Reengineering-Werkzeugen
rWeeien

Johannes Martin, University of Victoria
er Martin, echnische Universitt arstat

a onnef, Mai

Webangebot-Reengineering: Warum?

- zunehmendes Alter von Webangeboten
- mmer grer erdende atenmengen
- erausorderung r Anbeter
  - Wartung und Aktualisierung des nratinsestandes
  - ereitstellung einer ansreenden und gut ugnglien enutungssnittstelle
- Misserfolg -> Gefährdung des Überlebens

Web-Analysewerkzeuge kerzell

- großes Angebot, aber funktional beschränkt auf:
  - Webseiten-Prüfer
    - Syntaxanalyse von HTML Seiten
    - n von ennen
  - Websttisti-Weree
    - eten SeveLoe as
    - Statistien e tn eines eaneotes

Web-Analysewerkzeuge
rsungsberge

- Chung und Lee (WISE 2000)
  - UP und UML zur Darstellung von analysierten eangeoten
  - UMLDiagrae zur eransauliung von ernungs und erzeinisstrutur
- Ricca und Tonella (WSE 2000, ICSM 2000)
  - raen zur Darstellung von trutur und ntsteungsgesite von eangeoten
  - nalyse von eseiten it *Frames*

Warum das Rad zweimal erfinden?

- Software-Reengineering und Web-Reengineering ind erwandt
  - Unübersichtlichkeit der Daten wegen großer Datenenge elltet ebseiten
  - Daten ndern sich laend
- Frage: sind Software-Reengineering Werkzeuge feie genug u zu Reengineering on Weangeoten genutzt zu werden

Fallstudie: iialseeeeieeil

- Rigi besteht aus:
  - Parsern zur Sammlung von Informationen über zu analsierene Ssteme z Programmuelltete
  - raitor zur arstellung ieser Informationen
  - atenaustaus zisen Parsern un raitor gesiet ur gemeinsames ateiformat un Sema
- Alles, was wir brauchen, sind ein Schema und ein arser r ebseien

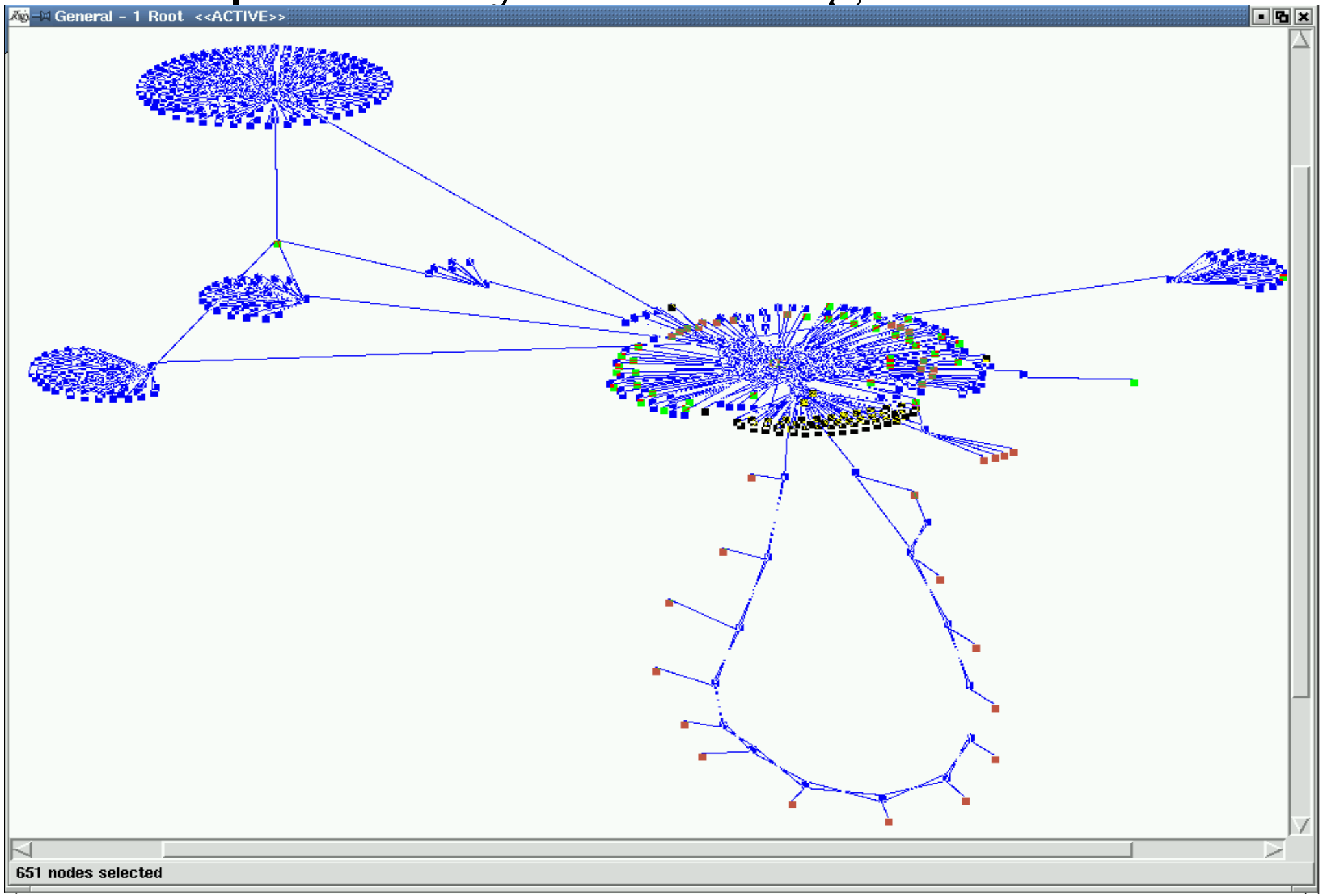
Rigis Webseiten-Schema

- Knoten:
  - HTML-Seiten, Bilddateien, Applets, Webserver, ...
- Kanten:
  - ernpnen isen den versiedenen nten
- ttte:
  - Ls der nten
  - MM-Tpen der nten
  - evtl. riseler

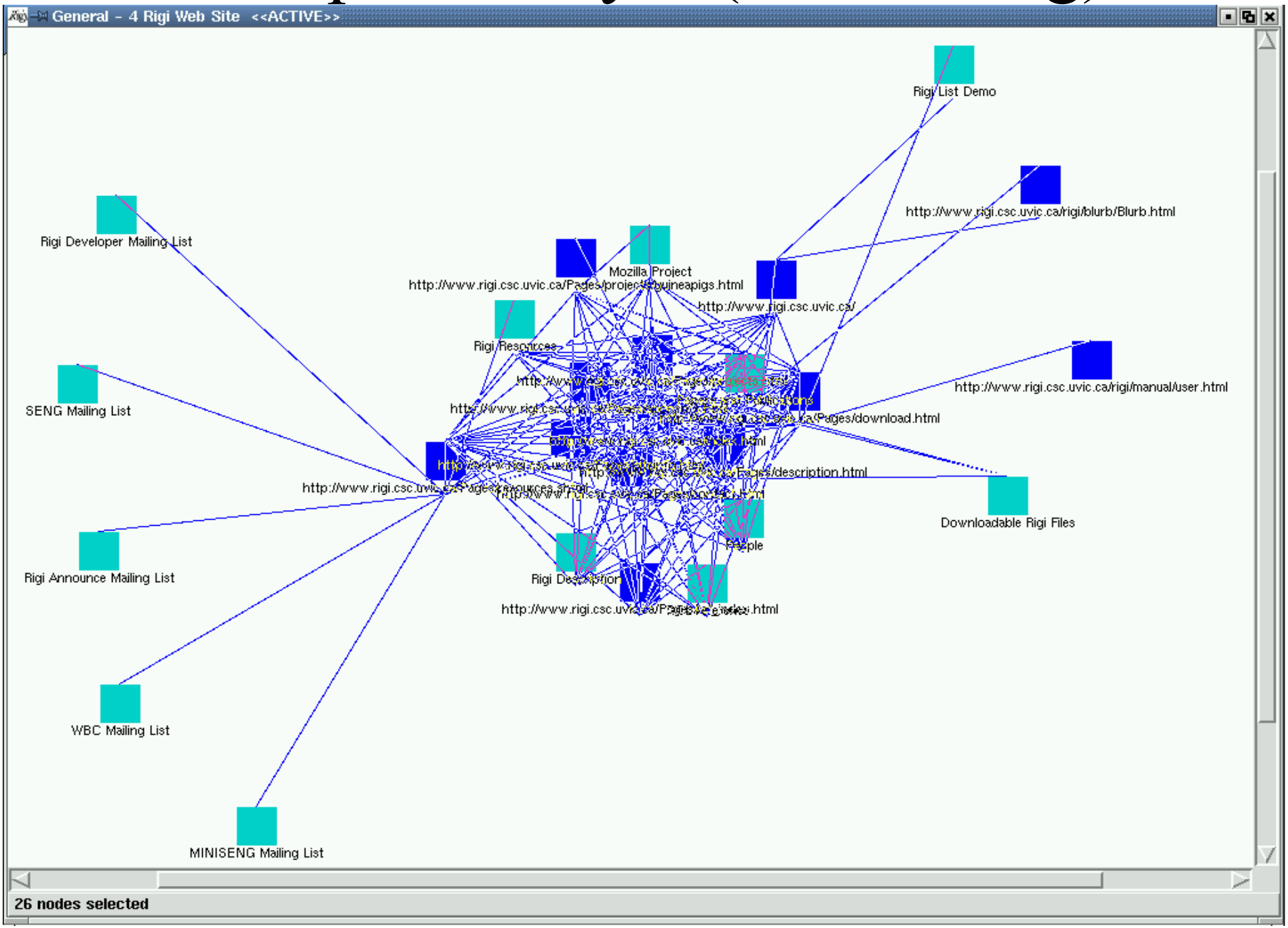
Rigis Webseiten-Parser

- Verfahren: beginnend mit einem URL
  - lade URL herunter
  - uhe nah L a
  - ehne ernunen u anderen URL au
  - unteruhe dee URL enau
- Optionen:
  - nhrnun au ette URL
  - Lterun der uhtee
- po ietin

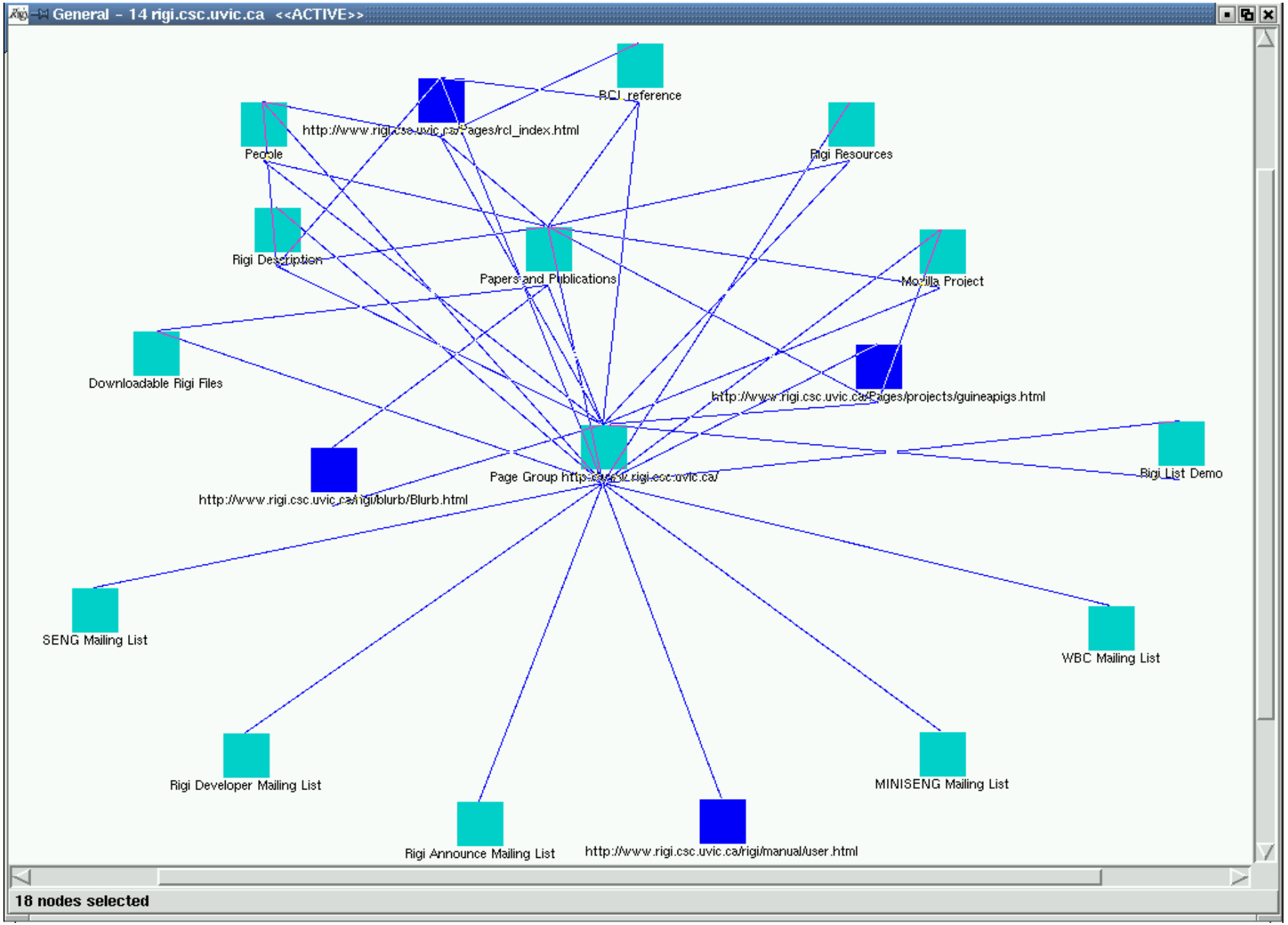
Beispiel-Analyse: www.rigi.csc.uvic.ca



Beispiel-Analyse (Fortsetzung)



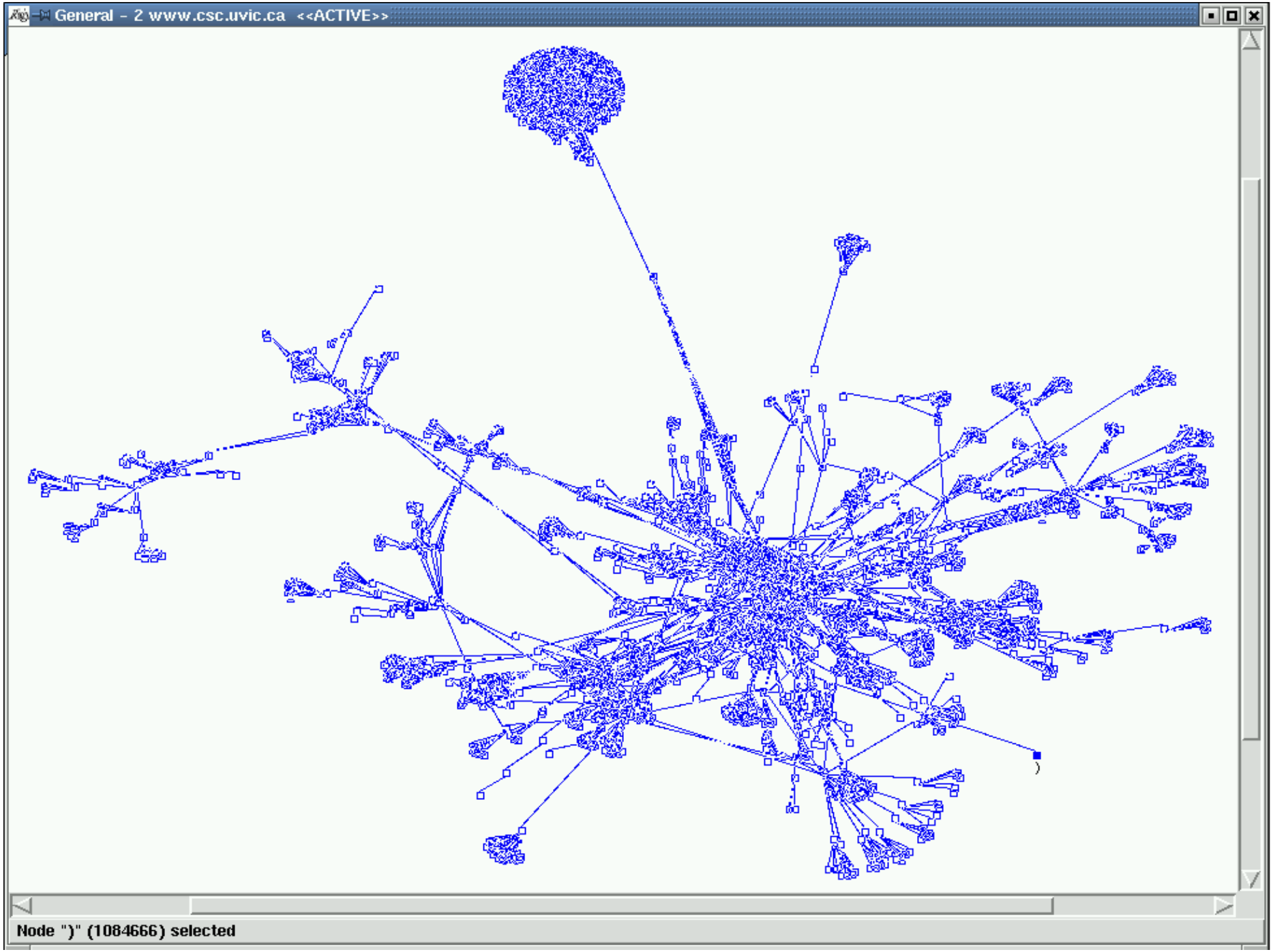
Beispiel-Analyse (Fortsetzung)



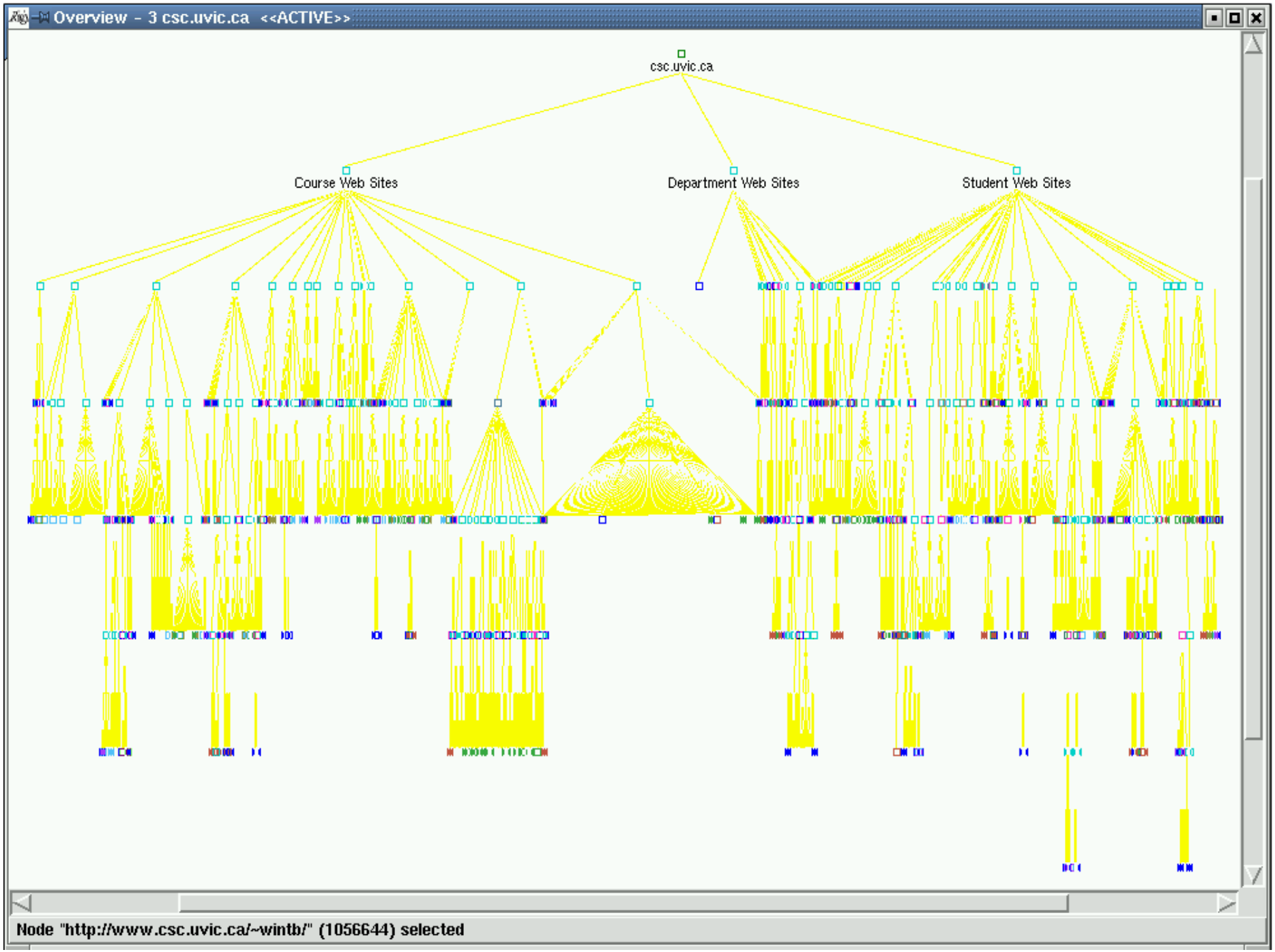
Zusammenfassung:
gsua

- zu analysieren wie ein kleineres uerrra
  - Layout-Algorithmus zum Gewinnen eines Überblicks
  - anuelle estrukturierung es Grahen
- Unterschied: Navigationsmenü und daraus resutierende hohe ernüung von entraen eseiten

Beispiel-Analyse: www.csc.uvic.ca



Beispiel-Analyse (Fortsetzung)



Zusammenfassung: www.csc.uvic.ca

- vergleichbar mit mittelgroßem Computerprogramm
  - Springlayout nicht mehr sehr hilfreich, lange Laufzeit, argestellte Strutur ist nicht ie logische Strutur
  - orhanenes issen er Strutur es Systems ann zur lieerung erenet eren hier erzeichnisstrutur
  - einige etails mssen on orneherein ignoriert eren Liraries in rogrammen, lipartSammlungen hier
  - manuelle anipulation es raphen ist nicht mehr sinnoll oer urchfhrar, utomatisierung ist ntig

Offene Probleme und Ausblick

- unser Parser spiegelt weder die exakte ateistruktur n die enuteransit wieder
  - keine Sonderbehandlung oder Erkennung von Server Side nlude u
  - keine Sonderbehandlung oder Erkennung von rae
- Parser-Performanceprobleme:
  - ava ah e einah u bearbeien edoh
    gib e Seiherroblee bei groen aenengen
  - lange auei da onload der ebeien nig
  - uirkungen au erieb der eberver
- nbeenonebserer-ofles
  - er nu da ngebo ann ie o und ie o