

anemone

ANALYSIERTE ARBEIT an experiment in organic information visualization



AUTOR Benjamin Fry
URL <http://acg.media.mit.edu/people/fry/anemone>

INHALT

- 1 Aufgabenstellung
- 2 Ansatz
- 3 Informationsdimensionen
- 4 Wahrnehmung
- 5 Resumé
- 6 Literatur

1 Aufgabenstellung

Die Arbeit Benjamin Frys setzt sich mit der Visualisierung von großen Mengen an dynamischen Informationen auseinander. Dynamisch nicht nur im Sinne von ständig dem System zufließenden neuen Daten, sondern insbesondere einer sich kontinuierlich ändernden Informationsstruktur, die sich aus veränderten Größenordnungen und Beziehungen untereinander ergibt. Erprobte Methoden zur Visualisierung statischer Informationen scheitern hier, da sie selbst inhärent statisch sind.

Gerade für das qualitative Verständnis hoch-komplexer Systeme mit vielen heterogenen Informationen würde sich aber eine Visualisierungsmethode als sehr hilfreich erweisen, um durch den Wust aus quantitativen Informationen hindurchzusehen.

Anemone nimmt sich exemplarisch dem Benutzungsmuster von Websites an. Es verarbeitet die auf jedem Webserver anfallenden *Access Logs* - Informationen über Seitenaufrufe mit Herkunft und Zeitstempel - und setzt diesen Informationen die Struktur der Website gegenüber. Das Resultat stellt also eine Verbindung aus der hierarchischen Struktur und der aktuellen Nutzung einer Site über die Zeit her.

```
ACCESS-LOG APACHE 127.0.0.1 - frank [10/Oct/2000:13:55:36 -0700] „GET /apache_pb.gif HTTP/1.0“ 200 2326 „http://www.example.com/start.html“ „Mozilla/4.08“
CLIENT IP USER DATE REQUEST STATUS BYTE REFERER USER AGENT
NAME CODE COUNT
```

2 Ansatz

Benjamin Fry entwickelte in seiner Master-Arbeit am Media Lab des MIT einen vollautomatischen Visualisierungsprozess, den er als *Organic Information Design* bezeichnet. Anstatt die zu verarbeitenden Informationen in eine vordefinierte Form zu bringen, entsteht die Form eigenständig durch einen mitgegebenen Satz von Regeln und Eigenschaften. Dies wird der Komplexität der avisierten, relativ strukturlosen Informationen gerechter und bleibt dennoch gut steuerbar.

Fry greift dabei auf die Forschung an dezentralisierten, adaptiven Systemen zurück, speziell den simplen Organismen. Er überträgt deren Eigenschaften wie Struktur, Aussehen, Anpassung, Verdauung, Stagnation, Wachstum, Reaktion, Bewegung und Fortpflanzung auf die Visualisierung und macht sie auf diese Weise zu einem lebenden Subjekt, das auf den sich stetig ändernden Informationsstrom reagieren kann.

Das Ergebnis des Prozesses ist eine Visualisierung, die die Erkenntnis der qualitativen Eigenschaften der absorbierten Daten erleichtern soll. Fry liefert eine konkrete Definition:

“Organic Information Visualization - a system that employs simulated organic properties in an interactive, visually refined environment to glean qualitative facts from large bodies of quantitative data generated by dynamic information sources.” [Fry, 19]

Oder laxer ausgedrückt:

“One way to look at it is that I make a software program that I can ‚feed‘ data to, and watch how the data is digested.” [Dodge, 2]

Die Simulation solcher organischen Systeme ist in der Informatik bereits gut erforscht und wird häufig eingesetzt. Fry nutzt das Konzept des zellulären Automaten, eines dezentralisierten, regelbasierten Systems. Die Stärke an solch verteilten Systemen ist, dass die Bedeutung des Ganzen aus der Interaktion der einzelnen Elemente als Epiphänomen hervorgeht [Hofstadter, 331]. So entsteht wie bei lebenden Organismen aus vielen simplen, wenig intelligenten Teilen eine äußerst komplexe, sich selbst organisierende Struktur.

Die Herausforderung besteht nun darin, dieses Modell auf eine Visualisierung zu übertragen. Als Basis dient die Darstellung mittels eines Graphen mit Knoten als kleinster Einheit (hier: Webseite) und verknüpfenden Kanten (hier: Hyperlink innerhalb der vorgesehenen Sitestruktur). Zusätzlich dienen Bögen der Darstellung von außerordentlichen Sprüngen. Der Stimulus ist das erwähnte *Access Log*.

Um das Verhalten des Systems zu steuern, definiert Fry u.a. folgende Regeln und Eigenschaften:

KNOTEN

WERTE
REGELN

Position (graphisch), Frequenz (Seitenaufrufe), Label (Seitenname)

Fortpflanzung - aufgerufene Seite wird als Knoten hinzugefügt

Verdauung - zusätzliche Aufrufe der Seite erhöht Frequenz-Wert

Aussehen - Durchmesser wächst entsprechend der Frequenz,
bei Klick Anzeige des Seitennamens

Bewegung - Abstand zum Vaterknoten halten, Abstand zum Nachbarn halten

Stagnation - Grenzwert für visuelle Ausdehnung

	KANTEN
WERTE	von, nach (Sprünge innerhalb der vorgesehenen Sitestruktur)
REGELN	<p><i>Fortpflanzung</i> - bei Sprung zu neuer Seite Kante zu neuem Knoten anlegen</p> <p><i>Wachstum</i> - länger nicht mehr besuchte Seiten verwittern (Senken der Frequenz) und werden schließlich entfernt</p> <p><i>Bewegung</i> - Gruppierung von verbundenen Seiten</p> <p><i>Aussehen</i> - Dicke entsprechend der Durchmesser der verbundenen Knoten</p>

Die Komposition wird durch das Zusammenspiel von Wachstum und Bewegung zum Leben erweckt. Zu Beginn wandert die Figur in einer eher wirren Weise über den Bildschirm, stabilisiert sich aber mit der Verdauung immer neuer Daten zusehends.

3 Informationsdimensionen

Anemone erreicht trotz seines minimalistischen Aussehens eine erstaunliche Informationsdichte.

FORM - Die Form repräsentiert die zuletzt besuchte Teilstruktur der analysierten Site. Sie muss nicht zusammenhängend sein, da Zwischenstücke evtl. längere Zeit nicht besucht wurden und verwittert sind. Zusammenhängende Strukturen sind hierarchisch aufgebaut entsprechend der Website-Hierarchie.

DETAILFORMEN - Dickere Knoten zeigen ein höheres Gewicht der von Ihnen repräsentierten Seiten an. Dies wird unterstützt durch die sich der Form durch Spreizung anpassenden Kanten, die ansonsten geradlinig verlaufen. Eine andere Art von Verbindung wird durch die in der Dicke unveränderlichen bogenförmigen Linien angezeigt. Sie repräsentieren Sprünge innerhalb der Website, die nicht entsprechend der eigentlichen Hierarchie geschehen.

ABSTÄNDE - Unterschiedliche Knotenabstände und damit Kantenlängen ergeben sich automatisch durch die Regeln zum Abstandhalten. Somit bedeutet eine größere Kantenlänge eine höhere Anzahl von Aufrufen.

DICHTE/STÄRKE - Je dichter der weiße Anteil in einem Bereich, desto größer ist dort das Besucheraufkommen. Eine solche gesteigerte Farbdichte entsteht entweder durch eine größere Strichstärke, die eine viel frequentierte Seite anzeigt, oder durch das kollektive Wachsen einer von vielen Besuchern aufgerufenen Teilstruktur.

RAUM/POSITION - Die zweidimensionale Grafik verteilt sich selbst übersichtlich über den verfügbaren Raum. Der genauen Position der einzelnen Elemente ist keine Information zugewiesen und entsteht vielmehr durch Zufall und durch Einwirkung des Benutzers.

FARBE - Das tiefe Rot für nicht informationsrelevante Fläche bietet einen guten und angenehmen Kontrast zum Weiß der Struktur. Zur besseren Unterscheidung der außerordentlichen Sprünge sind diese außer durch die Form durch eine orange Tönung gekennzeichnet. Ansonsten ist Farbe kein Informationsträger. Auf diese Weise erhält die Grafik jedoch eine angenehm Ästhetik und ist schnell zu überblicken.

BEWEGUNG - Bewegung ist hier der bedeutendste Informationsträger. In erster Linie bedeutet sie eine Veränderung des Verhältnisses einer Menge von Daten innerhalb der Komposition, wie Wachstum oder Neuentstehung. Direkt deutet sie Aktivität an und zeigt so die Lebendigkeit des Organismus. Außerdem aber erregt sie Aufmerksamkeit, da das menschliche visuelle System stark von Bewegung angezogen wird. So ist jederzeit ersichtlich, in welchen Bereichen Änderungen auftreten. Eine überflüssige und ablenkende Bewegung entsteht, wenn neue Knoten geboren werden. Sie entstehen stets im Zentrum der Grafik und springen dann an den Punkt innerhalb der Struktur.

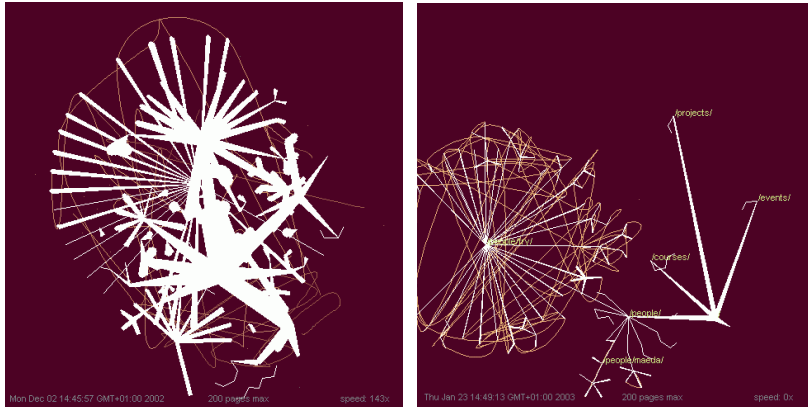
ZEIT - Die durch die Bewegung implizierte zeitliche Veränderung zeigt die Entwicklung der Struktur. Der Benutzer vergleicht ständig den aktuellen Stand ganz unterbewusst mit vorherigen Zuständen und mit abgespeicherten mentalen Modellen und erlangt so Rückschlüsse über Trends.

INTERAKTION - Alle Knoten sind sensitiv für Benutzeraktionen und liefern auf Klick ihren Bezeichner. Desweiteren ist es möglich, durch das ‚Festpinnen‘ einzelner Knoten, deren Entwicklung im Kontext der restlichen Struktur zu beobachten. Dies erlaubt eine aktive, explorierende Betrachtung der Daten.

Die Internet-Version von *Anemone* bietet leider nicht alle Features des Originals. Insbesondere ist die sichtbare Fläche und die Auflösung sehr begrenzt, so dass die Struktur bei viel Datenaufkommen (*canned version*) sehr schnell unübersichtlich wirkt und nicht mehr viel Information bereitstellt. Eventuell ist dies aber auch ein prinzipielles Problem. Die Live-Version ist in ihrer

Entwicklung aufgrund der geringen Aktivität deutlich besser zu verfolgen, insbesondere mit den Möglichkeiten der *Direct Manipulation*. Die teilweise lose herumliegenden Teilstrukturen sind auch visuell verwirrend, besitzen jedoch einen Sinn.

ÜBERINFORMATION (L)
DIRECT MANIPULATION (R)



4 Wahrnehmung und Interpretation

Durch die behutsam eingestellten Organismen-Regeln und das austarierte Gleichgewicht von Informationsdichte und Übersichtlichkeit bleibt *Anemone* ständig sehr gut lesbar. Und das, obwohl weder Zahlen und Zeichen, noch Farben zum Einsatz kommen, um die Information zu charakterisieren. So sind keine den Verarbeitungsprozess störenden aktiven, bewußten Interpretationen nötig. Vielmehr geschieht ein Großteil der Verarbeitung vollkommen unterbewußt, da die Visualisierung auf automatische Prozesse seitens des menschlichen Betrachters setzt. Zum Verständnis ist jedoch ein vorbereitendes Auseinandersetzen mit dem Visualisierungssystem notwendig. Nur auf diese Weise werden auch die feinen Nuancen der Strukturen verständlich. Mit diesem Vorwissen reicht dann aber ein kurzer Blick auf den aktuellen Zustand für einen unmittelbaren Eindruck über die aktuelle Lage und ein zweiter für momentane Trends. Die wirkungsvollen Methoden zur Erregung von Aufmerksamkeit lassen das System auch gut als Hintergrundanwendung eignen, die dann noch immer effektiv ist.

Die auf diese Weise passive Art der Betrachtung lässt sich bei Bedarf in eine aktive, explorierende Weise durch Interaktion umschalten. Nach dem Prinzip der *Direct Manipulation* reagiert die Visualisierung unmittelbar auf direkte Eingriffe in ihre Struktur durch den Benutzer ohne den Umweg über kryptische Befehle und Neuberechnung. Sie liefert hier erst spezifische Information (Seitenamen) und auf einfache Weise die Möglichkeit zur Analyse der Struktur. Zudem bereitet das Spielen mit dem Organismus durchaus Vergnügen und lädt zum Experimentieren ein.

Zuletzt bleibt diese Art der Visualisierung aber eine rein qualitative, die lediglich etwas über den relativen Zustand der Besucherströme aussagt und nur Rückschlüsse auf breite Trends zulässt. Auch wenn mit der Zeit Lerneffekte durch das aufmerksame Betrachten entstehen, die auch feinere Bestimmungen zulassen. Eine quantitativ stärkere Aussage wäre durch zusätzliche Tags oder ähnliches denkbar, würde jedoch die qualitative Aussage mindern. Als explorierendes Analysemittel, das bei Bedarf hinzugeschaltet werden kann, wäre es aber sicherlich eine Verbesserung.

5 Resumé

Die Idee des organischen Informationsdesigns ist eine äußerst interessante Methode zur Visualisierung komplexer, dynamischer Strukturen und wurde von Fry auch auf andere solche Strukturen angewandt. Für den frühen Stand der Forschung zeigt diese Methode schon sehr beeindruckende Resultate, die sich auch mit den bekannten Thesen über menschliche Informationsrezeption messen lassen können.

Die zunächst komplex wirkende Form der Visualisierung lässt sich nach kurzer Erläuterung ausgezeichnet verstehen und scheint der menschlichen internen Repräsentation von Informationen sehr zu schmeicheln. Der Betrachter erzielt bereits durch einen kurzen Blick ein tiefes Verständnis eines sehr komplexen Verhältnisses zwischen den Elementen, wie es ihm durch herkömmliche Methoden kaum möglich ist. Erst diese Komplexität der zugrundeliegenden Datenstruktur rechtfertigt auch die organische Visualisierung.

Die dargestellten Zusammenhänge haben zwangsweise immer eine Bedeutung, da sie erst aus ihr Entstehen. Es ist also das Pferd von hinten aufgezäumt: Die Daten erstellen ihre visuelle

Struktur selbst. So ist kein Pixel sinnlos und Anomalien werden nicht ausgeblendet sondern bekommen sogar eine wichtige Bedeutung.

Die Tatsache, dass dieses Werk bisher hauptsächlich in Kreisen der Medienkunst Beachtung gefunden hat und keinen ökonomischen Einsatz erfahren hat, ist vielleicht auf die rein qualitative Wirkung und die besondere Ästhetik zurückzuführen. Die in Anemone behandelte Thematik, nämlich die Gegenüberstellung von Site-Struktur und tatsächlichem Benutzerstrom ist eigentlich für Website-Architekten eine durchaus interessante Analysemöglichkeit. Sie erfahren, welche Bereiche zu welcher Zeit auf welche Weise mit welcher Intensität besucht werden und können daraus gegebenenfalls Änderungen in der Site-Architektur ableiten. Leider ist es aber kaum möglich, quantitative Aussagen zu treffen, zum Beispiel welche Seite in den letzten fünf Minuten wie viele Besucher empfangen hat und eignet sich daher nur bedingt zur Entscheidungsfindung [Tuftte, 27].

So steht der künstlerische und der forschende Aspekt zur Zeit noch im Mittelpunkt dieser Betrachtung. Dies soll sich laut Fry aber ändern, indem mehr quantitative Analysemöglichkeiten integriert werden sollen, weniger Zufälligkeit beim Aufbau der Struktur eine Rolle spielen soll und vor allem andere komplexe Themen angegangen werden sollen: Genforschung, Ökonomie, Audiodaten, Verkehrsfluss, Spieltheorie oder Software-Architektur.

6 Literatur

Apache Documentation. *Apache HTTP Server Log Files*. Apache Foundation, 2002.

[HTTP://HTTPD.APACHE.ORG/DOCS/LOGS.HTML](http://httpd.apache.org/docs/logs.html)

Dodge, Martin. *Anemone: Mapping How People Use A Website*. Map of the Month. Mappa Mundi, Juni 2001.

[HTTP://MAPPA.MUNDI.NET/MAPS/MAPS_022](http://mappa.mundi.net/maps/maps_022)

Fry, Benjamin J. *Organic Information Design*. M.S. Thesis. Massachusetts Institute of Technology, Program in Media Arts and Sciences, 2000.

[HTTP://ACG.MEDIA.MIT.EDU/PEOPLE/FRY/THESIS/](http://acg.media.mit.edu/people/fry/thesis/)

Hofstadter, Douglas R. *Gödel, Escher, Bach*. New York: Basic Books, 1979.

[HTTP://WWW.PSYCH.INDIANA.EDU/PEOPLE/HOME/PAGES/HOFSTADTER.HTML](http://www.psych.indiana.edu/people/homepages/hofstadter.html)

Tuftte, Edward R. *Visual Explanations*. Cheshire, Connecticut: Graphics Press, 1997.

[HTTP://WWW.EDWARDTUFTTE.COM/TUFTTE/BOOKS_VISEX](http://www.edwardtuftte.com/tuftte/books_visex)