

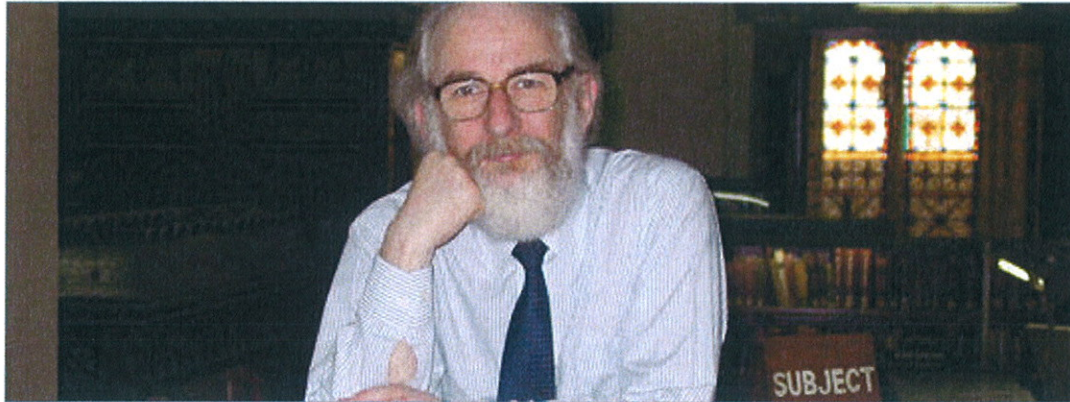
[« zurück zur Übersicht <URL: /wuv de infocenter/specials>](http://wuv.de/infocenter/specials)

[« zurückvorwärts » <URL: /wuv de infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)/1>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)/1)

[1](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)/1) | [2 <URL: /wuv de infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)/1>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)/1) | [3 <URL: /wuv de infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)/2>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)/2)

[/wuv de infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)/2>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)/2)

[Drucken <URL: javascript:window.print\(\);>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)/2) | [Versenden <URL: /content/tipafriend/136819>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)/2) | [An die Redaktion <URL: mailto:leif.pellikan%40wuv.de>](mailto:leif.pellikan%40wuv.de)



Sprachguru David Crystal

Sprachguru David Crystal zu Targeting

veröffentlicht am 25.06.2009 • [Specials <URL: /wuv de infocenter/specials>](http://wuv.de/infocenter/specials) • Specials

Sir David Crystal, Jahrgang 1941, ist einer der anerkanntesten Sprachwissenschaftler in Großbritannien. Zu seinen zahlreichen Veröffentlichungen zählt The Cambridge Encyclopedia of Language, die er als Chefredakteur mehrere Jahre betreute. Hieraus entstand das Kategorisierungssystem Sense-Engine für Sprache. Das Tool hat er 1996 an eine holländische Firma verkauft, die es für eine Suchmaschine nutzen wollte, aber eine typische Dot-com-Pleite erlitt. Er kaufte das System zurück und fand 2006 mit dem Online-Vermarkter Ad Pepper einen neuen Abnehmer, der das System heute unter dem Label iSense semantisches Targeting anbietet. Heute berät Crystal zahlreichen Firmen und Institutionen, darunter auch die BBC, Hüterin der Englischen Sprache. 2005 wurde er von der Queen mit dem Orden Order of the British Empire ausgezeichnet.

W&V Von automatisiertem, kontext-sensitivem Targeting versprechen sich Online-Werber enorm viel. Aber funktionieren solche Systeme überhaupt, reichen Keywords aus, um einen Text logisch zu erfassen?

Crystal Die Bedeutung von Worten können im Moment nur Menschen verstehen - daher brauchen Sie Linguisten, die den Computern die Bedeutung erklären – diesen semantischen Ansatz habe ich verfolgt. Aber ich habe nie versucht Computern Sprache beizubringen - das versuchen andere Anbieter mit ihren Kontext-Ansätzen.

W&V Das müssen Sie erklären.

Crystal 1986 arbeitete ich mit 350 Spezialisten an einer große Cambridge-Enzyklopädie. Um dieses Projekt zu steuern, musste ich eine Klassifikationssystem mit Kategorien entwickeln. Etwa zehn Jahre später wurde das Projekt eingestampft und die holländische Firma AND kaufte das Konzept 1996. Sie hatte die Idee diese Klassifikation auf das Web zu übertragen. Damals entstanden bereits Suchmaschinen wie Lycos, Excite und Altavista. Im Zuge dessen haben wir das Categoriesystem von etwa 1000 auf 1500 Bereiche erweitert – denken Sie an all die Begriffe, die im Web wichtig sind wie Sex, Autos oder Urlaub, die in der Enzyklopädie nicht vorkamen. Bis 1999 funktionierte das System ganz gut, schaffte es aber vor dem Platzen der Dot-com-Blase nicht auf den Markt. Ich haben dennoch weitergemacht, und mit dem Unternehmen Crystal Reference einen Weg gesucht, das System zu nutzen und auch um Geld zu verdienen.

W&V Aber Werbung war sicher nicht ihr erster Gedanke, oder?

Crystal Wir hatten zunächst keine Idee, wer sich dafür interessieren würde. 2002 waren wir in Silicon Valley und die Suchmaschinen-Anbieter zeigten sich interessiert. Sie fanden den Ansatz toll, aber er lies sich schlicht nicht mit den eigenen Systemen verbinden. Google nutzt beispielsweise einen Algorithmus, der alle möglichen Informationen verarbeitet, um einer Seite einen Page-Rank zu geben – ein rein

technischer Ansatz, der Metatags, die Überschriften, die Zahl der Links und Begriffe bewertet. Doch diese Systeme haben je nach Suchanfrage eine 40 oder 50 prozentige Treffergenauigkeit. Auf den ersten Seiten der Suchergebnisse mag das mehr sein, aber wenn Sie sich ein- oder zweitausend Treffer ansehen, finden sie in der Hälfte aller Fälle nichts Sinnvolles. Ein Suchmaschinenanbieter sagte damals sogar, „wir wollen gar nicht 100 Prozent genaue Treffer liefern“ Eine gewisse Menge an irrelevanten Treffern sorgt schlicht für weitere Suchanfragen und weitere Werbekontakte.

W&V Wieso kamen Sie auf die Idee, Werber könnten sich für ihr System interessieren?

Crystal Zunächst kamen wir selbst nicht auf den Gedanken. Der Online-Vermarkter Ad Pepper Media kam auf uns zu. Wir erklärten unser System damals anhand von folgendem Beispiel: Ein Interview mit einem Tennisspieler kurz vor dem Wimbledon-Turnier. Unser System erklärt uns, der Text dreht sich um Tennis und BMW. Wie passt das zusammen? Ganz einfach: Der Reporter hatte am Ende keine Tennisfragen mehr gestellt. Er fragte nach der Autosammlung und dem Auto mit dem der Spieler kam – einer seiner vielen BMWs. Wenn sie den Inhalt des Textes auf die Überschrift und die ersten Zeilen reduzieren, fehlt Ihnen der Rest an Informationen. Und Ad Pepper erklärte mir daraufhin, dass dies durchaus ein sehr brauchbares Werbeumfeld für BMW sein könnte, da der Münchner Autohersteller durchaus im Tenniseumfeldern mit Sponsoring aktiv ist. Und umgekehrt machte uns Ad Pepper klar, es gibt Firmen, die in bestimmten Umfeldern überhaupt nicht erscheinen wollen. Stellen Sie sich eine Anzeige für einen besonders günstigen Flug nach Brasilien vor, der in einem Bericht über die Air-France-Maschine steht, die eben abgestürzt ist. Es geht also darum, die logischen Inhalte zu identifizieren bevor etwas wie ein Flugzeugabsturz passiert. Ob das grundsätzlich geht, war die Frage von Ad Pepper? Meine Antwort lautet klar: ja.

W&V Wie gut ist ihr System – jeden Zusammenhang können auch Sie nicht erklären?

Crystal Auch wenn Ad Peppers iSense-Produkt nicht perfekt ist, können wir in etwa 97 Prozent aller Seiten semantisch korrekt bewerten. Faktisch fragen wir hierbei ab, in welche der 3000 Kategorien eine Seite fällt. Das einzige Problem ist, wenn zuwenig Text und zu wenige Metatags und Bildunterschriften auf einer Seite stehen. Aber trotz aller Einschränkungen glaube ich, dass es viel besser funktioniert, als andere Systeme.

W&V Aber die Werbung ist doch ziemlich weit von der Linguistik entfernt.

Crystal Das schöne, bislang war Linguistik ein fast rein wissenschaftliches Problem. Aber plötzlich gibt es mit der Online-Werbung ein Feld, in dem die Sprachwissenschaft eine immense Hilfe ist. Der Unterschied der Welten ist durchaus bemerkenswert – ich auf der einen Seite mit Krawatte, Brille und Anzug und auf der anderen Seite junge Menschen – zumeist 50 Jahre jünger, die immense Geldbeträge in der Werbung bewegen. Aber deren Interesse und die Begeisterung für die Möglichkeiten der Linguistik haben mich durchaus beeindruckt.

In Teil 2 erklärt David Crystal wie semantisches Targeting funktioniert und warum Web 3.0- das semantische Web wohl noch lange eine Fantasie bleiben wird.

« zurück **vorwärts** » <URL: /wuv de infocenter/specials/sprachguru david crystal zu targeting(offset)/1>
1 | 2 <URL: /wuv de infocenter/specials/sprachguru david crystal zu targeting(offset)/1> | 3 <URL: /wuv de infocenter/specials/sprachguru david crystal zu targeting(offset)/2>

[Drucken](#) <URL: javascript:window.print();> | [Versenden](#) <URL: /content/tipafriend/136819> | [An die Redaktion](#) <URL: mailto:leif.pellikan%40wuv.de>

[« zurück zur Übersicht](#) <URL: /wuv de infocenter/specials>



Leif
Pellikan

Redakteur

[An den Redakteur](mailto:leif.pellikan%40wuv.de) <URL: mailto:leif.pellikan%40wuv.de>

Social Bookmarking

 <URL: http://www.mister-wong.de>  <URL: http://de-de.facebook.com>  <URL: http://www.google.com>  <URL: http://www.yahoo.com>  <URL: http://www.webnews.de>  <URL: http://www.oneview.de>  <URL: http://linkarena.com>

[« zurück zur Übersicht <URL: /wuv.de/infocenter/specials/>](http://wuv.de/infocenter/specials/)

[« zurück <URL: wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting_vorwaerts/> » <URL: wuv.de/infocenter](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting_vorwaerts/)

[1 <URL: wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting_1/> | 2 | 3 <URL: wuv.de/infocenter/specials](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting_1/)

[sprachguru_david_crystal_zu_targeting_offset_2/>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting_offset_2/)

[specials/sprachguru_david_crystal_zu_targeting_offset_2/>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting_offset_2/)

[Drucken <URL: javascript:window.print\(\);> | Versenden <URL: /content/tipafriend/136819/> | An die Redaktion <URL: mailto:leif.pellikan%40wuv.de/>](http://wuv.de/infocenter/specials/sprachguru_david_crystal_zu_targeting_offset_2/)

Sprachguru David Crystal zu Targeting - Seite 2

veröffentlicht am 25.06.2009 • [Specials <URL: /wuv.de/infocenter/specials/>](http://wuv.de/infocenter/specials/) • Specials

Teil 2 Interview mit Sprachwissenschaftler Sir David Crystal

W&V Werber versprechen sich durch Targeting-Systeme eine höhere Effektivität. beispielsweise arbeitet Behavioral Targeting mit mathematischen Wahrscheinlichkeiten und errechnet ob sich ein User womöglich für ein Thema oder Produkt interessiert...

Crystal ...Ich weiß nicht allzu viel über Werbung. Aber sicher ist, ein mathematischer Ansatz kann nicht funktionieren. Zumindest heute nicht. Der semantische Ansatz konkurriert aber auch nicht mit Behavioral Targeting, er ergänzt ihn. Allerdings haben sie weit aus weniger Probleme, es gibt hierbei keinerlei Datenschutzbedenken. Alle Targeting-Varianten, die auf User-Profilen basieren könnten früher oder später aus rechtlichen Gründen zusammenbrechen, was ich auch befürchte. Daneben verfügen Behavioral-Targeting-Systeme nur über ein begrenztes Kategoriensystem. Sie können einfach nicht alle Unterseiten oder User-Interessen vernünftig kategorisieren, das kann wie gesagt heute kein Computer leisten.

W&V Ist der Bewertungsprozess nicht enorm aufwändig?

Crystal Sie können Seiten vorab kategorisieren, aber natürlich auch im Live-Modus. Wenn eine Person eine Seite ansurft, dauert der Prozess den Inhalt auf 3000 Kategorien abzuklopfen, nur Bruchteile einer Sekunde

W&V Aber wie erkennen Sie den Zusammenhang von Texten?

Crystal Nun wir haben jedem Begriff weitere Begriffe zugeordnet – und zwar einfach nur mit menschlicher Hirnleistung. Dafür haben wir mit 40 Mitarbeitern schlicht Wörterbücher durchgeackert und sehr viel nachgedacht, auch über Marken, die zu den Begriffen passen. Das hat drei Jahre gedauert. Wir wissen also heute, welche Begriffe zu „Flug“ und „Brasilien“ passen und welche nicht und welche positiv und welche negativ besetzt sind. Aber das reicht nicht. Sie müssen diese Begriffszusammenhänge permanent aktualisieren. Die Welt dreht sich weiter. Denken Sie nur, wofür etwa „Irak“ 1999 stand – das waren Öl, eine alte Hochkultur und Urlaub. Inzwischen steht „Irak“ für Krieg, Terror und Selbstmordattentäter.

W&V Das erklärt aber noch nicht die Zusammenhänge innerhalb eines Textes, um das System für Werbeschaltung brauchbar zu machen?

Crystal Begriffen müssen Werte zugeordnet werden. Manche Wörter sind schlicht wichtiger als andere. Nehmen Sie beispielsweise das Wort „Quarterback“. Der Begriff spielt in keiner Kategorie eine Rolle – außer es dreht sich um American Football. In einem solchen Zusammenhang erhält „Quarterback“ 100 Prozent Relevanz. Ein Computer kann diese Bewertung unmöglich leisten, das müssen Menschen tun.

W&V Ist Web 3.0 – das sogenannte semantische Web – also eine Fantasie?

Crystal Der Begriff Semantik stammt aus der Linguistik. Hinter „semantisches Web“ steht eine viel breitere, soziologische und philosophische Sichtweise: Sie beinhaltet Wissen, logische Beziehungen und mehr. Aber alles bleibt Teil des semantischen Problems. Einfache Texte, etwa einen Reiseplan kann ein Computer durchaus heute schon verstehen. Aber bis ein Rechner über einen Witz lachen kann, müssen wir noch etwas warten.

In Teil 3 erklärt David Crystal warum Computer nicht lachen können und ob die Semantik auf

[« zurück zur Übersicht <URL: /www.de.infocenter/specials>](http://www.de.infocenter/specials)

[« zurück <URL: /www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)1 >vorwärts »](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)1)

[1 <URL: /www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting >](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting) | [2 <URL: /www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)1 >](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)1) | [3](#)

[Drucken <URL: javascript:window.print\(\); >](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting) | [Versenden <URL: /content/tipafriend/136819 >](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting) | [An die Redaktion <URL: mailto:leif.pellikan%40wuv.de >](mailto:leif.pellikan%40wuv.de)

Sprachguru David Crystal zu Targeting - Seite 3

veröffentlicht am 25.06.2009 • [Specials <URL: /www.de.infocenter/specials>](http://www.de.infocenter/specials) • [Specials](#)

W&V Die Online-Werber versprechen sich viel von der Video-Werbung. Ist semantisches Targeting auch auf bewegte Bilder übertragbar?

Crystal Es ist durchaus möglich, aber nicht einfach. Vor fünf Jahren hätte ich gesagt, die Spracherkennung wird dazu niemals in der Lage sein. Aber die Fortschritte waren enorm. Viele dieser Spracherkennungssysteme können heute durchaus auch Dialekte verstehen. Die Genauigkeit ist heute irgendwo im Bereich von 80 Prozent.

W&V Aber die Menschen machen doch beim Sprechen viel mehr Fehler.

Crystal Deshalb würde auch ein Keyword-Ansatz nicht funktionieren. Semantische Ansätze können mit gelegentlichen Fehlern wesentlich besser umgehen. Sie gehen schlichtweg in der Gesamtanalyse unter, weil sie auf alle Wörter achten.

W&V Fehler und Umgangssprache zeichnen auch Social-Media-Angebote aus. Dort ist meist sehr wenig Text – etwa Twitter, dort stehen nur 140 Zeichen zur Verfügung.

Crystal Ich habe mir Twitter angesehen. Und mit 140 Zeichen kann ein User sehr viel sagen. Er gibt sich auch sehr viel Mühe, verständlich zu sein. Dann reichen auch zehn bestimmte Begriffe um den Inhalt zu verstehen. Das Problem ist vielmehr, dass sich die Themen sehr schnell ändern. Was sie im Moment noch nicht schaffen, ist einen Blog jedes Mal zu qualifizieren, wenn ein neuer Eintrag kommt. Das sprengt im Moment schlichtweg die Kapazitäten.

W&V Wie hoch ist die Sprachbarriere – oder müssen künftig alle Menschen, die Zugang zum Wissen und zum semantischen Web haben wollen, Englisch lernen?

Crystal Genau an diesem Punkt könnte die Idee des gesamten semantischen Webs scheitern. Jeder Web 3.0-Ansatz stolpert immer über die Sprache. Es ist nicht möglich, die Denkweise und das Wissen der Menschen von einer Sprache in eine andere zu übertragen. Führen Sie sich schlicht vor Augen, dass sich schon britisches und amerikanisches Englisch deutlich unterscheiden. Chinesen oder Russen haben wiederum eine ganz andere Sichtweise auf die Welt. Und dann denken Sie an all die anderen Sprachen und Dialekte auf diesem Planeten. Daneben produziert jede Sprache zwangsläufig andere Sichtweisen. Alle Fragen, die etwas mehr abstrakt sind, können Computer ebenfalls nicht beantworten. Sie scheitern schon bei simplen Fragen: Wieviele Länder gibt es auf der Erde? Ist Palestina ein Land und was ist mit Schottland? Jeder der von einem großen, semantischen Web träumt, ist im Moment ziemlich naiv. Ich bin auch sehr pessimistisch, dass wir in absehbarer Zeit ein semantisches Web konstruieren können.

W&V Wie lange wird das dauern bis Maschinen sprechen können – in Hollywood-Filmen ist so etwas ja seit Jahren Realität.

Crystal Lange. Linguisten und die Computerexperten werden sicher noch eine Generation brauchen, um Computern das Sprechen beizubringen, aber sie verstehen heute zumindest Sachverhalte. Und über Sachverhalten, die sich unmissverständlich darstellen lassen, mag schon heute eine rudimentäre Mensch-Maschine-Konversation möglich sein. Aber wenn sie Sarkasmus, Ironie oder Begriffe mit Doppeldeutigkeiten wie Metapher benutzen, verzweifelt jede Maschine sehr schnell. Die Linguistik ist schlichtweg ein sehr komplexes Feld. Den amüsanten Beweis liefert Ihnen jedes Übersetzungsprogramm.

[« zurück <URL: /www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)1 >vorwärts »](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)1)

[1 <URL: /www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting >](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting) | [2 <URL: /www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting\(offset\)1 >](http://www.de.infocenter/specials/sprachguru_david_crystal_zu_targeting(offset)1) | [3](#)