

Adaption von Webseiten

Methoden und Ansätze

Mobile Business Seminar

Universität Mannheim
Lehrstuhl für Praktische Informatik IV
im Wintersemester 2005/2006

Ali Ikinci
ali@ikinci.de

Agenda

- ***Architektur***

Inhaltsanalyse

Anpassung

XML-Framework

Motivation Adaption

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

- Zunehmende Verbreitung von Personal Digital Assistants, Smart Phones und anderer heterogener Geräte
- Bessere Anbindung ans Internet über GPRS, UMTS und WLAN
- Zentrale Frage: Wie kann man Webseiten, die für den PC entwickelt worden sind, auf mobilen elektronischen Geräten unter Berücksichtigung der Beschränkungen im Hinblick auf Input/Output-Kapazitäten darstellen?

Motivation Adaption

- Anbieter verwenden spezielle Repräsentationen für die Inhalte
- Anwender haben einen spezifischen Client mit besonderen Eigenschaften

Architektur

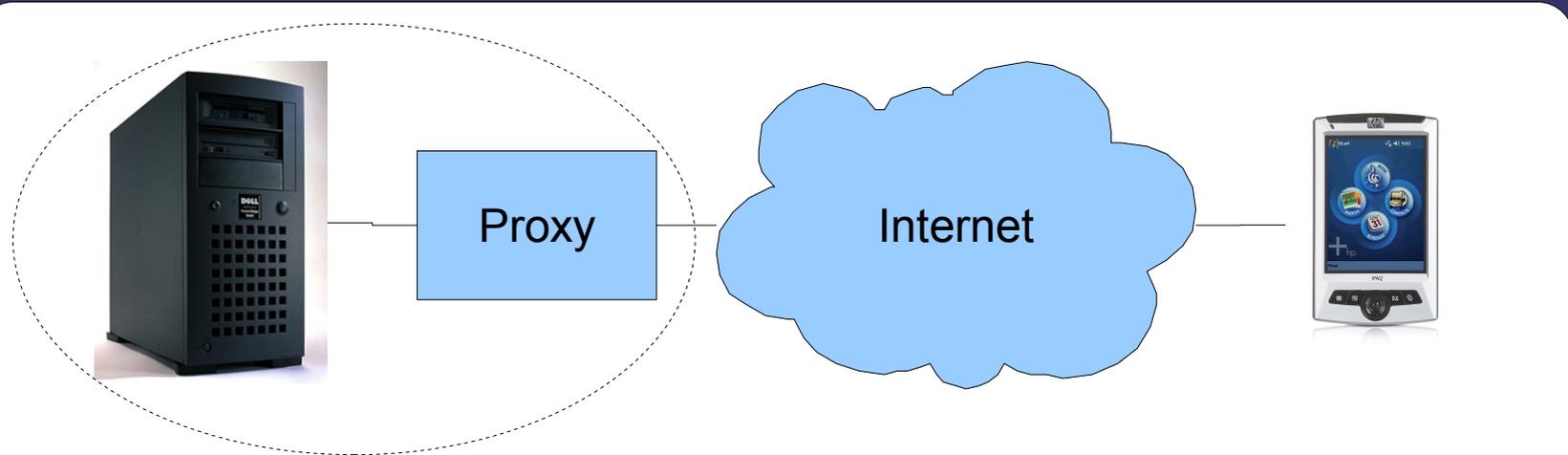
Inhaltsanalyse

Anpassung

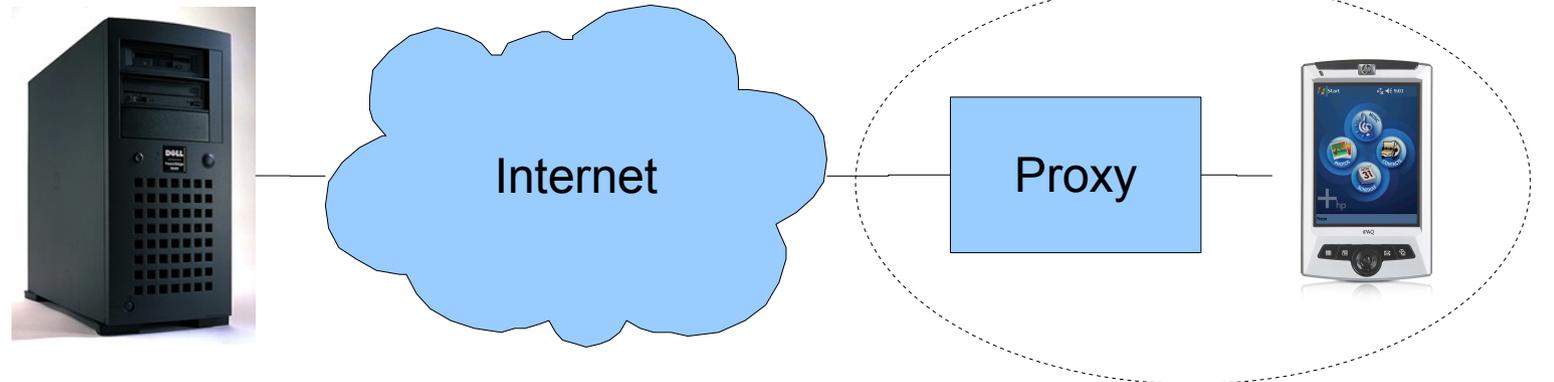
XML-
Framework



Architektur



Adaption beim Anbieter



Adaption beim Anwender

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

Mögliche Arbeitsschritte

- **Analyse der Struktur**
 - **Bestimmung der Inhaltsblöcke und der Hierarchie**
 - **Tabellenanalyse**
 - **Imagemap-Analyse**
- **Analyse des Inhalts**
 - **Textanalyse**
- **Anpassung einer Webseite**
 - **Textanpassung bzw. Textzusammenfassung**
 - **Medienanpassung (nicht Teil dieser Arbeit)**
 - **Erzeugung von Thumbnails**
 - **Anpassung der Abfolge von zusammengesetzten Seiten (Page-Splitting)**
- **Zusammensetzung einer Webseite**
 - **Automatische Neupositionierung**
 - **Berücksichtigung des Benutzerverhaltens**
 - **Berücksichtigung von Benutzerpräferenzen (nicht Teil dieser Arbeit)**

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

Agenda

Architektur

 ***Inhaltsanalyse***

Anpassung

XML-Framework

Identifikation der Inhaltsblöcke

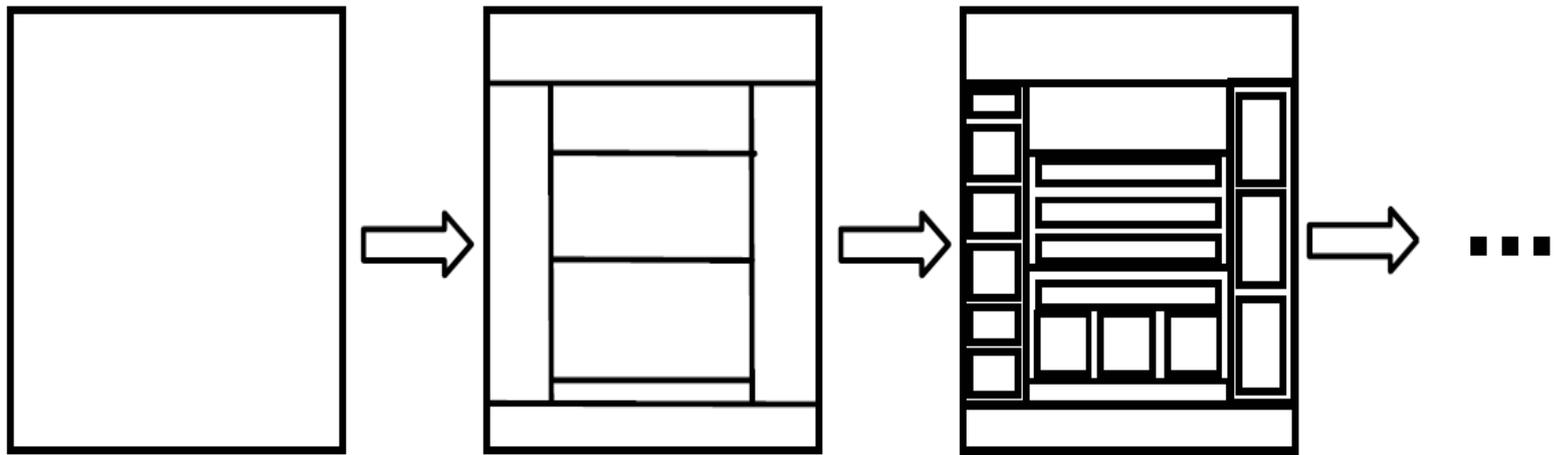
- Informationseinheiten als Inhaltsblöcke
- Iterative Vorgehensweise
 - durch sukzessives Aufteilen des initialen Inhaltsblocks in immer kleinere Blöcke
 - Klassifizierung durch Kopfzeile, Fußzeile, linker Rand, rechter Rand bis zum atomaren Inhaltsblock

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

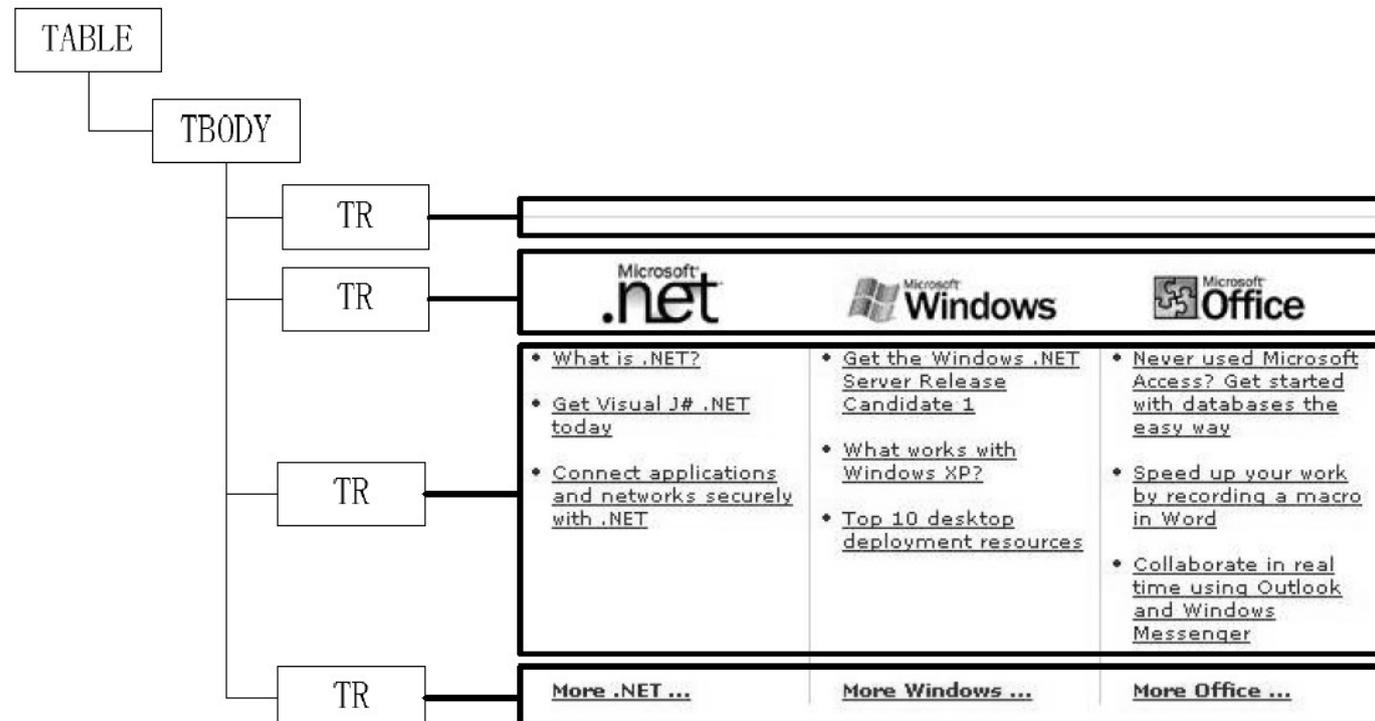


Tabellenanalyse

Identifikation von Begrenzungen zur weiteren Aufteilung der Inhaltsblöcke:

- <HR> wird oft horizontaler Begrenzer
- <TR> und <TD> Block Begrenzung
- Bilder können auch als Begrenzung dienen wenn sie charakteristische Eigenschaften aufweisen

Architektur
Inhaltsanalyse
Anpassung
XML-
Framework



Tabellenanalyse

Architektur
Inhaltsanalyse
Anpassung
XML-
Framework

Microsoft .NET	Microsoft Windows	Microsoft Office
<ul style="list-style-type: none">• What is .NET?• Get Visual J# .NET today• Connect applications and networks securely with .NET	<ul style="list-style-type: none">• Get the Windows .NET Server Release Candidate 1• What works with Windows XP?• Top 10 desktop deployment resources	<ul style="list-style-type: none">• Never used Microsoft Access? Get started with databases the easy way• Speed up your work by recording a macro in Word• Collaborate in real time using Outlook and Windows Messenger
More .NET ...	More Windows ...	More Office ...

Imagemap Analyse

Aus Imagemaps werden Indexseiten generiert, der Name der Links wird aus dem <ALT>-Tag oder den Links selbst extrahiert

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework



ALABAMA
ALASKA
ARIZONA
ARKANSAS
CALIFORNIA
COLORADO
CONNECTICUT
DISTRICT OF COLUMBIA
DELAWARE
FLORIDA

Textanalyse: Keywords

Zusammenfassende Informationen über Texte in den atomaren Inhaltsblöcken mit:

Keywords:

- geben den Inhalt des Textes stichwortartig wieder
- jedes Wort hat eine *importance*
- *importance* = $\frac{\text{Anzahl der Häufigkeit im Text}}{\text{Anzahl der Häufigkeit in Gesamtdokument}}$

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

Textanalyse: Summary Sentence

Summary Sentence:

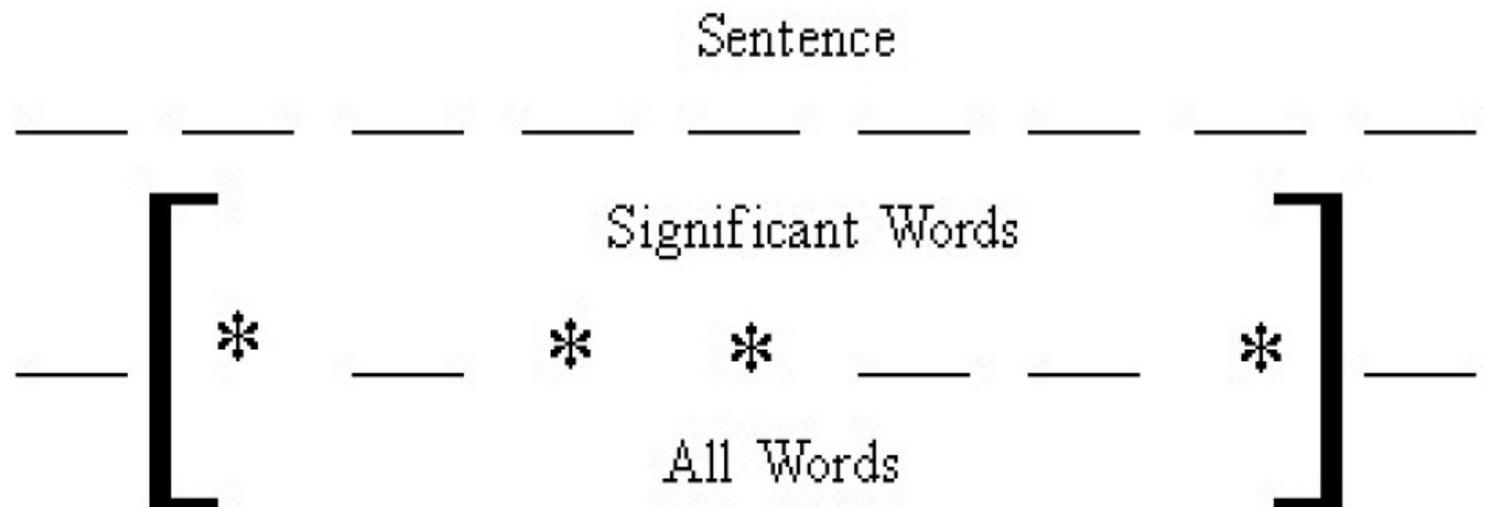
- geben den Inhalt des Textes bestmöglich wieder
- jeder Satz hat einen *significance factor*

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework



Agenda

Architektur

Inhaltsanalyse

 ***Anpassung***

XML-Framework

Textzusammenfassung

Sukzessive Darstellung in vier Variationen:

- *Incremental*: drei Zustände; die erste Zeile; die ersten drei Zeilen; der gesamte Text
- *Keyword*: die wichtigsten *keywords*; die ersten drei Zeilen des Texts; der gesamte Text

Incremental

- Want a Vaccine With That?
- ◐ Want a Vaccine With That?
Genetic engineers from
Bayer Thompson Institute at
- Want a Vaccine With That?
Genetic engineers from
Bayer Thompson Institute at
Cornell University have
developed an edible vaccine
against potentially deadly

Keyword

- vaccine diseases diarrhea
cholera
- ◐ Want a Vaccine With That?
Genetic engineers from
Bayer Thompson Institute at
- Want a Vaccine With That?
Genetic engineers from
Bayer Thompson Institute at
Cornell University have
developed an edible vaccine

Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

Textzusammenfassung

- *Summary*: zwei Zustände; der wichtigste Satz; der gesamte Text
- *Keyword/Summary*: die *keywords*; der wichtigste Satz; der gesamte Text

Summary

- Genetic engineers from Bayer Thompson Institute at Cornell University have developed an edible vaccine against potentially deadly diseases.
- Want a Vaccine With That? Genetic engineers from Bayer Thompson Institute at Cornell University have developed an edible vaccine against potentially deadly diseases. They combined genes from bacteria and viruses with potato cells and

Keyword/Summary

- vaccine diseases diarrhea cholera
- Genetic engineers from Bayer Thompson Institute at Cornell University have developed an edible vaccine against potentially deadly diseases.
- Want a Vaccine With That? Genetic engineers from Bayer Thompson Institute at Cornell University have developed an edible vaccine against potentially deadly

Architektur

Inhaltsanalyse

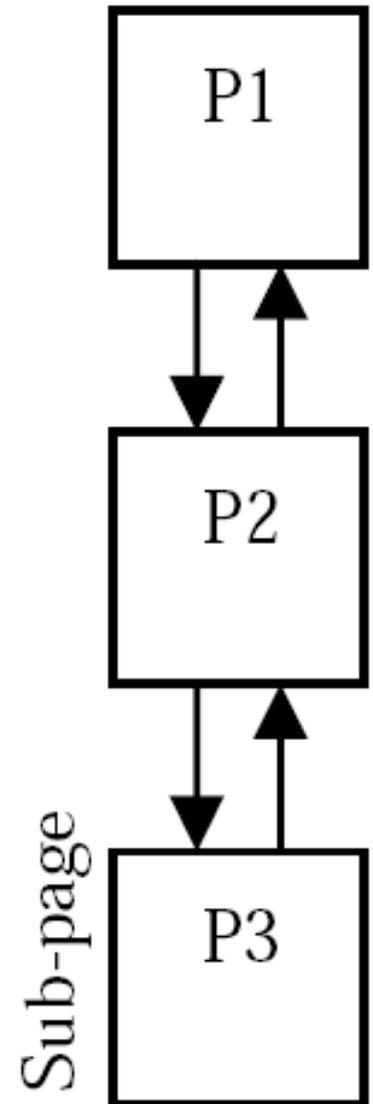
Anpassung

XML-
Framework

Page-Splitting

Beim Page-Splitting wird ein Inhaltsblock auf mehrere Seiten verteilt:

- Single-subject splitting: Um einen längeren Inhaltsblock ohne Scrollen darzustellen wird er auf mehrere Seiten aufgeteilt und mit next/previous Links ausgestattet



Architektur

Inhaltsanalyse

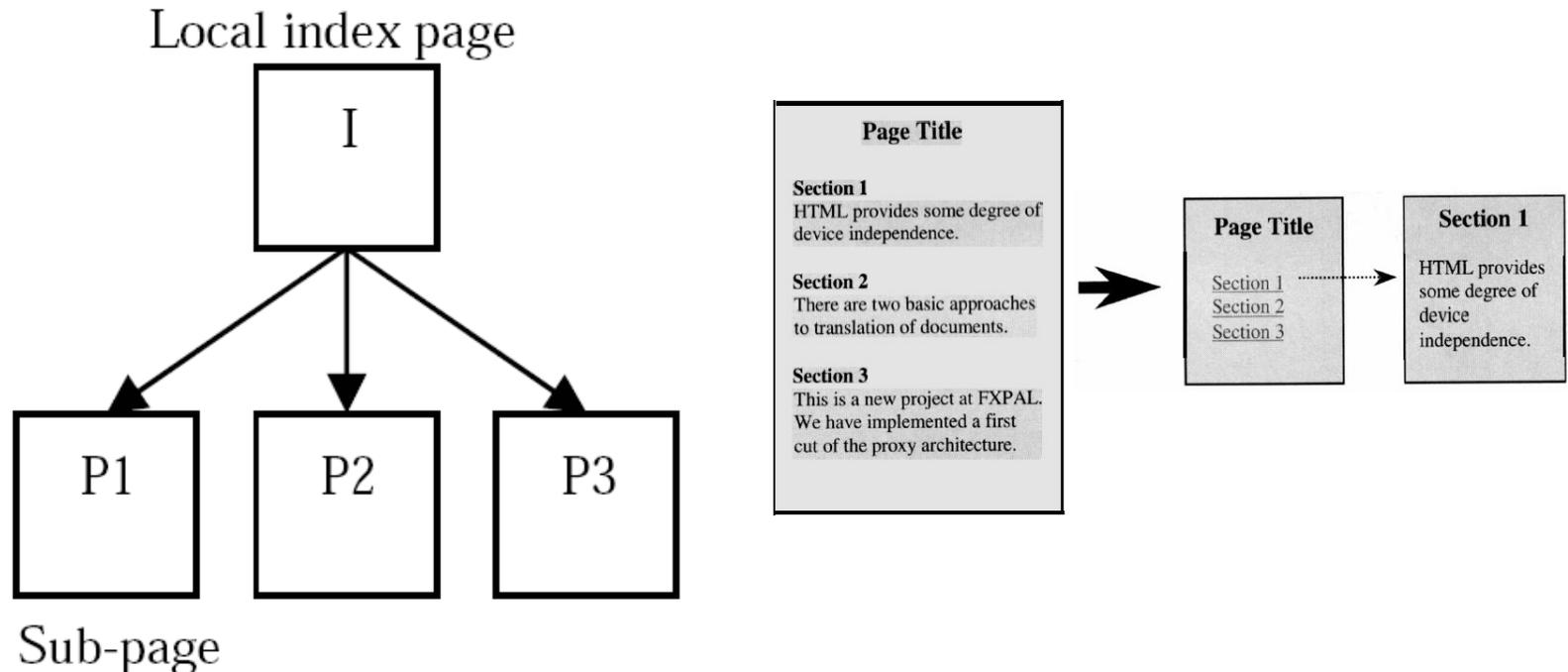
Anpassung

XML-
Framework

Page-Splitting

- Multi-subject splitting generiert eine neue Indexseite für verschiedene Unterseiten
- Section-outlining transformation: Kapitel erhalten eigene Seiten, eine neue Indexseite wird für sie generiert

Architektur
Inhaltsanalyse
Anpassung
XML-
Framework



Thumbnail-Browsing

- komplette Darstellung der Webseite als Thumbnail zur Indizierung und besseren Navigation



Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

Auto-Positioning vs. Page-Splitting

- Page-Splitting kann u.U. den Bildschirm nicht voll ausnutzen
- ein Link zurück erschwert die Navigation



Architektur

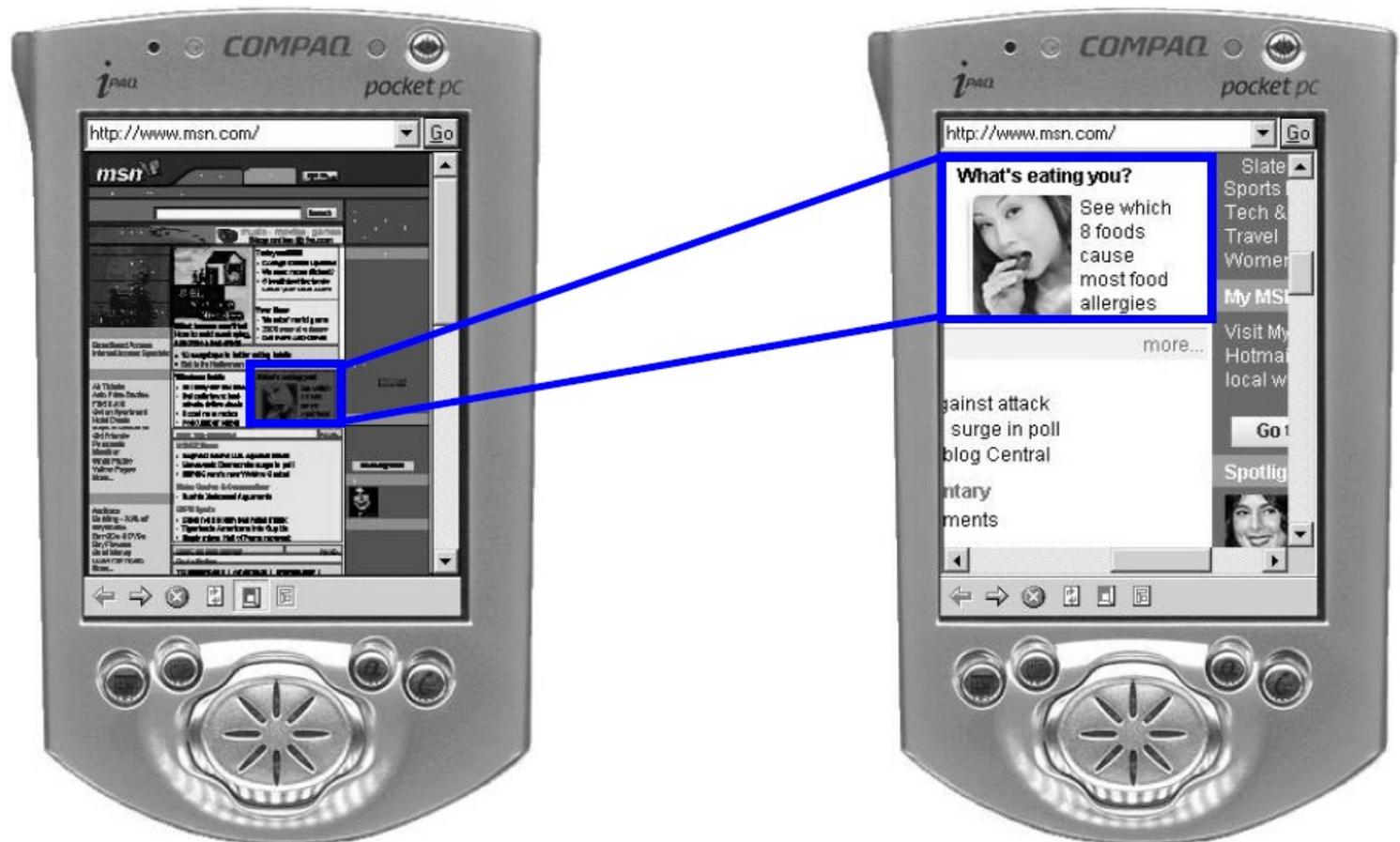
Inhaltsanalyse

Anpassung

XML-
Framework

Auto-Positioning vs. Page-Splitting

- Auto-Positioning vereinfacht die Navigation und erleichtert die Bedienung durch zusätzliche Scrolloptionen



Architektur

Inhaltsanalyse

Anpassung

XML-
Framework

Agenda

Architektur

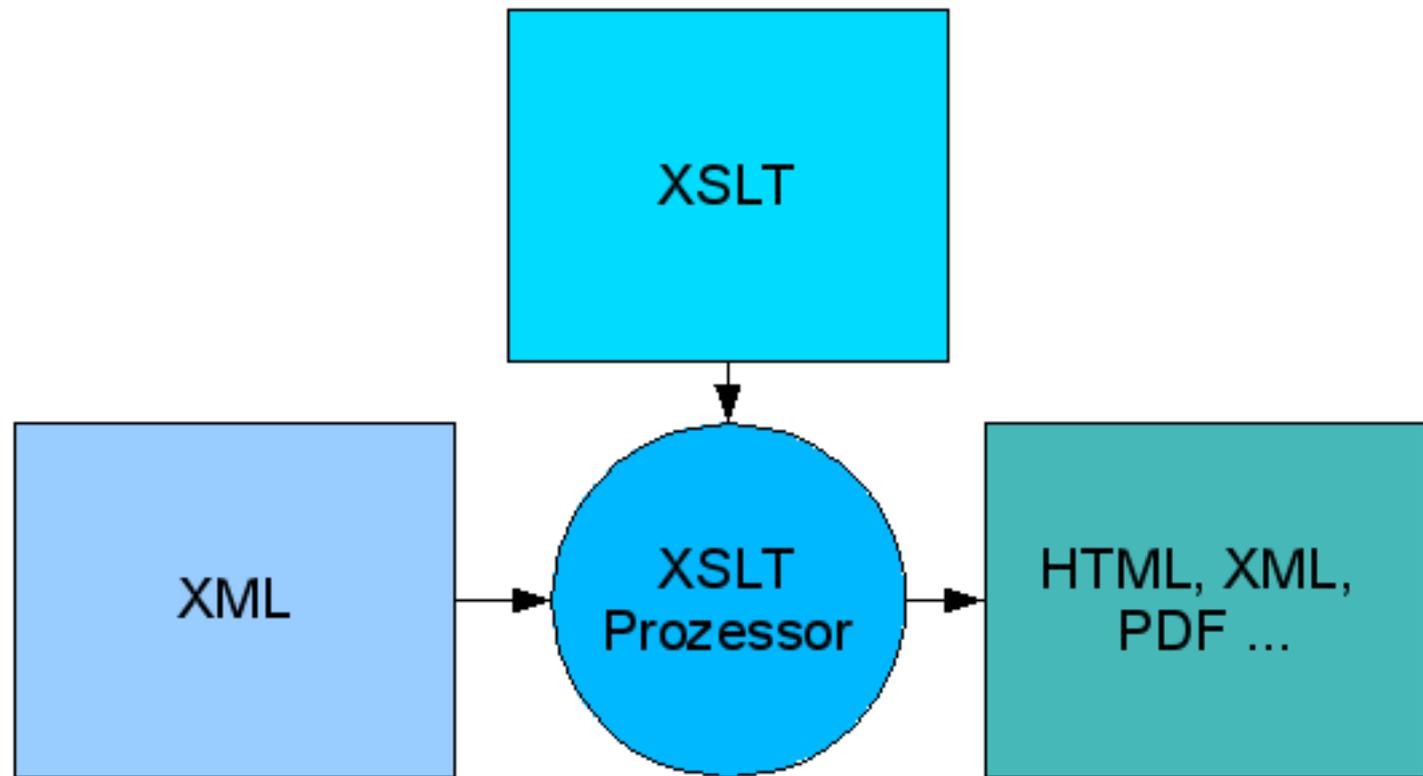
Inhaltsanalyse

Anpassung

 ***XML-Framework***

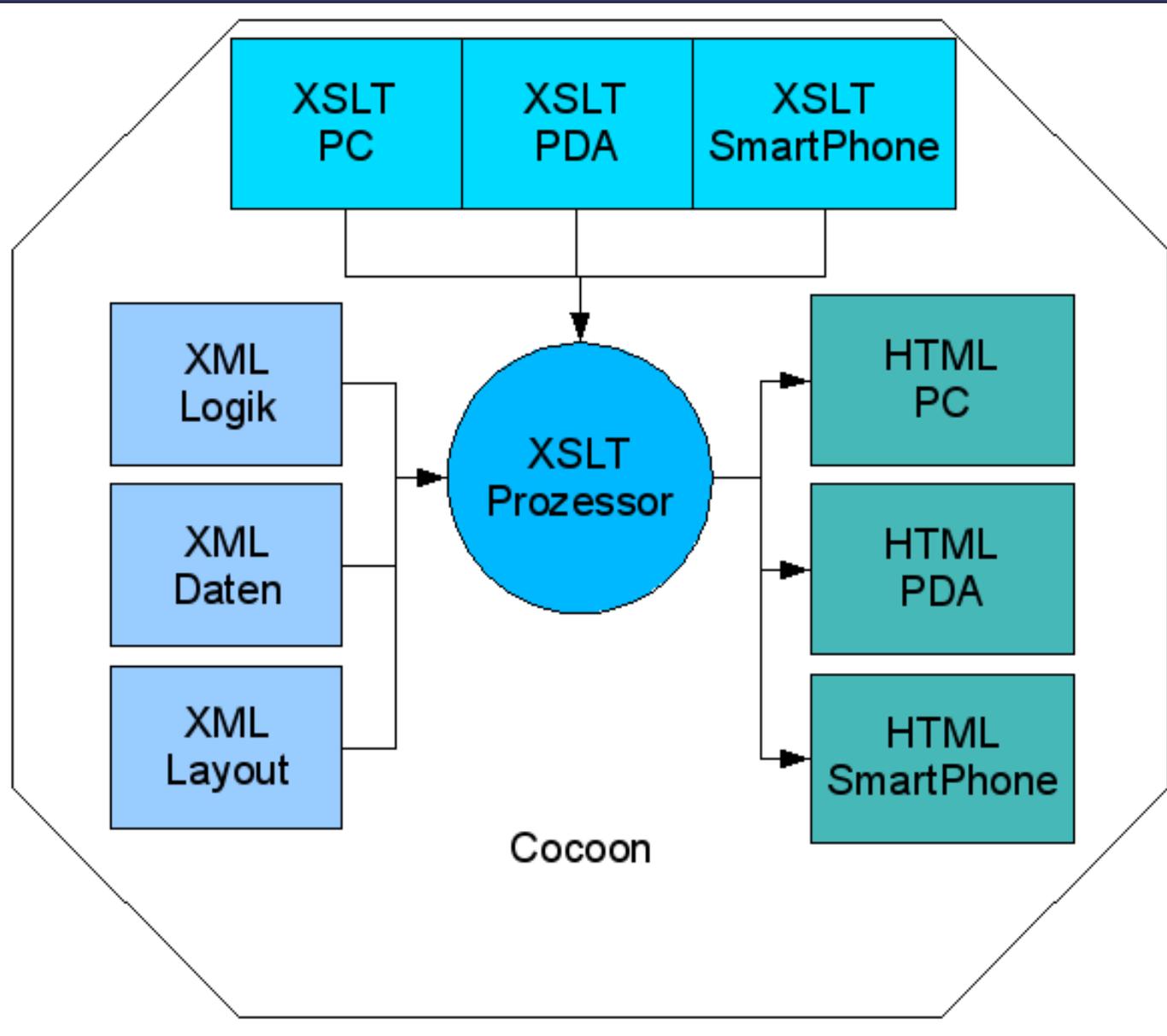
XSLT

Eine XSLT-Description kann eine Umwandlung von XML-Dokumenten in viele andere Formate ermöglichen



Architektur
Inhaltsanalyse
Anpassung
***XML-
Framework***

Cocoon



Architektur
Inhaltsanalyse
Anpassung
**XML-
Framework**

Zusammenfassung

- Viele Teillösungen
- Endgültige Lösung für Adaption nicht möglich
- Oft Überdimensionierung der Clients statt Anpassung an Gegebenheiten
- Lösung evtl. durch neue Displaytechnologien und neue Eingabetechnologien
- Großes Interesse von kommerziellen Anbietern führt zu „Portaloasen“ und intensiven Forschungstätigkeiten

Noch Fragen ?