W3C Video on the Web  
Towards a Standard Specification for Web Video  
Ranga Muvavarina  
Architect, Comcast Interactive Media

Video distribution networks are clearly shifting away from ‘traditional’ models that are characterized by closed supplier networks, complex IP protection and entitlement, clear metadata specifications [e.g. cablelabs], very coarse-grained content-packaging and physically bounded distribution networks... to more open and participatory models that are characterized by increasingly open supplier networks, evolving IP protection and entitlement regimes, multiple metadata specifications [cablelabs, movielabs], fine-grained content-packaging and peer-to-peer distribution via broadband networks.

Comcast Interactive Media [CIM] is looking to leverage its reach in broadband and television to create a cross-platform experience that spans this spectrum. Specifically, CIM is focused on building content distribution platforms that aggregate premium content [www.fancast.com] for subsequent distribution via multiple channels -- including online [fancast, ziddio], in-theater [fandango], mobile and on-tv [fancast.com]. Additionally, CIM has built platforms for aggregation and distribution of UGC [www.ziddio.com].

Searching for and Navigating Web Video
The consumption of video content on the web is both fine-grained and highly personalized. In order to effectively deliver personalization, it is imperative that web video content be both findable and navigable.

- **Standard schema:** The lack of a widely-adopted, open metadata standard for web video is a significant impediment. Access to the current repositories of web video content requires that content distributors implement multiple provider specific ingestion frameworks [aol.videos.com, youtube.com, fancast.com etc]. Beyond the basic video attributes [title, release date...] these assets have divergent schemas. The economics of web video suggest an imminent explosion in the number of premium content providers... putting significant pressure on content aggregators. A common framework will significantly improve syndication web video [particularly by distributors]

- **Video to entity graphs:** Tied to this is an inability to efficiently correlate web video content to other entities on the web [e.g. actor to spoof]. At present, collaborative techniques [e.g. tagging] partially satisfy the need to leverage explicit relationships in a cross content-graph query. A standard schema for web-video would enable queries across large content graphs -- significantly improving findability and navigation.

- **Topical segmentation:** Although noth short form “snack” size and long form (e.g. 15~20 minutes or more) video content is typically considered as one contiguous asset, for a significant portion of long form content its consumption [and attendant social value] is largely driven by topically related sub-segments. Even though speech to text technologies can extract/generate a lot of metadata from web video, the lack of a consistent schema [e.g. text to time-offset mapping] reduces the opportunities available to leverage this metadata. For example, YouTube content is dominated by topically specific segments of longer-form content.
Delivering Web Video
Although the consumption model for web video continues to evolve, a large proportion is delivered via download-to-play. According to a report published by YouTube, 16% of Americans now routinely watch web video; with 2% of Americans having either watched or downloaded a full length video. A larger proportion of web video is delivered via streaming protocols. As broadband [DSL, Cable, Fiber, wireless] penetration grows [~100m households, 300m+ media-enabled phones], we foresee an attendant increase in video consumption via streaming delivery.

- **Dynamic quality management:** Bandwidth and quality of service characteristics of these varying broadband networks suggest that streaming web videos [particularly to the ever increasing number of wireless/mobile devices] will require variable bit-rate encoding schemes -- that leverage network intelligence to dynamically tune-up/down bitrates.

- **Protocols:** The attendant proliferation of codecs, media players and protocols presents significant [but surmountable] operational challenges for content distributors.

Conclusion
In order to establish web video as a first-class, linked, citizen of the web, we need to ensure that global repositories of web video are findable and navigable. This requires that we establish a common set of specifications for managing web video metadata. Additionally, we need to ensure that the web video consumption experience continues to close the quality gap with traditional video.