.

Camino pa la Yerba Buena

Camino para el Castillo

same-origin policy

things with the same origin can mess with each other

access to members of HTMLDocument by scripts with different origin forbidden

“origin”

scheme

domain name

port

https
example.com
443

whose
authority
caused a script
to exist?

e.g., javascript: URI in stylesheet

e.g., Document generated from data: URI that was redirect target

relaxing origin restrictions through document.domain

x.example.com

=>

example.com

side-effect: set “effective script origin”

```
host: example.com
port: “manual
  override”
```

x.example.com
y.example.com

example.com

Activate Your Web Pages

5th Edition
Includes Ajax and DOM Scripting

JavaScript

The Definitive Guide

O'REILLY®

David Flanagan

4th edition
Chapter 21.3
page 402

“the same-origin policy does not actually apply to all properties of all objects”

exceptions: members of Window

location
postMessage
frames
xxx4

location:
“navigation
policy”

setting

location.href

i.e., navigating elsewhere

introducing browsing contexts

A.com

A.com

B.com

A.com

B.com

C.com

A.com

B.com

A.com

fragment-based messaging

send messages by setting other frames' location

A.com

E.com

B.com

A.com

E.com

E.com

postMessage

implemented in all major browsers

part of Window interface

causes

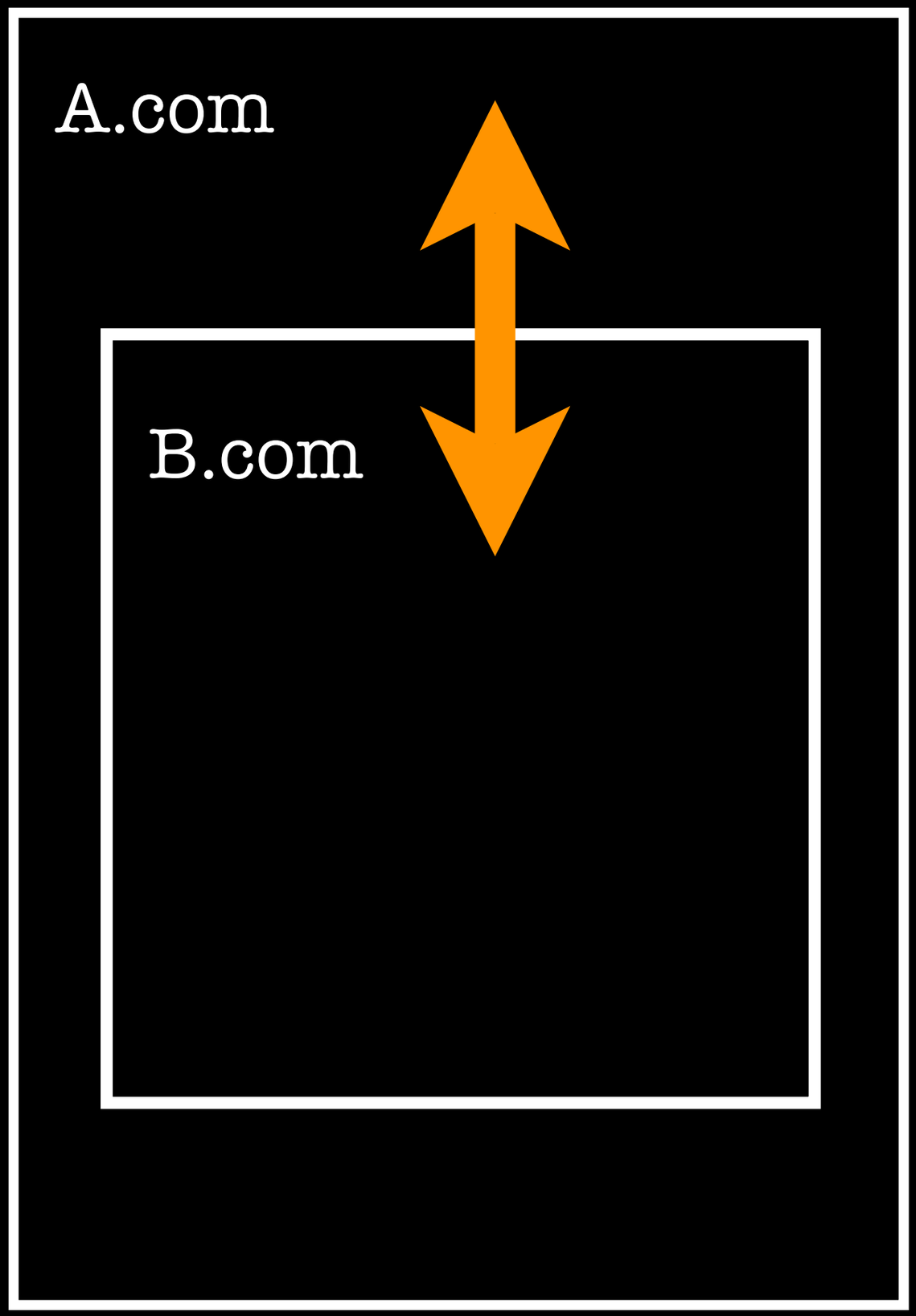
message event

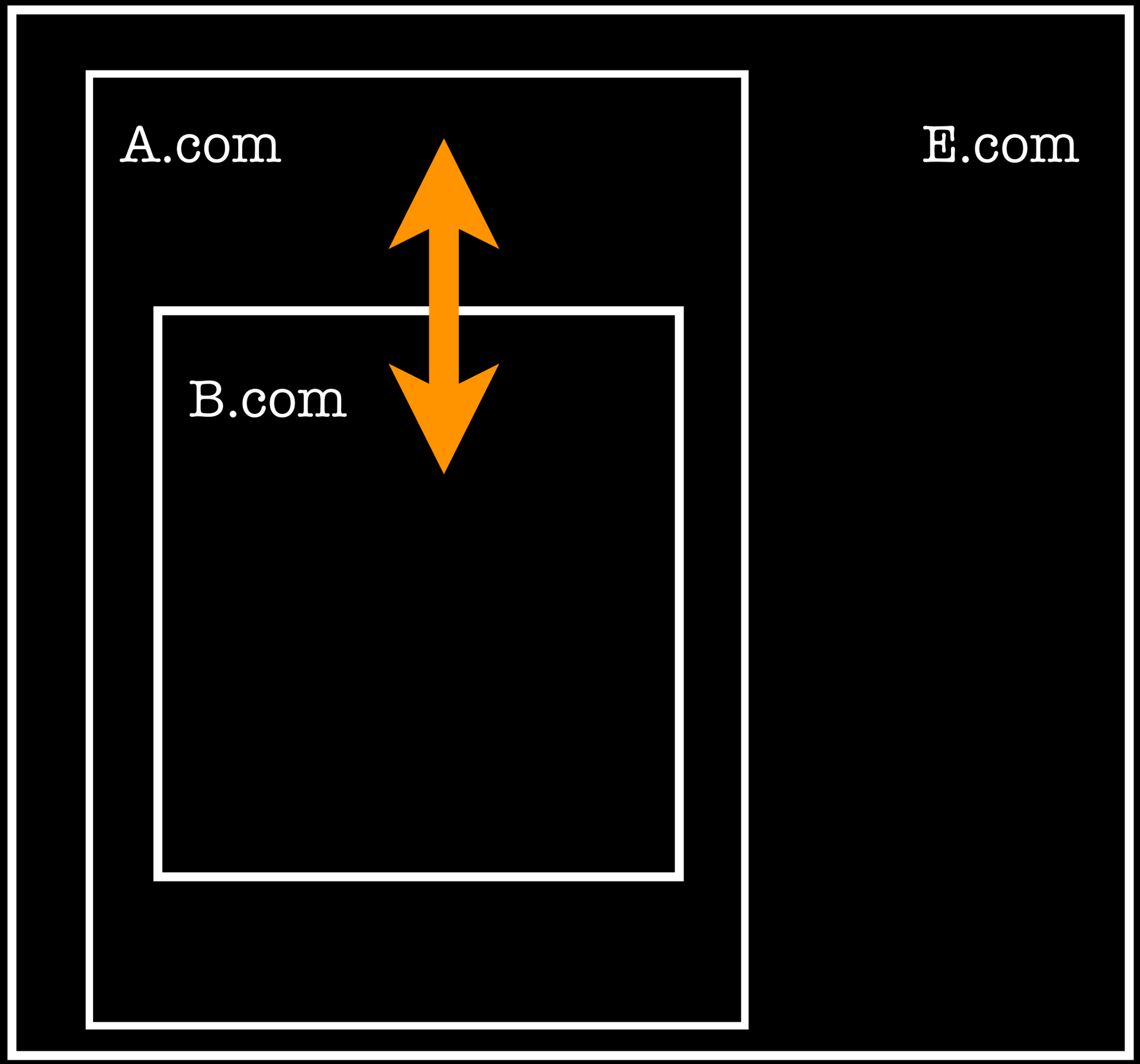
passes a string

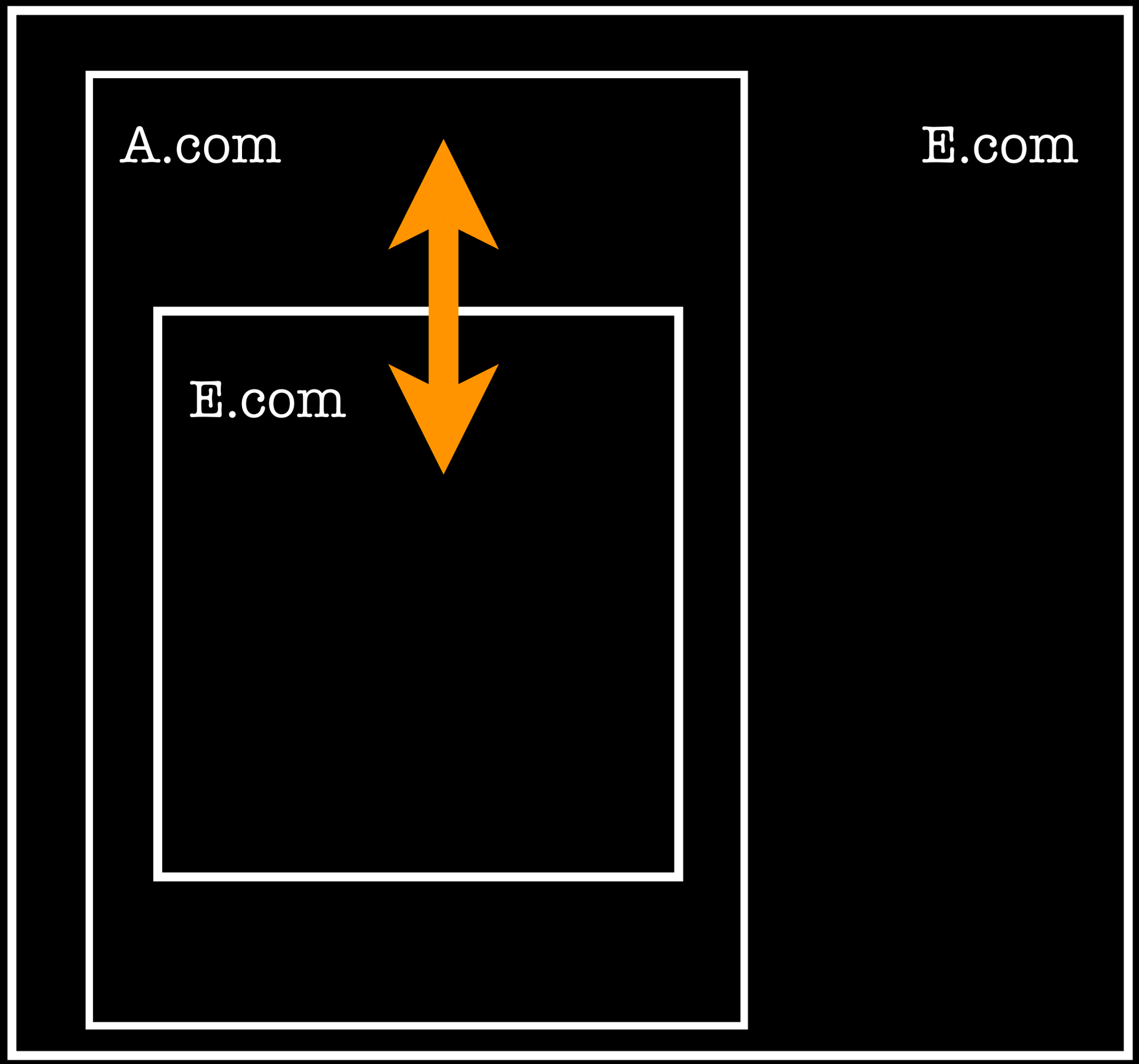
no same-origin restrictions

cross-domain communication

```
void postMessage (
    in DOMString message,
    in DOMString targetOrigin);
```


A.com
B.com


A.com
E.com
B.com


A.com
E.com
E.com

messages are bound to targetOrigin

(Barth, Jackson, Mitchell 2008)

```
document.addEventListener('message',
    receiver, false);
function receiver(e) {
  if (e.origin == 'http://example.com') {
    if (e.data == 'Hello world') {
      e.source.postMessage('Hello',
                           e.origin);
    } else {
      alert(e.data);
    }
  }
}
```

note:

string only!

JSON + eval?

cross-domain communication on the wire

XMLHttpRequest (not quite html5; Last Call)

same-origin constraint

a.com can't read data from b.com with the user's credentials

<idea>

let b.com authorize
(a) a.com reading data
(b) a.com causing unsafe
HTTP requests

“access-control”

+

XMLHttpRequest

Level 2

arbitrary methods

arbitrary content types

ambient authentication

responseXML API

allow from w3.org except people.w3.org

GET

Access-Control HTTP Header

POST, …

pre-flight OPTIONS

Referrer-Root HTTP header on all cross-site requests

separate specification from XMLHttpRequest

almost shipped in Firefox 3

But: ambient authentication and cookies.

Microsoft

XDomainRequest

surprise

separate API:

XDomainRequest()

strings

text/plain

(JSON + eval?)

anonymous

requests

Additional header:

XDomainRequest:1

GET, POST

no other methods

http://flickr.com/photos/mwboeckmann/2313632431/

server-side

dom events

implemented

by Opera

use HTTP to pull an event stream over the network

same-origin

+

access-control

persistent HTTP connection or reconnecting

text/event-stream

the Connection interface

bluetooth?
irda?
broadcast?
tcp? udp?

“On TCP/IP networks, broadcast connections transmit data using UDP over port 18080.”

TCP-based protocol

target host must be subdomain of current origin

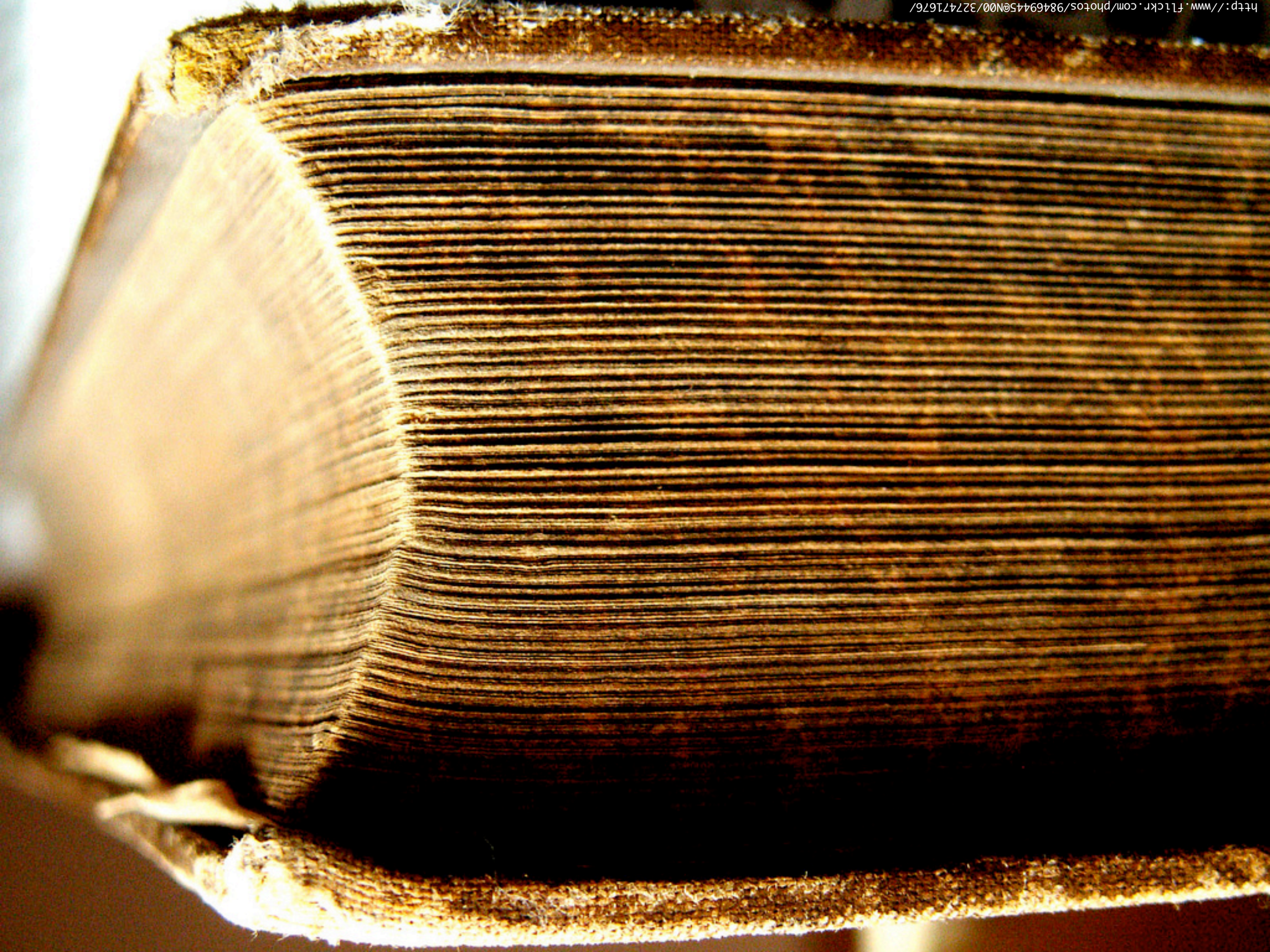
C->S: Hello\n
S->C: Welcome\n
C->S: foo.com\n
S->C: foo.com\n

\0x02
<data>
\0x17

\0x00, \0x17
replaced by
\0xffff

no browser implementations known so far

candidate for removal



sessionStorage

localStorage

Database

localStorage
sessionStorage

shared interface

getItem()

setItem()

clear()

localStorage

scoped by origin

persistent

sessionStorage

scoped by origin
and browsing
context
not persistent

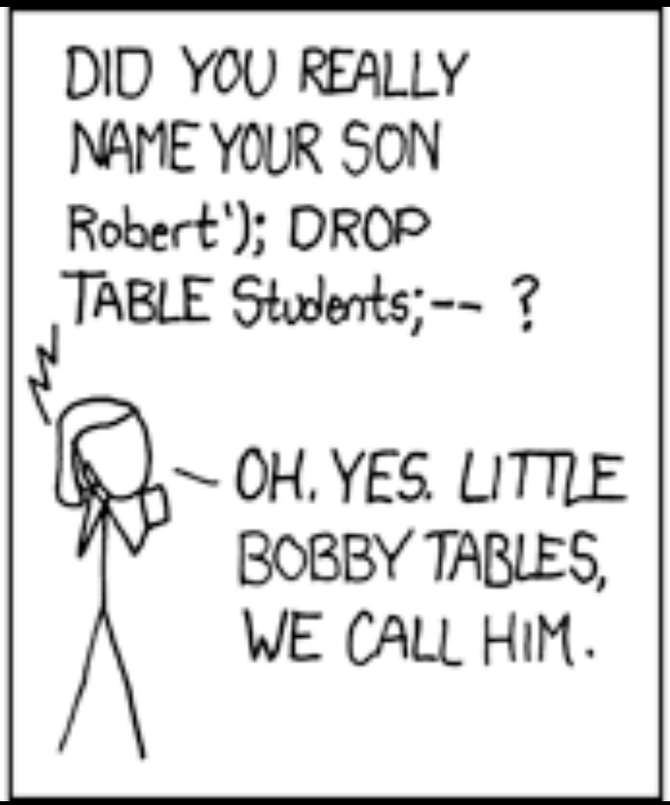
Easy and reliable client-side state.

even more
business logic on
the client?

Database


HI, THIS IS YOUR SON'S SCHOOL. WE'RE HAVING SOME COMPUTER TROUBLE.


OH, DEAR - DID HE BREAK SOMETHING?
IN A WAY-


DID YOU REALLY NAME YOUR SON Robert'); DROP TABLE Students;-- ?
OH. YES. LITTLE BOBBY TABLES, WE CALL HIM.


WELL, WE'VE LOST THIS YEAR'S STUDENT RECORDS. I HOPE YOU'RE HAPPY.
AND I HOPE YOU'VE LEARNED TO SANITIZE YOUR DATABASE INPUTS.

SQL databases
persistent
scoped by origin

asynchronous API

transaction-based

“A future version of this specification will probably define the exact SQL subset required in more detail.”

implemented in

Webkit (Safari)

Offline Web-Apps

more control over the browser cache

cause
download of
additional
resources

cause use of cache to retrieve these in offline mode

detect whether browser is in offline mode

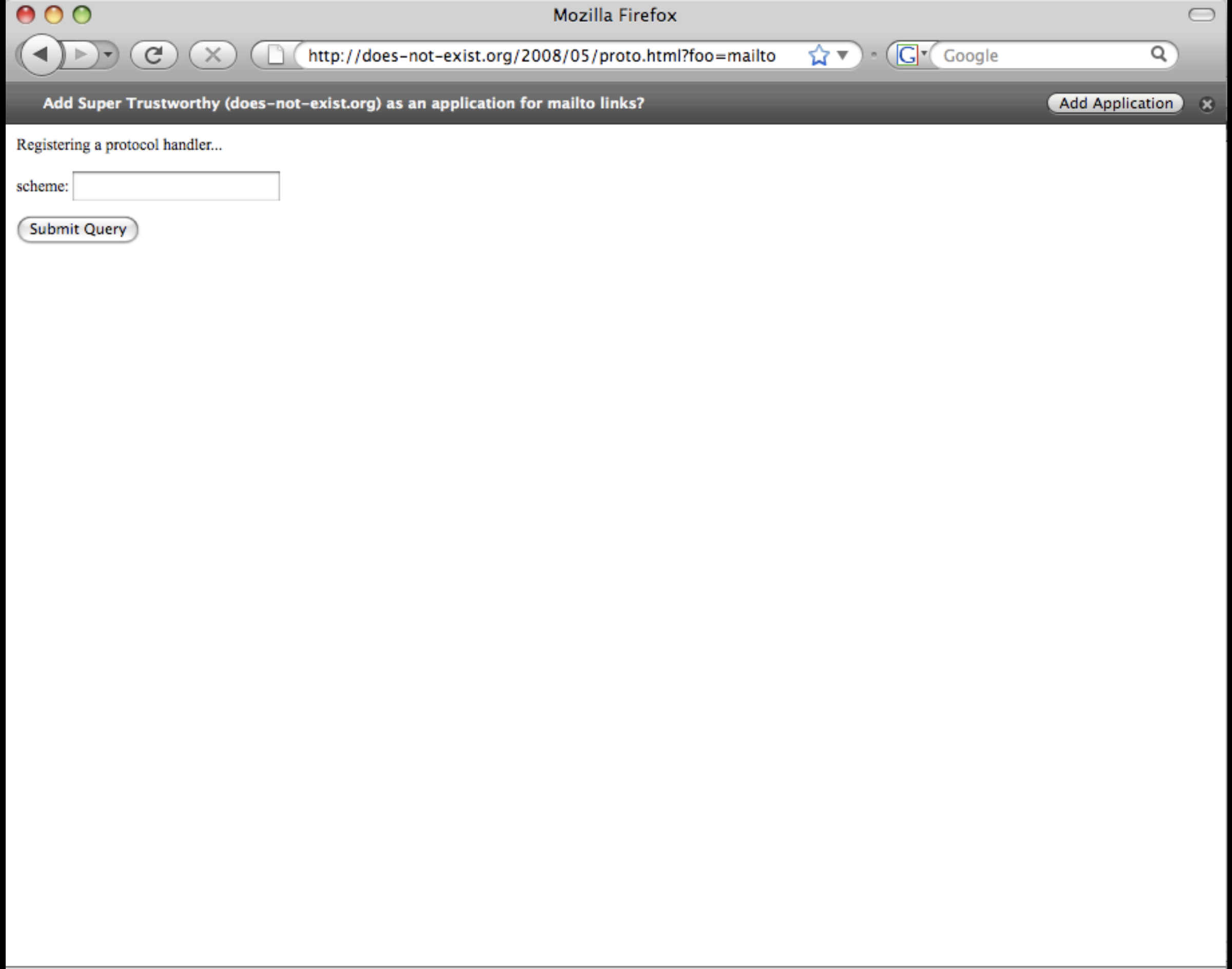
ships in Firefox 3

web-based protocol handlers

ships in Firefox 3

```
navigator.registerProtocolHandler
   ("mailto",
   "https://www.example.com/?uri=%s",
   "Example Mail");
```

http, https
cause security
exceptions


Mozilla Firefox
http://does-not-exist.org/2008/05/proto.html?foo=mailto
Google
Add Super Trustworthy (does-not-exist.org) as an application for mailto links?
Add Application
Registering a protocol handler...
scheme:
Submit Query
Done

Mozilla Firefox

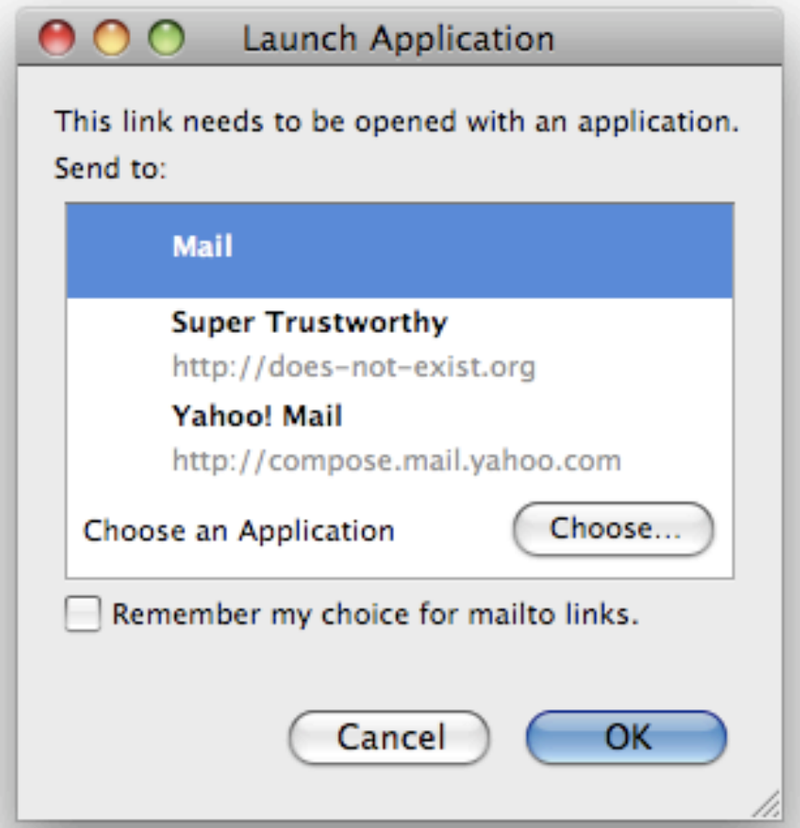
mailto:tlr@w3.org

Google

Registering a protocol handler...

scheme:

Submit Query



Done

registerContentHandler

http://www.flickr.com/photos/lassmatazz/325437141/

W3C HTML WG

http://www.w3.org/html/wg

everybody is
welcome

apply today

tlr@w3.org