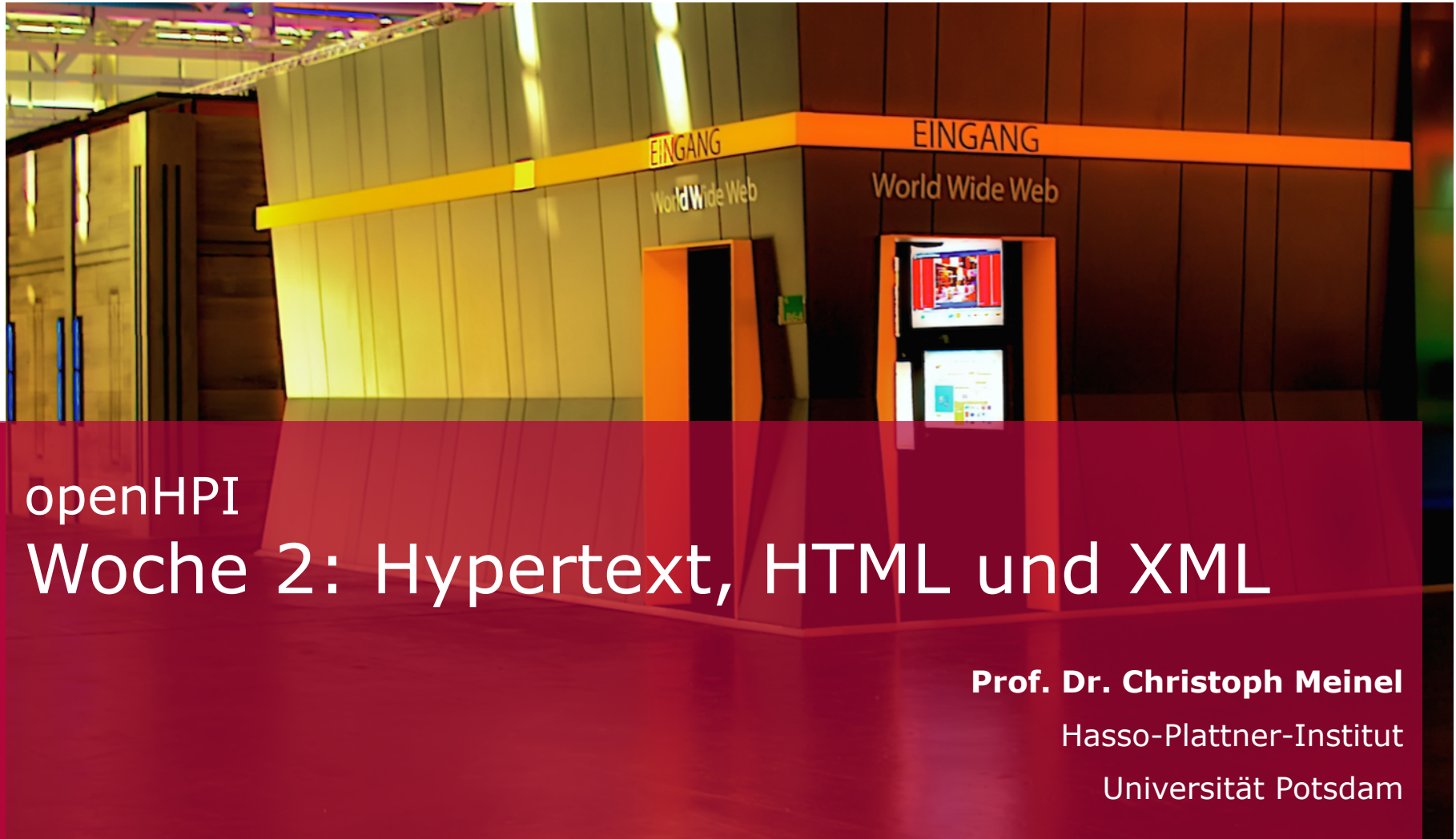




openHPI

Woche 2: Hypertext, HTML und XML

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

WWW ist ein **Hypermedia-System** auch **Hypertext-System**, **WWW-Dokumente** sind durch **Hyperlinks** untereinander verbunden und auf über die Welt verteilten Servern gespeichert. Das WWW wird so zum **verteilten** Hypermedia-System

- Hypermedia-Dokumente überwinden die traditionelle, lineare Dokumentenstruktur – mit dem Link-Mechanismus werden sie Teil eines riesigen Informationsnetzwerks
- Nutzer können ein WWW-Dokument direkt von einem anderen WWW-Dokument aus erreichen, selbst wenn diese auf unterschiedlichen Servern gespeichert sind
- Für die auf verschiedenen Servern gespeicherten WWW-Dokumente ist es schwierig, die Konsistenz von Hyperlinks zu gewährleisten

Hypertext und Hypermedia (2/2)

WWW-Dokumente werden typischerweise mit einer speziellen deskriptiven Sprache beschreiben: der **Hypertext Markup Language – HTML**

Mit HTML können Autoren:

- **Struktur** eines Dokuments beschreiben, also z.B. Unterteilung in (Unter-)Überschriften und Absätze, Listen, Tabellen, ...
- Hyperlinks in die Dokumente einbetten,
- Multimedia-Komponenten einbinden und
- (bis zu einem gewissen Grad) die graphische Aufbereitung beeinflussen

Als WWW-Dokumente werden aber nicht nur HTML-Dokumente bezeichnet, sondern alle Dokumente, auf die im WWW zugegriffen werden kann

HTML-Dokumente (1/3)

- HTML-Dokument heißt **Page** (dt. Seite)
- Die Startseite oder Eingangsseite für die Navigation durch ein Informationsangebot im WWW nennt man **Homepage**.
Zunehmend hat sich dieser Begriff aber auch als Bezeichnung für das gesamte Informationsangebot eines Anbieters (die **Website**) durchgesetzt
- Browser fordern ein HTML-Dokument an, interpretieren es und stellen es dann anhand der dort angegebenen HTML-Elemente dar.
- Jedes HTML-Element ist durch einen **Tag** gekennzeichnet. Tags sind als **Markups** in das Dokument eingebettet

HTML-Dokumente (2/3)

Markup-Sprachen sollten nur für die Beschreibung der Dokumentenstruktur genutzt werden, nicht jedoch für die Beschreibung der Präsentation des Dokuments

- Grundlegende Idee von Markup-Sprachen ist die **Trennung** von **Dokumentenstruktur** und **Dokumentengestaltung**
- Mit **HTML** wird die Dokumentstruktur beschrieben
- **Cascading Stylesheets – CSS** – sind verantwortlich für die Beschreibung der Gestaltung (das Layout) dieser Struktur
 - Wichtigste Anwendungsgebiete für CSS:
 - exakte Definition des Layouts eines HTML-Dokuments
 - Anpassung an unterschiedliche Ausgabemedien
 - zentrales Management von Layouts

HTML-Dokumente (3/3)

- Mit Hilfe spezieller Tags können Referenzen als **Links** in HTML-Dokumente eingebaut werden, die dort als passive „Wegweiser“ zu anderen Dokumenten dienen
- Die global eindeutige Identifikation der verlinkten Dokumente erfolgt mit Hilfe von **URIs** (Uniform Resource Identifiers)
- URIs können als Zeichenketten („Strings“) kodiert werden und geben in der Ausprägung als **URLs** den „Ort“ (Server) an, an dem ein Dokument gespeichert ist:
 - Name des Zugriffsprotokolls (z.B. http oder https),
 - Name des Servers (meist Domain Name),
 - Pfad (in der Verzeichnisstruktur des Servers)
 - und (Datei-)Name des Dokuments

Dynamisches HTML

- **DHTML – Dynamisches HTML:** Ein HTML-Dokument kann während der Anzeige durch Benutzerinteraktion (oder andere Ereignisse) verändert werden, z.B. Ausklappen eines Menüs, Runterzählen eines Countdowns, ...
- Typische Technik dahinter ist das → **DOM-Skripting** (→ DOM – Document Object Model). Das sind Programme (Skripte), die im Browser ausgeführt werden und auf bestimmte Ereignisse (z.B. Maus-Clicks) reagierend das HTML-Dokument manipulieren
- Während der Anzeige eines Dokuments können Browser und Server mittels → **XMLHttpRequest** kommunizieren, bekannt z.B. aus:
AJAX = **A**synchronous **J**avaScript **a**nd **X**ML

Werden uns in Kurswochen 3 und 4 detaillierter mit DOM-Skripting und AJAX bzw. allgemeiner mit der **Client-seitigen Web-Programmierung** befassen

- Hauptkritikpunkt an HTML ist eingeschränkte Flexibilität aufgrund der in ihrer Bedeutung unveränderlich festgelegten Markups
- Mit der Einführung von **XML – Extensible Markup Language** – als Meta-Markupsprache ist es möglich, alle nur erdenklichen Dokumenttypen mit einer eigenen, speziellen Syntax zu beschreiben
- XML ist Ausgangspunkt für Reihe Anwendungs-spezifischer Markup Sprachen, die genau auf die jeweiligen Anwendungsdomänen oder spezialisierte Ausgabegeräte zugeschnitten sind. Beispiele:
 - **MathML** – Mathematical Markup Language
 - **SVG** – Scalable Vector Graphics
 - **ODF** – Open Document Format for Office Applications
 - ...

Themen der zweiten Kurswoche (1/2)

HTML – Hypertext Markup Language

- **HTML-Dokumente:** Wie funktioniert HTML grundlegend und wie sind HTML-Dokumente aufgebaut?
- **HTML-Syntax und guter Stil:** Welche Regeln sind beim Aufbau von HTML-Dokumenten zu beachten und welche Konventionen gibt es?
- **HTML-Elemente:** Welche Tags gibt es für welche HTML-Elemente, (z.B. Listen, Bilder, Überschriften) und wie funktionieren Attribute?
- **HTML5:** Welche Neuerungen bringt die aktuelle HTML-Version – HTML5 – und welche Tags sollte man nicht mehr verwenden?

Themen der zweiten Kurswoche (2/2)

CSS – Cascading Stylesheets

- **CSS:** Wie funktioniert die Gestaltung von HTML-Dokumenten mit Cascading Stylesheets und wie baut man diese in HTML ein?
- **CSS Box Modell:** Wie funktioniert die Positionierung von HTML-Elementen mit CSS?
- **Responsive Web Design:** Wie kann man mit HTML und CSS Webseiten für verschiedene Endgeräte (z.B. Desktop und Smartphone) bauen?
- **CSS Präprozessoren:** Wie kann das syntaktisch etwas umständliche CSS eleganter formuliert werden?

XML – Extensible Markup Language

- **XML:** Wie kann man mit XML spezielle Dokumententypen für den generellen Austausch von Daten zwischen Computersystemen definieren?