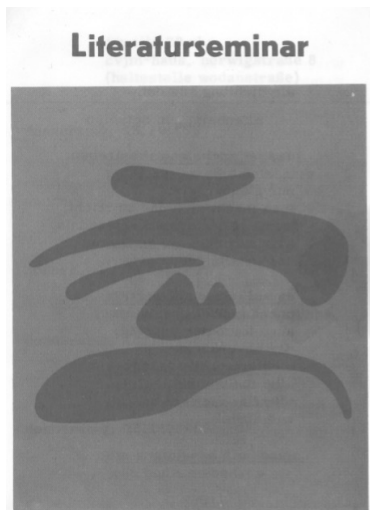
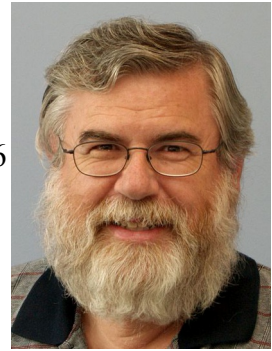


## Von der Kindersprache zur Computersprache und zurück

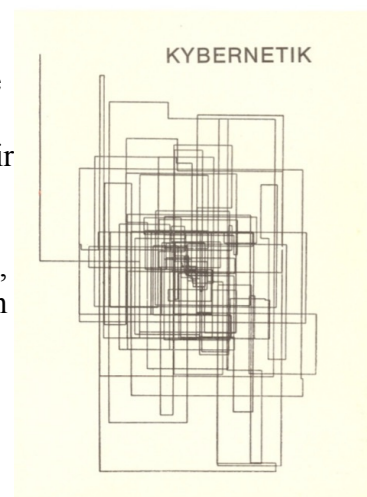
Lieber Gerd,

vor ziemlich genau 50 Jahren lernten wir uns kennen: das Schuljahres 1965/66 brachte uns zusammen, als wir beide in die Klasse 10a des Martin-Behaim-Gymnasiums in Nürnberg eingeteilt wurden. Vom ersten Tag an machten wir so einiges zusammen, zwischen Tanzkurs und Geigenunterricht, wobei letzterer auf Deinen Vorschlag hin zum Bratschenunterricht mutierte. Mir ist die Bratsche ja geblieben, Dir ja nun nicht.

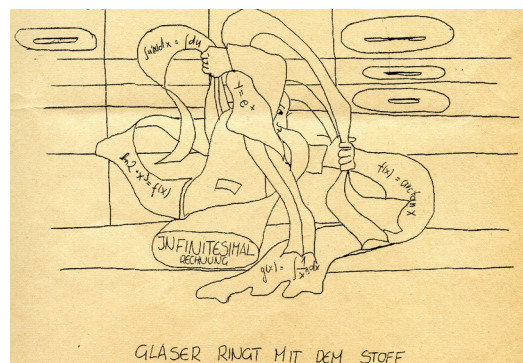


Und dann gab es den CVJM Nürnberg-Lichtenhof, wo man sich im Rahmen von Seminaren mit allerhand Themen beschäftigt hat. Es gab ein Literaturseminar, und es gab ein Seminar zum Thema Kybernetik. Informatik gab es noch nicht: in der Schule erlernte man den Umgang mit Logarithmentafeln und Rechenschiebern, Rechenmaschinen waren mechanische Ungetüme, und die 68-er Generation stand noch bevor. Immerhin fanden wir es ziemlich interessant, das seltsame Thema Kybernetik einmal genauer

anzusehen, hatten aber aus heutiger Sicht nicht die geringste Ahnung, welche Dynamik das Thema in der vor uns liegenden Zeit entwickeln würde, und dass das Thema 50 Jahre später zu einer Cloud für gerdantos führen könnte.



Dann kam, was kommen musste: die 13. Klasse. Glaser ringt mit dem Stoff, wie unserer Kollege Fuhrmann anschaulich darstellt. Wir ringen auch mit dem Stoff, und zwar erfolgreich, nur unterbrochen von der Klassenfahrt nach Berlin. Im Sommer wurde uns das Reifezeugnis ausgehändigt und wir wurden vor die Schultüre gesetzt. Du fandest Aufnahme in der Germanistik, ich in der eigentlich noch gar nicht vorhandenen Studienrichtung Informatik. Gab es da etwas Verbindendes? Konnte man sich etwas Verbindendes vorstellen?



Wir versuchten es -wieder auf Deinen Vorschlag hin- mit Aphasie und Kindersprache, wir versuchten es mit der Frage, wie ein Kind Sprache erwirbt, welche Voraussetzungen im Gehirn angelegt sein müssten, damit Spracherwerb möglich ist. Wir beschäftigten uns mit generativen Grammatiken und dem Herrn Chomsky, der sie erfunden hatte.

Ich erlernte zur gleichen Zeit Programmiersprachen. Und siehe da, auf einmal war Chomsky auch dort zu finden mit seiner Chomsky-Hierarchie für formale Sprachen. Kurz darauf war ich in der Mustererkennerei zu finden. Auf einmal ging es nicht mehr darum, wie Kinder Sprache erwerben,

sondern wie man Computern beibringen kann, Sprache zu verstehen. Was heisst das eigentlich, Sprache verstehen? Wenn man dieses Konzept „Verstehen“ einer Maschine beibringen will, kommt erst einmal völlig durcheinander, weil es eben gar nicht so einfach festgestellt werden kann, ob etwas verstanden worden ist.

Wie Du weisst, hat mich das Thema bis heute begleitet. Im Jahr 2011 hat die IBM Forschung, bei welcher ich mich viele Jahre mit den Grundlagen der Informationswissenschaften befasst habe, ein System vorgestellt, das Antworten auf Fragen geben kann, die in natürlicher Sprache formuliert sind. Das System Watson [https://de.wikipedia.org/wiki/Watson\\_\(Künstliche\\_Intelligenz\)](https://de.wikipedia.org/wiki/Watson_(Künstliche_Intelligenz)) dürfte für den derzeitigen Stand der technischen Möglichkeiten auf dem Gebiet „Automatisches Verstehen natürlicher Sprache“ repräsentativ sein.

Klarerweise nutzt so ein System die Erkenntnisse aus der Sprachforschung, es braucht aber natürlich eine Umsetzung in technische Bausteine. Die Umsetzung wird über Patente dokumentiert. Im Grunde ist das auch nur eine Textsorte. Lass mich nochmals die Frage aufgreifen, wie man einem Computer beibringen könnte, Sprache zu verstehen. In einer Implementierung muss explizit gemacht werden, wie der Vorgang „verstehen“ ablaufen soll. Man könnte fast von einer Definition sprechen, allerdings nur für die gerade vorliegende Situation. Es gibt da ein Patent von einem Kollegen und mir, welches einen einfachen solchen Fall von „verstehen“ implementiert. Im Patent heisst der Vorgang „Suche nach relevantem Text“. Man könnte wahrscheinlich ewig diskutieren, was dies für eine Definition sein soll. Um es kurz zu machen, wir suchten nach einer Möglichkeit, als Ergebnis einer Suchanfrage über eine Suchmaschine nur solche Dokumente (es ging um Artikel) anzuzeigen, die für den Nutzer inhaltlich relevant sind. Ich hatte vorgeschlagen, dies über eine Art graphentheoretische Methode anzugehen. Die verwendeten Graphen entstanden aus den Syntaxbäumen (sozusagen ein Syntaxwald) der (englischen) Sätze in den Artikeln, wobei zwei Knoten in den Syntaxbäumen, die durch eine anaphorische Referenz miteinander verbunden sind, als identisch angesehen werden. Das machten wir sowohl mit den Suchergebnissen der Suchmaschine und wie auch mit den auf der Computerplatte des Benutzers gespeicherten Dokumenten. Der Differenzgraph war die Grundlage für die Berechnung der Relevanz. Hier findest Du das zugehörige Patent:

<https://docs.google.com/viewer?>

[url=patentimages.storage.googleapis.com/pdfs/b3ea5115a895c8ebc93c/EP1225517A2.pdf](https://patentimages.storage.googleapis.com/pdfs/b3ea5115a895c8ebc93c/EP1225517A2.pdf)

Es ist sicher ein einfaches Beispiel. In gewisser Hinsicht sind wir zurück bei der Kindersprache, für Computerkinder, sozusagen.

Lieber Gerd, für uns ist es an der Zeit, den Stab an die nächsten Generationen weiterzugeben. Deshalb sollte die Abschiedsvorlesung eigentlich besser Stabsübergabevorlesung heissen. In diesem Sinne gratuliere ich Dir zu all dem Erreichten und wünsche Dir noch viele gesunde Jahre ohne die Verpflichtungen des Amtes.

Dieter