



**Iss, damit du groß und stark wirst!**



Problematiken und Lösungen eines wachsenden IT-  
Unternehmens nachgezeichnet an



**WEB.DE**

Referentin:

Jennifer Stoll  
Technical Director  
WEB.DE AG



**Iss, damit du groß und stark wirst!**



## **Gliederung des Vortrags**

- Geschäftsfelder der Cinetic Medientechnik GmbH und WEB.DE AG
- Entwicklung des Portals
- Architektur V0.0: Die Anfänge von WEB.DE und C.O.P.S.
- Problematik: Externes vs. internes Hosting
- Architektur V1.0: Die erste C.O.P.S Release
- Skalierung die erste: Mit Hardware erschlagen
- Der Einzug von Linux, Problematik: Proprietäres vs. Open Source
- Skalierung die zweite: Verteilung der Architektur
- Die Gründung der AG
- Going Public
- Das neue Rechenzentrum
- Architektur V2.0: C.O.P.S. – The Next Generation
- Problematik: Performance und Ausfallsicherheit
- Problematik: Administration und Monitoring



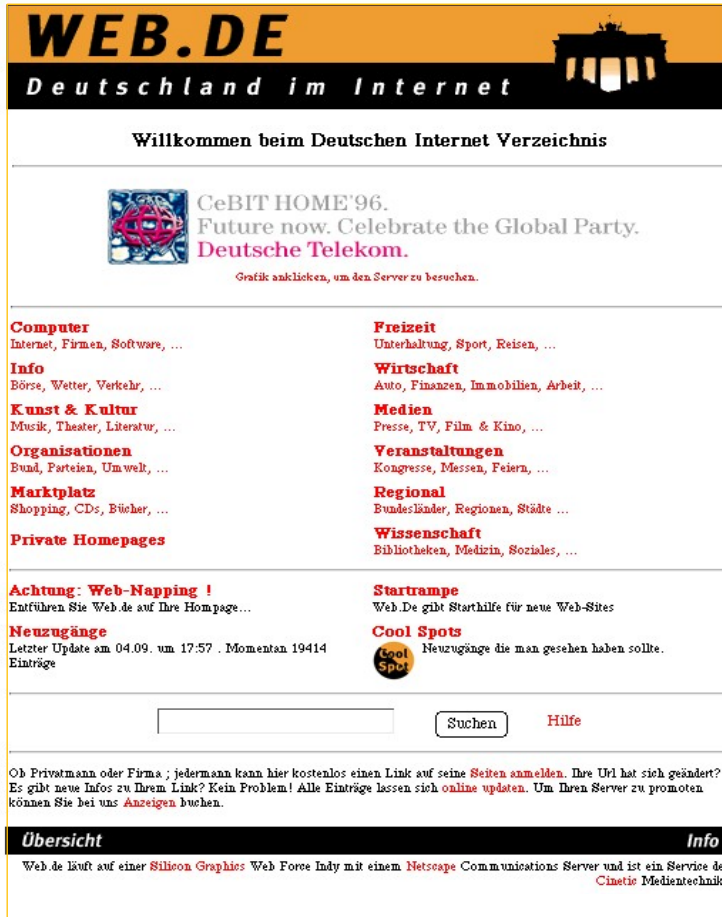
## Entscheidung für das Internet: 1996



### Die Cinetic Medientechnik GmbH

- Vor 1996
  - Entwicklung und Vertrieb der VideoMachine für den Macintosh (für FAST)
  - Diverse CD-Rom Projekte im Multimedia-Bereich
  
- Anfang 1996 hatte die Cinetic 3 Geschäftsfelder:
  - Das Internet Systemhaus
    - Realisierung und technische Betreuung von Web-Auftritten für Großkunden (z.B. BASF, Comparex, Pharma-Konzerne)
  
  - Die Touristik
    - Online-Flugbuchungssysteme für Endkunden (flug.de, lastminute.de)
  
  - Das Portal WEB.DE
    - Redaktionell betreuter Katalog von Web-Seiten

## Von der Suchmaschine zum Portal.



**WEB.DE**  
Deutschland im Internet

Willkommen beim Deutschen Internet Verzeichnis

 CeBIT HOME'96.  
Future now. Celebrate the Global Party.  
Deutsche Telekom.  
Grafik anklicken, um den Server zu besuchen.


<b>Computer</b> Internet, Firmen, Software, ...	<b>Freizeit</b> Unterhaltung, Sport, Reisen, ...
<b>Info</b> Börse, Wetter, Verkehr, ...	<b>Wirtschaft</b> Auto, Finanzen, Immobilien, Arbeit, ...
<b>Kunst &amp; Kultur</b> Musik, Theater, Literatur, ...	<b>Medien</b> Presse, TV, Film & Kino, ...
<b>Organisationen</b> Bund, Parteien, Umwelt, ...	<b>Veranstaltungen</b> Kongresse, Messen, Feiern, ...
<b>Marktplatz</b> Shopping, CDs, Bücher, ...	<b>Regional</b> Bundesländer, Regionen, Städte ...
<b>Private Homepages</b>	<b>Wissenschaft</b> Bibliotheken, Medizin, Soziales, ...

**Achtung: Web-Mapping !**  
Entführen Sie Web.de auf Ihre Homepage...

**Neuzugänge**  
Letzter Update am 04.09. um 17:57 . Momentan 19414 Einträge

**Startrampe**  
Web.De gibt Starthilfe für neue Web-Sites

**Cool Spots**  
Neuzugänge die man gesehen haben sollte.



[Hilfe](#)

Ob Privatmann oder Firma ; jedermann kann hier kostenlos einen Link auf seine [Seiten anmelden](#). Ihre Url hat sich geändert? Es gibt neue Infos zu Ihrem Link? Kein Problem! Alle Einträge lassen sich [online updaten](#). Um Ihren Server zu promoten können Sie bei uns [Anzeigen](#) buchen.

**Übersicht** **Info**

Web.de läuft auf einer [Silicon Graphics](#) Web Force Indy mit einem [Netscape](#) Communications Server und ist ein Service der [Cinetix](#) Medientechnik.

- 1995: WEB.DE geht online
- 1997/1998: zahlreiche Innovationen von „Live-Suche“ bis „FreeMail“
- 1999: Umwandlung in die WEB.DE AG
- 2000: Erfolgreicher Börsengang an den Neuen Markt
- 2001: Über 350 Mitarbeiter

# Das Portal (II)



## WEB.DE ist das „Vollprogramm“ im Internet.

The screenshot shows the WEB.DE homepage with a blue header containing the logo and navigation links (Suche, Themen, Dienste, FreeMail, Hilfe, Kontakt). Below the header, there are promotional banners for '1.000 Klingeltöne & Logos...' and 'Neu: WEB.DE Weckruf'. A search bar is prominently displayed with a search button and a 'Suchen' button. Below the search bar, there are several service categories listed in a grid, including 'Dienste', 'Nachrichten', 'Neu: Themen', and 'Cool: Asterix-Archiv'. On the right side, there is a vertical sidebar with more services like 'MillionenKlick', 'FreeSMS', 'FreeMail', 'SmartSurfer', 'Shopping', and 'Grußkarten'. At the bottom, there are footer links for 'allEurope Partnerguides', 'WEB-Napping', and 'English Information on WEB.DE AG'.

Navigation

Kommunikation

Information

Community

E-Commerce

### **WEB.DE hat, was die User im Netz suchen.**

- In Breite und Tiefe einzigartig: z. Zt. 37 Dienste
- 150 Themen-Portale in 18 Kategorien
- Das umfangreichste Internet-Verzeichnis deutscher Websites mit über 370.000 Einträgen
- Shopping-Portal mit vielen hunderttausend Angeboten in über 300 Online-Shops
- FreeMail mit über 5 Mio. registrierten Benutzern

### **WEB.DE ist der Innovator im deutschen Internet.**

- WEB.DE MillionenKlick
  - Erste kostenlose Internet-Lotterie
- WEB.DE FreeMail
  - Unified Messaging mit mehr als 5 Mio. registrierten Anwendern, 6 TB Message-Daten, 10 Oracle-Datenbanken mit bis zu 300 GB
- Internet-Telefonie
  - Ausbau des UMS zu Unified Communication (UC)
- Produkt-Suchmaschine
  - Mit mehr als 100.000 Produkten in mehr als 300 Shops
- WEB.DE SmartSurfer
  - Provider der Provider für Internetzugänge
- WEB.DE TrustCenter
  - Mit Zertifikaten für alle registrierten Anwender

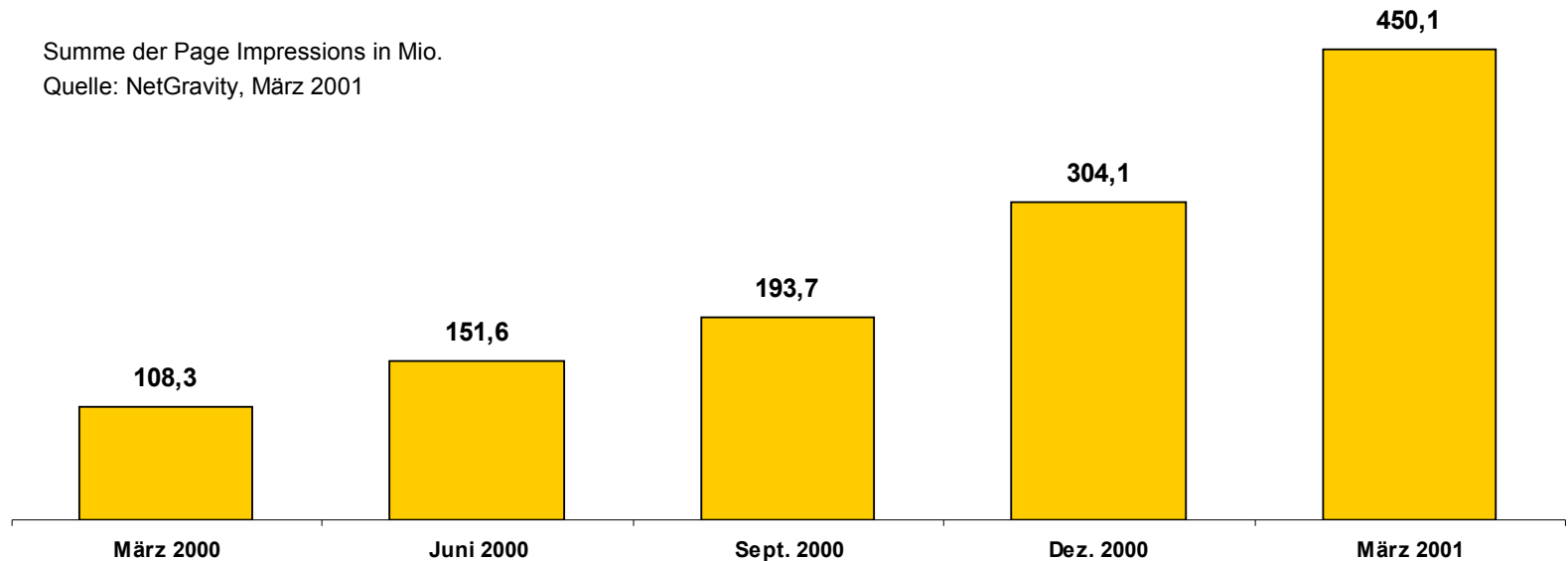
WEB.DE liefert Zahlen, die überzeugen:  
Über 450 Millionen Page Impressions pro Monat.

- Page Impressions im Vergleich:

Ende 03/2000: 108.272.336  
Ende 03/2001: 450.135.730

Wachstumsrate 03/2000 - 03/2001: **+ 316 %**

Summe der Page Impressions in Mio.  
Quelle: NetGravity, März 2001





Qualität setzt sich durch: WEB.DE wird im Wettbewerb immer häufiger bevorzugt.

- Die registrierten Anwender im Jahresvergleich:

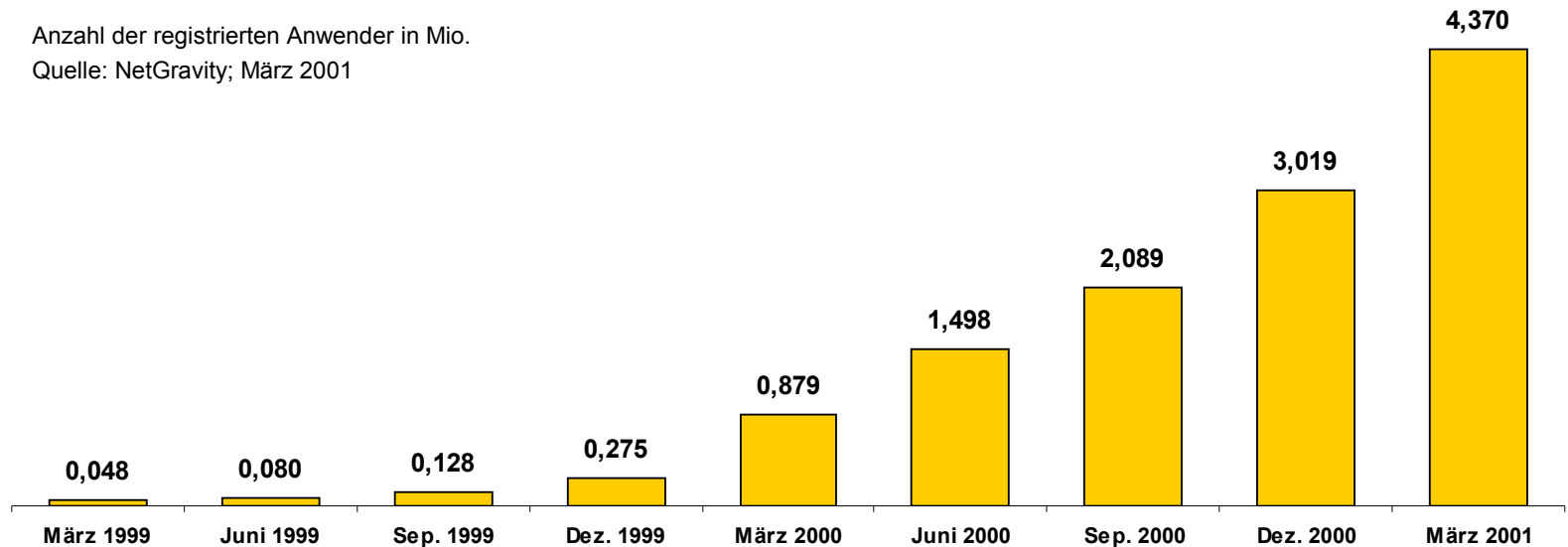
Ende 03/2000: 879.813

Ende 03/2001: 4.370.034

Wachstumsrate 03/2000 - 03/2001: + **397 %**

Anzahl der registrierten Anwender in Mio.

Quelle: NetGravity; März 2001



### Die Anfänge von WEB.DE und C.O.P.S.

- MacOS und IRIX waren die vorherrschenden Systeme aus der Multimedia-Vergangenheit der Firma
- Der Katalog und die komplexen Projekte des Systemhauses machten den Einsatz eines Redaktionssystems notwendig
- Für die damals schon innovativen Anwendungen zeigte sich der Einsatz einer existierenden Software als ungenügend
- Erste Version eines eigenen „Redaktionssystems“ auf der Basis von Netscape-Server, C-Plugin, Oracle und PL/SQL
- Hosting auf einer, später 2 Silicon Graphics R5000 Indies
- Anbindung zuerst über 64KBit, später 2MBit  
(-> Hosting war zunächst extern)



### Schlüsselfaktoren: Infrastruktur und Operabilität

- **Infrastrukturelle Voraussetzungen:**

- **Außenanbindungen**

Es lassen sich im gegenseitigen Wechsel 2 Phasen beobachten:

- **Phase der Dezentralisierung**

Bandbreite (-): Verteilung nahe zum Kunden oder Knoten  
Performance (-): Horizontale Verteilung auf mehrere Systeme

- **Phase der Zentralisierung**

Bandbreite (+): Aggregation in einem Rechenzentrum  
Performance (+): Skalierung über schnellere/größere Hardware

- **Bauliche Infrastruktur**

- Ausbaustand des Serverraumes (Klima, Strom & NEA, EMA, BMA)
- Ausbaustand der internen Infrastruktur (Netzwerk, Racks)

- **Organisationsstruktur:** Operativer Betrieb, Backup!

Unersetzbar: Nähe zum System in rasch wachsendem Umfeld.

## Architekturübersicht der ersten Release

- Layer 1: Einsatz der Oracle-Datenbank hat sich bewährt, jedoch eliminieren der PL/SQL-Teile
- Layer 2: Redaktionssystem implementiert als Plug-In zum Netscape-Server
  - Funktionalität jetzt komplett durch Plug-In realisiert
  - Datenbank verbleibt als intelligenter Datenspeicher, trägt jedoch keine Applikations-Logik mehr
  - Anbinden eigener Module an das Redaktionssystem erlaubt ein erstes Verteilen auf mehrere Systems  
Erstes relevantes Beispiel: die Suchserver für das Verzeichnis.
- Die gesamte Suite ist immer noch auf einer R5000 nebst eines NT-Suchservers lauffähig



## Skalierung die erste



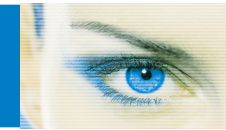
### Mit Hardware erschlagen

- Umzug von Datenbank und Web-Server auf eine R10k Doppelprozessormaschine (Origin 200)  
Vorteil: keine Änderungen an der Architektur  
Nachteil: kostenintensiv
- Weiterer Ausbau des Such-Backends
- Einsatz eines Anzeigen-Management-Systems
- Trennung von Homepage, Suche und Verzeichnis aus Gründen der Performance und Stabilität
- Dezentralisierung der Oracle-Datenbanken auf Windows NT aus Kostengründen (WEB.DE ist zu diesem Zeitpunkt profitabel)
- Zusätzliche Dienste werden geschaffen: Seite1, Chat, FreeMail, NetNews
- Aufschalten einer zweiten 2MBit-Leitung, Ausbau des Serverraums.



## Problematik: proprietäres vs. Open-Source-System

- Portierungsaufwand
  - Übertragung des Redaktionssystems auf den Apache Web-Server (IRIX)
  - Anpassung des Quellcodes an Linux/glibc
- Passende, stabile Hardware
  - Durchgängiger Einsatz von Raid-Controllern
  - Abstimmung zwischen Komponenten des Systems
  - Performance von Disk-I/O und Net-I/O Treibern
- Wartbarkeit und Support
  - Vorher: Silicon Graphics mit Supportvertrag und 24/7-Support (HW/SW)
  - Nachher: ?
- Der Wendepunkt
  - Compaq's Release des Linux-Treibers für den SmartArray-Controller
  - Integration des Treibers in SuSE Distribution und Live-Filesystem-CD



### Verteilung der Architektur

- Frontenden sind nun auf Linux-Basis kostengünstig skalierbar
- Performance-Gewinn erlaubt komplexere Logik in der Applikation
- Open-Source-Elemente sind nun einfach integrierbar  
z.B.: News-Server, Relay-Infrastruktur
  
- Abspalten der Benutzer-Datenbank
- Aufteilen der Datenbanken auf die Applikationen
- Zusammenfügen via DB-Links
  
- Neue Problematik: Platz  
Erhöhter Bedarf an Maschinen sprengt den alten Serverraum
- Umzug der 25-Mann-Firma ende 1998 in eine Etage der alten Fabrikhalle von Pfaff
- Bau eines neuen Serverraums mit Platz für 15 Racks

## Neue Herausforderungen für die IT

- Vorher: Lösungen angepasst auf eigene Bedürfnisse
- Jetzt:
  - Anforderungen seitens der Versicherungsgesellschaften
  - Audits durch unabh. Gesellschaften im Auftrag von Investoren
  - Audits im Rahmen von Beteiligungen und Kooperationen
- Allmähliches Entstehen von Standards in der IT-Infrastruktur
  - Sicherheitskonzept
  - Bauliche Normen gewinnen immer größere Bedeutung
- Unabdingbar sind nun:
  - Zugangskontrolle (ZKS), Einbruchmeldeanlage (EMA), Brandmeldeanlage (BMA), etliche Redundanzen, Netzersatzanlage (NEA) etc.
  - Fest definierte Workflows
  - Umfassende technische Dokumentation



### Immer mehr Nähe zur Öffentlichkeit

- Technische Probleme finden sich in Zukunft auf heise.de dokumentiert
- Einfluss der technischen Performanz auf den Kapitalmarkt
- Einfluss von Gerüchten auf den Kapitalmarkt
- Alle Zeichen deuten auf enorm schnelles Wachstum hin
  
- Schaffung von performanteren, ausfallsicheren Systemen
- Schaffung von Redundanzen
  - Mehrere Homepage-, Verzeichnis-, Freemail-Server
  - Standby-Datenbanken für wichtige Systeme
  - Zweite 34M-Leitung
  - Kein System mehr ohne RAID oder Ersatzsystem Out-of-the-Box
  
- Deja-Vu-Problem: Platz!
  - Planung, Bau und Inbetriebnahme eines neuen Rechenzentrums

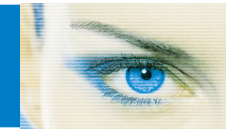


## WEB.DE setzt weiter auf die Nähe zum System

- Bau der ersten RZ-Zelle der Firma Lampertz
  - Fassungsvermögen: 30 Racks
  - 160KW Klimaleistung (1/7-Redundanz)
  - FM200-Löschanlage, Brandfrühsterkennung
  - gasdichte Schottung, faradayscher Käfig
  - Endausbau: 5 Zellen
- Ausbau der Stromeinspeisung auf 2000 kVA RZ + 1000 kVA Gebäude
  - gepuffert durch USV-Systeme (10 Min bei Vollast im Endausbau)
  - Netzersatz durch zwei Diesellaggregate mit je 1000 kVA
- Bau eines Sicherheits-Rechenzentrums
  - Schutzraum nach VDMA
  - Rechenzentrums-Zelle in Form eines begehbaren Tresors

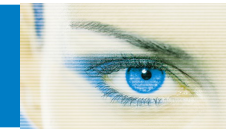
## Implementierung einer klassischen 3-Layer-Architektur

- Layer 1: Storage-Infrastruktur und Low-Level-Komponenten
  - Oracle kommt inzwischen auf Linux/Intel zum Einsatz
  - Zusätzliche „Datensenken“ (z.B. NAS, Filesysteme, externe Content-Quellen)
- Layer 2: Middleware auf Basis von CORBA
  - Einfache Implementierung von verteilten Systemen
  - Objekte aus der Middleware sind universell einsetzbar
  - Ablösen von Embedded-SQL durch OCI erhöht Skalierbarkeit
  - XML als Austauschformat ermöglicht einfache Integration
- Layer 3: Präsentationsschicht
  - Rendering via Stylesheets (XSLT, XALAN)
  - Integration von Python anstatt proprietärer Sprache erhöht Flexibilität und verringert Entwicklungsaufwand



## Neue Probleme verteilter Systeme

- Performance-Gewinn durch horizontales Skalieren der Teilkomponenten
  - Nachteil: Jede neu hinzugekommene Komponente erhöht die Wahrscheinlichkeit eines Ausfalls
  - Lösung im Frontend-Bereich: Linux Virtual Server (LVS)  
Bildung von Clustern auf Basis von Packet-Switching, Load-Balancer-Server selbst ist redundant
- Hoher administrativer Aufwand bei der Verteilung einzelner, zentraler Systeme
  - Einsatz von Filer-Clustern, Snapshots, Mirroring
  - Einsatz von Standby-Datenbanken und Real-Application-Cluster
  - Reduktion der Performance-Anforderung durch Verlagerung von Funktionalität in die Middle-Ware, Einsatz von khttpd
- Zusammenfassen einzelner Server in Segmente
  - Bildung von Zellen oder funktionalen Einheiten zur Vereinfachung des Sicherheitskonzepts und der Erhöhung der Performance



## Überblick und Handhabbarkeit

- Einsatz von Netsaint zur Überwachung der Systeme
  - Viele Erweiterungen zur Anpassung an die eigenen Bedürfnisse
  - Überwachung sowohl der Einzelkomponenten als auch aus Benutzer-Sicht
  - Datenbank-Unterstützung und Integration in ein Gesamtsystem
  - Überwachung der Funktion und Generieren von Meldungen und SMS
  - Überwachung der Performance zur Skalierungsplanung
  - Voraussetzung: Service Level Agreements (SLAs) mit sich selbst
- Administration verteilter Systeme: „Welche Software ist am besten?“
  - Grundproblematik ist nicht die Software, sondern die Klassifizierung und Typisierung der vorhandenen Systeme in ähnlich zu behandelnde Gruppen
  - Wichtig: Man muss automatisiert **Kommandos absetzen** können
  - Auf neuen Systemen kann von Anfang an ein „Hook“ implementiert werden
  - Alte Systeme müssen sowieso angefasst werden, egal ob mit einem Software-Paket gearbeitet wird, oder mit Scripting-Zugängen
  - Nicht überorganisieren, sonst drohen zu viele Ausnahmen



### Gut gegessen stark und gerüstet für die Zukunft

- Solide infrastrukturelle Basis
- Den Bedürfnissen angepasste Organisationsstruktur
- An vielen Stellen skalierbare Architektur
- Klarer Trend hin zum Einsatz von Open-Source-Systemen
- Linux trägt einen wesentlichen Teil zum Gesamtkonzept bei
- Ohne den Einsatz von Linux wäre ein so rasches Wachstum in gleicher Qualität nur unter enormem Einsatz von Kapital möglich gewesen.

Jennifer Stoll  
WEB.DE AG  
Amalienbadstr. 41  
76227 Karlsruhe  
[jst@web.de](mailto:jst@web.de)



**WEB.DE AG**



**WEB.DE**

Vielen Dank für Ihre Aufmerksamkeit!