Towards the Massive Media Matrix
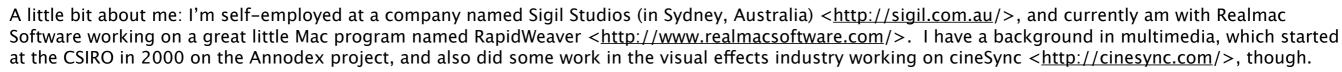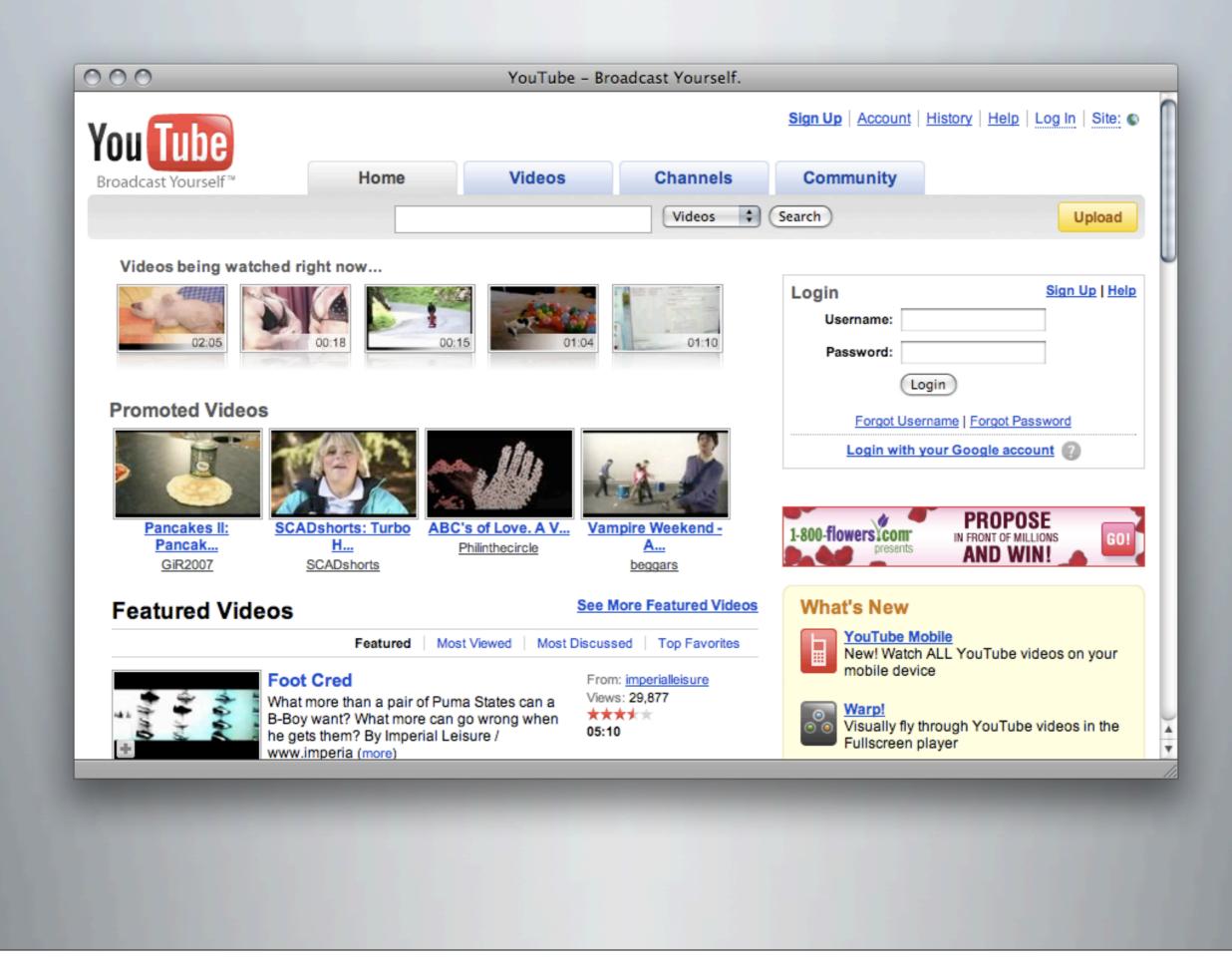
(or turning the www Upside Down)

This is a talk I gave at Kiwi Foo Camp (a.k.a. "Baa Camp"), on February 2, 2008.
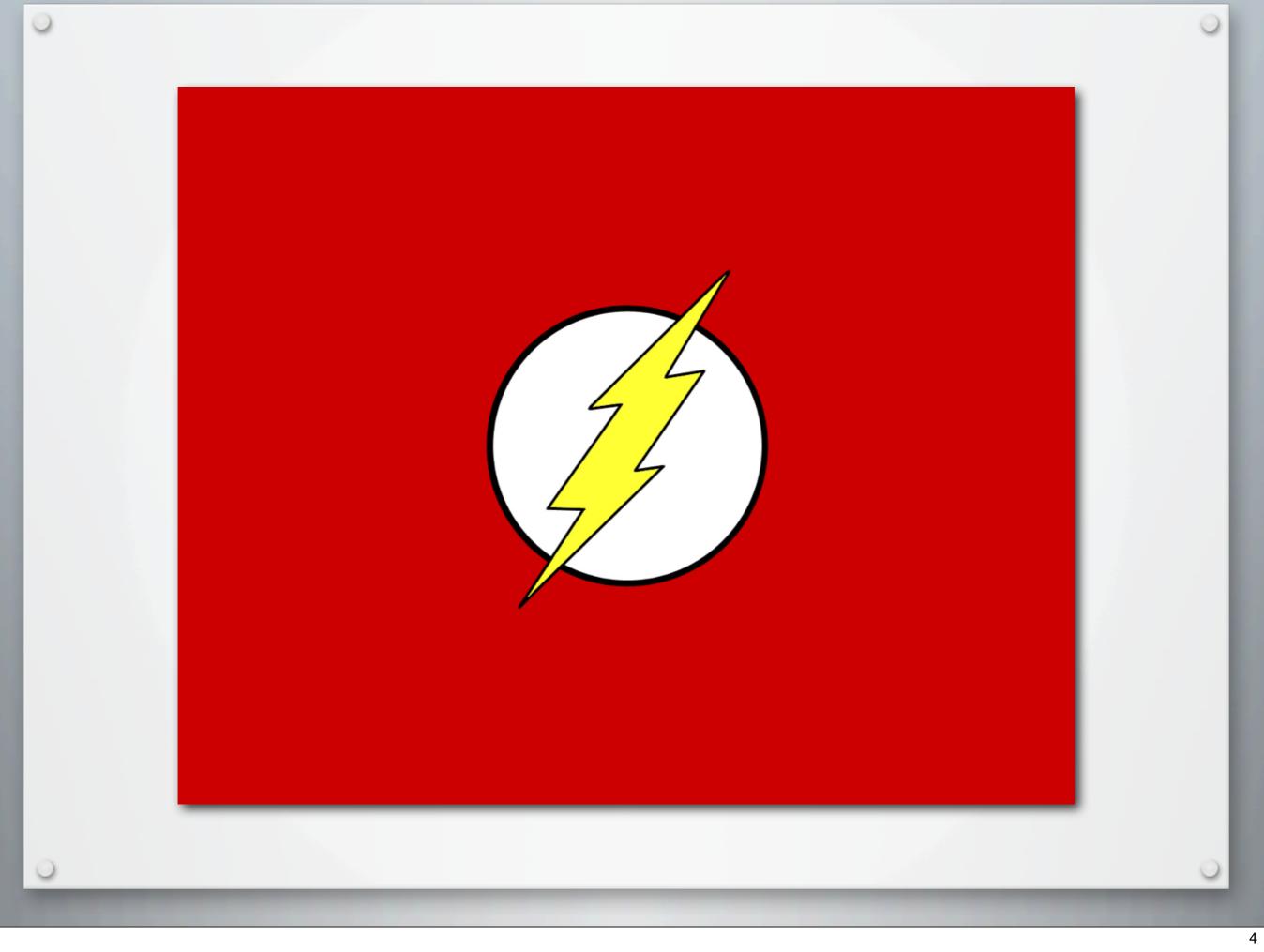
There's two main parts to this talk. The first part is about open standards for Web video. The second is about Annodex, and the general idea of making video a "first-class citizen" on the Web like HTML.

A little bit about me: I'm self-employed at a company named Sigil Studios (in Sydney, Australia) <http://sigil.com.au/>, and currently am with Realmac Software working on a great little Mac program named RapidWeaver <http://www.realmacsoftware.com/>. I have a background in multimedia, which started at the CSIRO in 2000 on the Annodex project, and also did some work in the visual effects industry working on cineSync <http://cinesync.com/>, though.

This is video on the Web as we know it today: YouTube. Mostly, it works…

… but this is the main problem with open-standards video on the Web: Flash.

Or rather, this Flash. Flash is only a problem if you care about open standards on the Web. In other words, if you'd be happy if HTML was proprietary and controlled by a single company, you probably won't care for the rest of this talk. Video on the Web may work, but it is not an open standard right now.

The World Wide Web Consortium (W3C) is an international consortium where Member organizations, a full-time staff, and the public work together to develop Web standards. W3C's mission is:

To lead the World Wide Web to its full potential by developing protocols and guidelines that ensure long-term growth for the Web.

W3C Develops Web Standards and Guidelines

W3C primarily pursues its mission through the creation of Web standards and guidelines. Since 1994, W3C has published more than 110 such standards, called W3C Recommendations. W3C also engages in education and outreach, develops software, and serves as an open forum for discussion about the Web. In order for the Web to reach its full potential, the most fundamental Web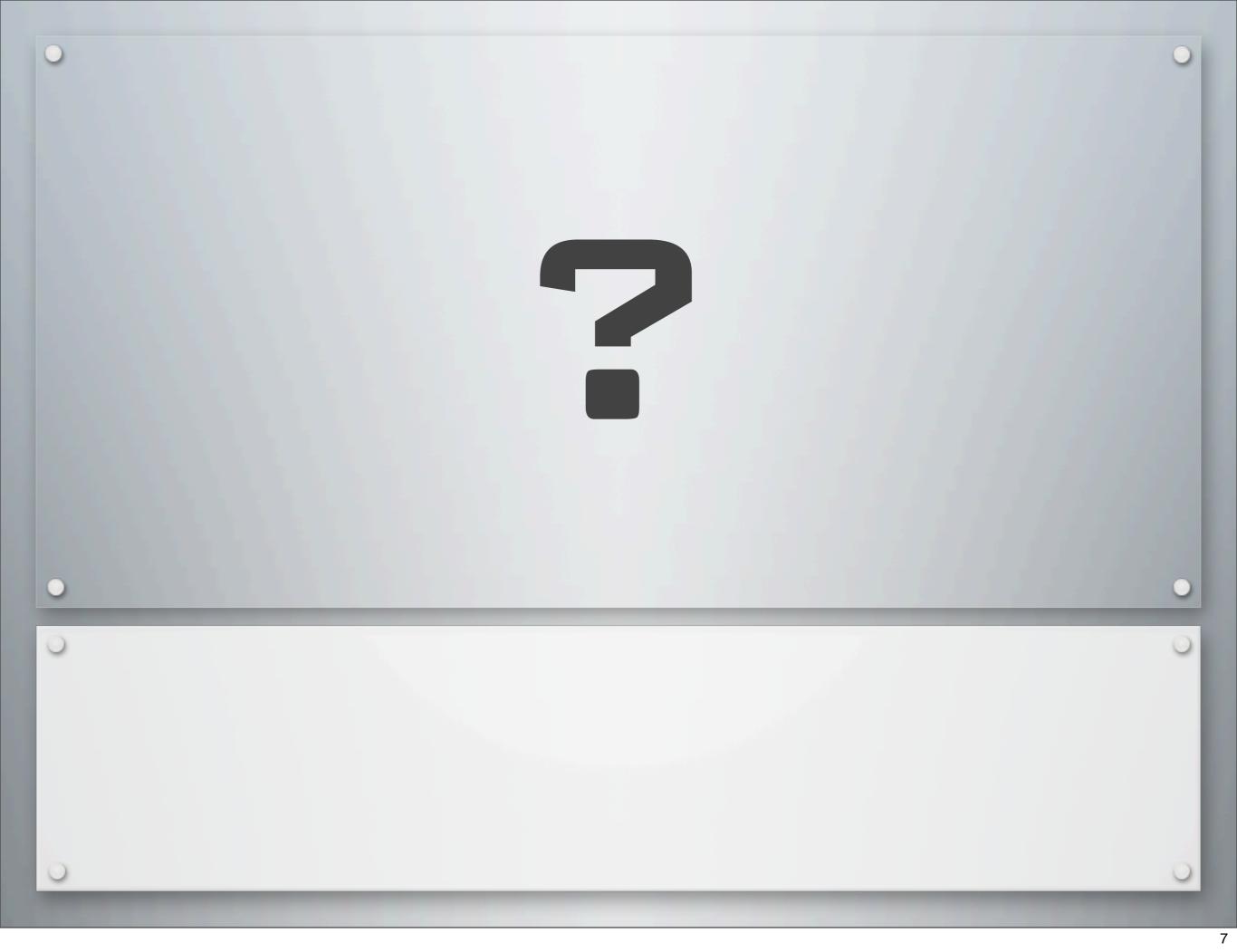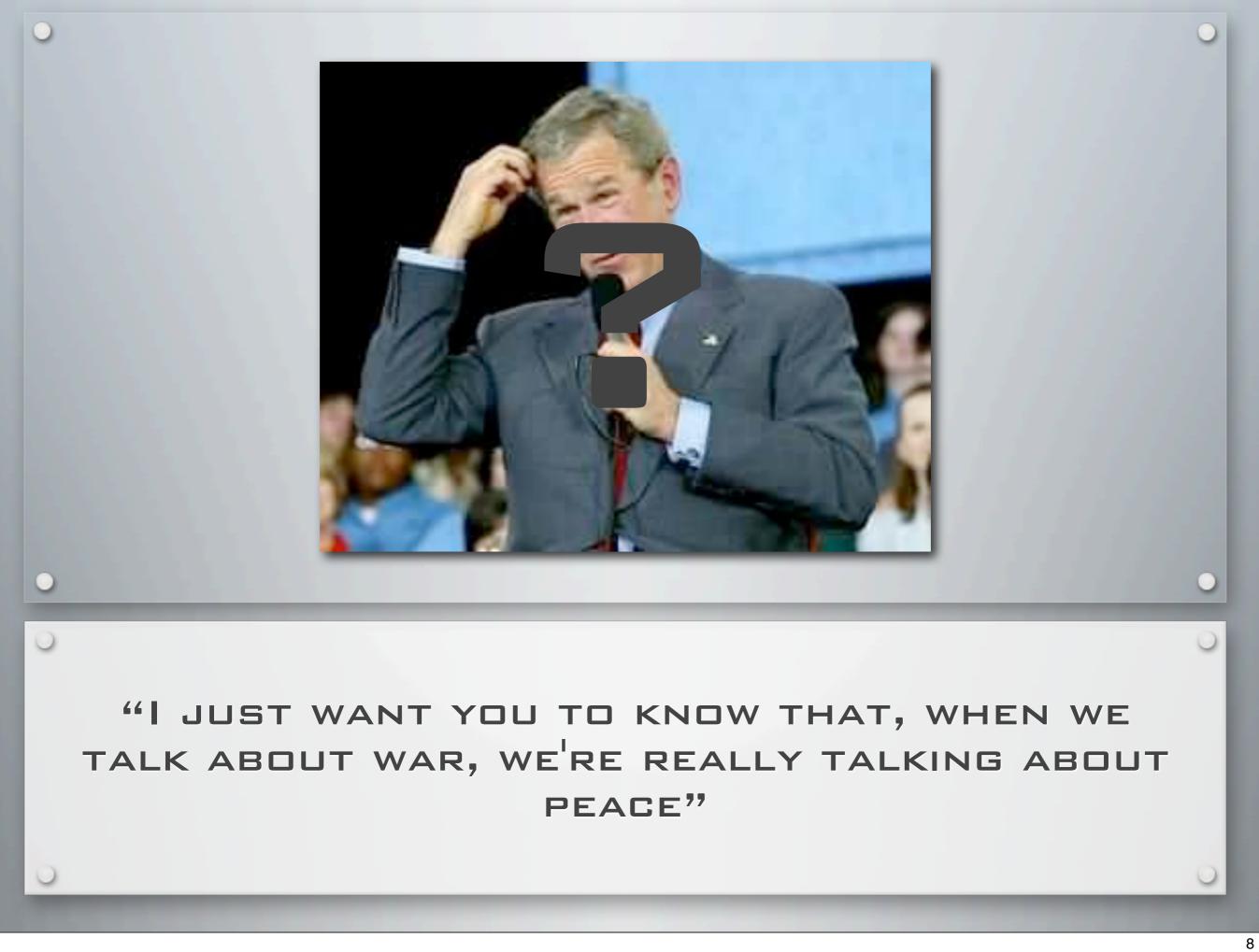 technologies must be compatible with one another and allow any hardware and software used to access the Web to work together. W3C refers to this goal as "Web interoperability." By publishing open (non-proprietary) standards for Web languages and protocols, W3C seeks to avoid market fragmentation and thus Web fragmentation.

Tim Berners-Lee, W3C Director and inventor of the World Wide Web

Open standards are important: arguably, the Web was only successful because it was a completely open standard. The W3C requires its standards to be non-discriminatory, provide accessibility, and provide interoperability across a broad range of applications. (Notice that the word "open" is highlighted there.) Market fragmentation is bad; talk to practically any Web video providers, and I bet you that all of them will start ranting about how much fun it is to provide videos in Flash, be held ransom to Adobe's Flash server licensing, and also provide three or four different video formats (often Flash, QuickTime/MPEG-4, and Windows Media). A very big problem is that being proprietary also precludes open-source software from supporting that standard, which is why most Linux distributions will not support MP3 audio or MP4 video out-of-the-box.

So, what's the challenge in making video on the Web follow open standards?

"I just want you to know that, when we talk about war, we're really talking about peace"

First challenge: this is a hard problem, and there are lots of uninformed people expressing opinions when they don't see the bigger picture.

- [whatwg] Removal of Ogg is *preposterous* Manuel Amador (Rudd-O)
  - [whatwg] Removal of Ogg is *preposterous* Dave Singer
    - [whatwg] Removal of Ogg is *preposterous* Manuel Amador (Rudd-O)
      - [whatwg] Removal of Ogg is *preposterous* Geoffrey Sneddon
      - [whatwg] Removal of Ogg is *preposterous* Manuel Amador (Rudd-O)
  - [whatwg] Removal of Ogg is *preposterous* Ian Hickson
    - [whatwg] Removal of Ogg is *preposterous* ddailey
    - [whatwg] Removal of Ogg is *preposterous* Maik Merten
      - [whatwg] Removal of Ogg is *preposterous* Ian Hickson
      - [whatwg] Removal of Ogg is *preposterous* Maik Merten
      - [whatwg] Removal of Ogg is *preposterous* Geoffrey Sneddon
      - [whatwg] Removal of Ogg is *preposterous* Charles McCathieNevile
      - [whatwg] Removal of Ogg is *preposterous* Maciej Stachowiak
    - [whatwg] Removal of Ogg is *preposterous* Manuel Amador (Rudd-O)
      - [whatwg] Removal of Ogg is *preposterous* Geoffrey Sneddon
      - [whatwg] Removal of Ogg is *preposterous* Robert Sayre
      - [whatwg] Removal of Ogg is *preposterous* Manuel Amador (Rudd-O)
      - [whatwg] Removal of Ogg is *preposterous* David Hyatt
      - [whatwg] Removal of Ogg is *preposterous* Manuel Amador (Rudd-O)
      - [whatwg] Removal of Ogg is *preposterous* David Hyatt
      - [whatwg] Removal of Ogg is *preposterous* Robert Sayre
      - [whatwg] Removal of Ogg is *preposterous* Manuel Amador (Rudd-O)
      - [whatwg] Removal of Ogg is *preposterous* Silvia Pfeiffer
      - [whatwg] Removal of Ogg is *preposterous* L. David Baron
      - [whatwg] Removal of Ogg is *preposterous* Ian Hickson
      - [whatwg] Removal of Ogg is *preposterous* Robert Sayre
      - [whatwg] Removal of Ogg is *preposterous*, SHOULD, and other matters ddailey
      - [whatwg] Removal of Ogg is *preposterous* Jeff McAdams

9

Long, long threads on WHATWG about Ogg vs MPEG-4 vs Flash vs whatever. There are "Free Software zealots" on one side, and immovable "corporate shills" on the other side. It's hard to say anything without sounding like you're in one of the two camps. Free Software and Open Source people have an emotional reaction to this issues; large corporations with a big investment in video (and arguably helped to create the video ecosystem on the Web in the first place) are not willing to use solutions that are very risky.

Welcome to the Jungle!

multimedia encapsulation formats

video & audio codec technology

legal (patent) status of everything

industry standards of everything (iso, mpeg, itu)

industry usage of everything
(not the same as industry standards!)

This is some of the prerequisite knowledge required to participate in debate, which a lot of people spouting their own opinion have no idea about. Not many people have this level of knowledge and can see the bigger picture and take all issues into account. (I'm not claiming I necessarily do, either…)

#include <stddisclaimer.h>

These are my personal opinions only, blah blah blah.
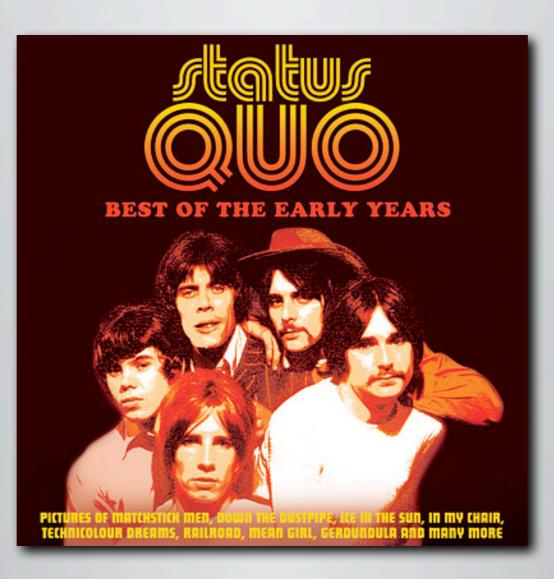
The three main players in this rat-race: Flash, MPEG, and xiph.org.

While everyone else is arguing about what to do, Flash is the status quo, and is more-or-less kicking everyone's arse right now. Flash mostly uses VP6, which is a proprietary, but they're switching to H.264 in recent Flash versions, which is a more advanced codec licensed by MPEG (see next slide). For every product that Adobe sells that supports Flash, something like $1 of that goes toward MPEG as a licensing cost.

MPEG isn't a big Web player, but is practically the completely dominant standard everywhere else, and have great influence. Everything is licensed and patented to the wazoo and is thus unsuitable for W3C adoption, but MPEG standards are used abso-bloody-lutely everywhere that video is, so their industry usage is not just massive; it's de-facto. Almost everything from DVDs to satellite TV to videos found in most files traded on BitTorrent uses MPEG-1, MPEG-2 or MPEG-4 (including that "DivX" stuff). They are a 10000-ton gorilla in the video space.
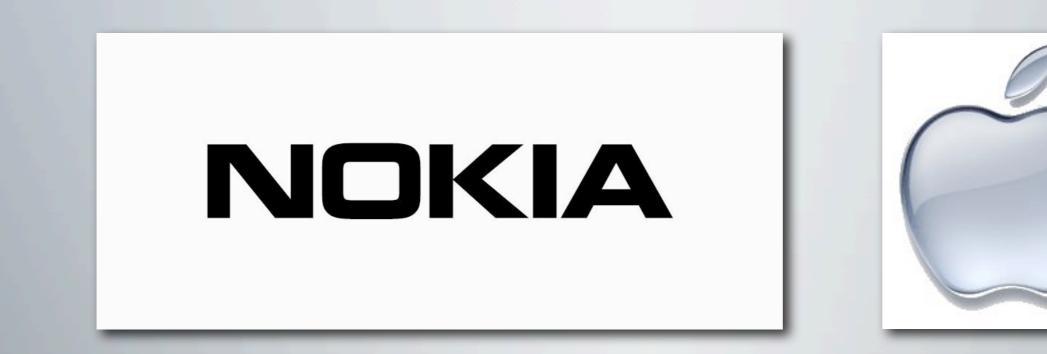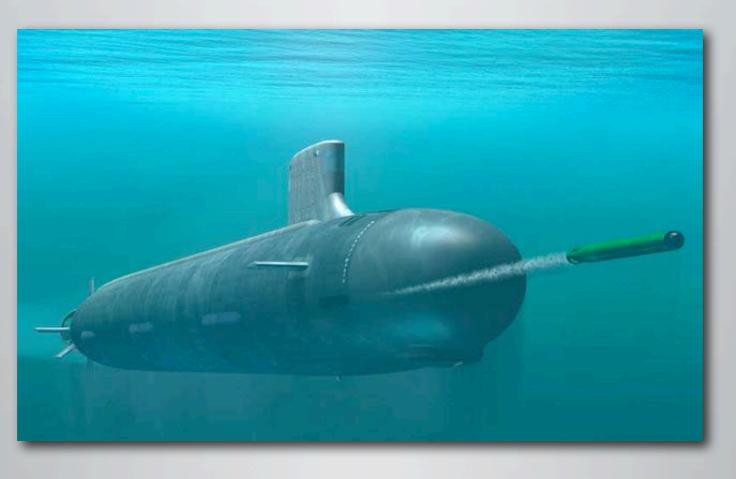
xiph.org provides the only truly free multimedia solution (video+audio) around: no licensing fees, no _known_ patents…

If xiph.org provide open multimedia standards, what's the problem with their standards (Ogg, Theora, Vorbis) being adopted for the Web?
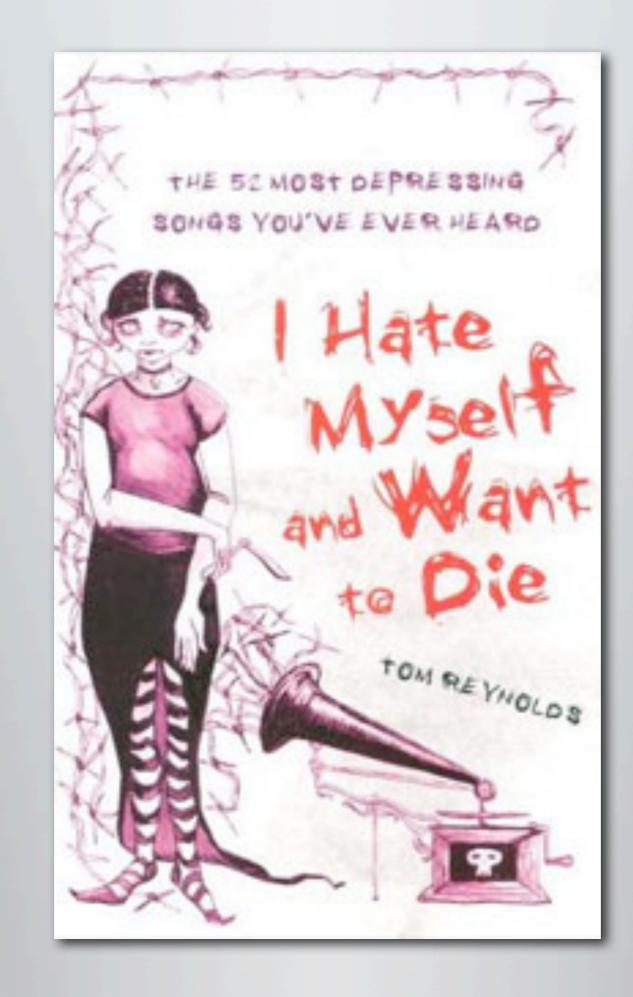
Apple and Nokia have very legitimate concerns about "submarine patents" in Theora (xiph.org's video codec) & Vorbis (xiph.org's audio codec). If Apple/ Nokia ship something with xiph.org codec support, they are the ones who will be sued if those codecs are violating any patents, because patent holders sue the companies that ship the product, not the developers. A W3C video codec standard is useless unless every single player is OK with the solution.

A W3C video workshop was held in December. It's likely that a video working group will be formed soon to move forward the discussion beyond flamewars on the whatwg mailing list.

The 52 Most Depressing Songs You've Ever Heard

I Hate Myself and Want to Die

Tom Reynolds

Legal issues are depressing, so let's move on to something else…

This is the Annodex project, which started around 2000.

What is Annodex?

Video is a passive entity right now: you sit there and watch it. This is completely unlike our normal behaviour when we browse the Web, where we can rapidly scan webpages for information, click through to find out about related information, and search for information via search engines such as Google and Yahoo!. Annodex wants to transform your Web video experience to be active rather than passive. In the same way that browsing the Web is an active experience vs reading a book, that's the case for Annodex vs the Web video of today. In short, it's shooting for the Massive Multimedia Matrix.

Video as a First-Class Citizen

- Searching

- Seeking

- Surfing

These three properties are what makes HTML special. Each of them has an analogy with Annodex. Annodex video can have embedded metadata inside them that make them amenable to text searching just like a standard web page. You can also rapidly navigate a video clip to chapter points and seek around—even though the video hasn't finished downloading yet—just like you can scroll a big webpage of text. Finally, Annodex can be a hyperlink source and destination. Not only can you jump from Annodex videos to normal webpages, but you can link to specific time points and chapter points in Annodex video so that video playback starts at fifteen minutes into it rather than the beginning, where the real information that you're after is.

#include <demodisclaimer.h>

Time for an Annodex demo. [I demoed some of the liboggplay Javascript testing here; see Shane Stephen's "Building a video remixing web-site using Annodex" talk at Linux.conf.au 2008 <http://linux.conf.au/programme/detail?TalkID=182> for a similar demo.] Note that this is all a work in progress and a tech demo, and it doesn't look very pretty at the moment since we don't have lovely designers doing spiffy graphics for us. [In the actual talk, the demo seemed to go mostly OK, woohoo.]

Video as a First-Class Citizen... on Web 2.0

Web 2.0 made webpages dynamic and interactive, and encouraged communities to emerge. YouTube may be a community, but it's hardly dynamic yet. Unfortunately, a five-minute demo of Annodex will do more to convince you about its video remixing and hyperlinking capabilities than twenty minutes of boring reading will...

Video as a First-Class Citizen... on Web 2.0

- Youtube 2.0

- dynamic composition of videos across websites