

Software

Häßlicher Amerikaner

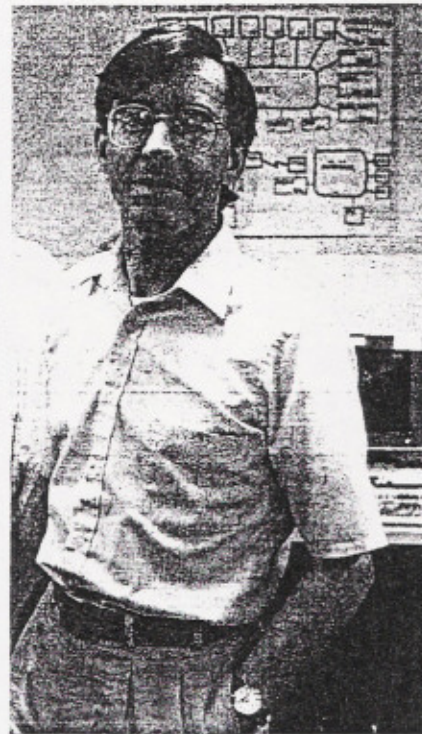
Ein neuer Zeichenstandard soll die babylonische Sprachverwirrung in der Computerwelt beenden.

In den sechziger Jahren, als es entwickelt wurde, habe „ASCII“, das ABC der Computersprache, „großartig dastanden“. Jetzt habe es nur noch das Image eines „häßlichen Amerikaners“. Gefühlvoll gedachte die US-Branchenzeitschrift *Datamation* der besseren Tage des „simplen, kleinen, alten ASCII“.

Der nostalgische Abgesang galt einer schlichten Tabelle mit 256 Buchstaben, Ziffern und Sonderzeichen, von der selbst die meisten EDV-Anwender kaum je etwas gehört haben. Dabei ist ASCII (gesprochen „asskie“) der gebräuchlichste Computercode der Welt.

Ohne ASCII, den „American Standard Code for Information Interchange“ (Amerikanischer Standardcode für Informationsaustausch), würde in der EDV heillos Sprachverwirrung herrschen. Der Zeichensatz schafft Ordnung in der Computersprache: Für viele gängige Schrift- und Grafikzeichen legt er jeweils einen kurzen, computergerechten Code aus Nullen und Einsen fest.

Fast ein Vierteljahrhundert nach ihrer Einführung scheinen die Tage der Tabelle nun gezählt. Der „wohlstrukturierte und übersichtliche Zeichensatz“ (*Datamation*) könnte durch „Unicode“ abgelöst werden, eine Art globales Super-ASCII mit über 65 000 Zeichen, in



„Unicode“-Sprecher Whistler
Freie Fahrt für „ü“ und „ñ“

dem die fernöstlichen Kanji-Symbole ebenso enthalten sind wie etwa das tibetische Alphabet.

Der neue Code wurde von 22 führenden US-Firmen, darunter IBM, Apple, Sun und Microsoft, gemeinsam entwickelt, weil der vergleichsweise bescheidene ASCII-Zeichensatz für die moderne EDV nicht mehr ausreichte. „Kleinigkeiten“, bemerkte dazu *Datamation*, „schaffen manchmal die größten Probleme.“



Computer an der Börse: „\$“ oder „£“?

So beherrscht der „häßliche Amerikaner“ ASCII, sehr zum Leidwesen etwa von international operierenden Firmen, Wissenschaftlern und Software-Entwicklern, wichtige Fremdsprachen nur mangelhaft oder gar nicht. „Weltweiter Datenaustausch“, erklärt Unicode-Mitentwickler Ken Whistler vom kalifornischen Software-Unternehmen „Metaphor“, „wird durch einen derart kurzen Zeichensatz unnötig behindert.“

Zwar hat die „Internationale Organisation für Standardisierung“ (ISO) den ursprünglichen ASCII-Code bereits von 128 auf 256 Zeichen aufgestockt, die aktuelle „Latin1“-Version enthält auch spezifisch deutsche, französische und spanische Buchstaben, beispielsweise „ü“, „é“ und „ñ“. Aber bereits zu Beginn der achtziger Jahre wurde der Computerindustrie klar, daß sie mittelfristig nicht ohne einen völlig neuen, viel umfassenderen Standard auskommen würde.

So muß etwa ein französischer Ingenieur immer noch auf die gewohnten Akzent-Buchstaben seiner Muttersprache verzichten, wenn er elektronische Post an einen Kollegen in Dänemark schicken will. Was auf dem Bildschirm in Paris korrekt als „é“ oder „â“ dargestellt wird, käme in Kopenhagen als wirres Kauderwelsch auf den Monitor.

Probleme gibt es sogar beim anglo-amerikanischen Datenverkehr. US-Computer setzen beispielsweise britischen Kunden ein Dollarzeichen (\$) auf die Rechnung, wo eigentlich britische Pfund (£) gemeint sind. Der Grund: Beide Währungssymbole werden auf der ASCII-Tabelle durch denselben Binär-Code repräsentiert.

Ursache des Problems: Der Zeichensatz läßt sich, aus rein mathematischen Gründen, nicht mehr beliebig erweitern. Weil die Buchstaben, Zahlen und Symbole jeweils als 8-Bit-Folgen, also als achsstellige 0/1-Kombinationen, festgelegt sind, muß es bei 256 Zeichen bleiben. Denn mehr als 256 (2^8) verschiedene Möglichkeiten, Nullen und Einsen zu kombinieren, hat der sogenannte 8-Bit-Code nicht zu bieten.

Unicode löst dieses Problem auf ebenso einfache wie wirkungsvolle Weise. In der neuen Tabelle wird jedem Zeichen eine Sequenz von Nullen und Einsen zugewiesen, die doppelt so lang ist wie bei ASCII. Auf diese Weise – in einem 16-Bit-Code – lassen sich insgesamt 65 536 Zeichen definieren (2^{16} verschiedene 0/1-Kombinationen). Das sind mehr als genug, um alle Schriftsprachen der Welt für die EDV zu standardisieren.

Doppelte Code-Länge bedeutet allerdings auch, daß jedes einzelne Zeichen im Arbeits- und Festpeicher des Computers zweimal soviel Platz beansprucht wie bisher. Die Unicode-Version eines

elektronischen Briefes beispielsweise wäre doppelt so lang wie dieselbe Mitteilung in ASCII-Lesart.

So ist der neue Zeichensatz der jüngsten Generation leistungsfähiger Arbeitsplatzrechner, den oftmals vielfach vernetzten sogenannten Workstations, gleichsam auf den Leib geschrieben. Großrechner-Betreiber mit riesigen Datenbanken hingegen, erklärt Unicode-Mitentwickler Whistler, können auf den neuen Code einstweilen noch verzichten: „Die werden doch ihre Bestände nicht aufs Doppelte aufpusten.“

Für Geschäftsleute und Wissenschaftler dagegen wäre der erweiterte Zeichensatz ein immenser Fortschritt. So ist in Unicode (neben vielen anderen) auch der kyrillische und der hebräische Zeichensatz integriert, dazu die chinesischen, japanischen und koreanischen Alphabete (insgesamt 4096 Schriftzeichen), für die es bisher verschiedene gesonderte Standards gab. Als größten Block enthält die Tabelle 27 000 chinesische Han-Symbole.

Der internationale Datenaustausch könnte durch den neuen Code in Zukunft erheblich vereinfacht werden, Hard- und Software für ausländische Märkte ließen sich sehr viel schneller als bisher an die jeweiligen Zeichensätze anpassen. „Auf lange Sicht“, meint Unicode-Sprecher Whistler, „wird sich das gewaltig bezahlt machen.“

Ob sich Unicode allerdings als Weltstandard durchsetzen wird, hängt vor allem von der ISO ab, die bereits seit 1983 ebenfalls an einem entsprechenden Super-Zeichensatz arbeitet, genannt „ISO 10646“. Den letzten ISO-Entwurf hatten 14 Länder, darunter die USA, China, Frankreich und Großbritannien, wegen deutlicher Design-Mängel abgelehnt.

Seit Montag letzter Woche wird nun in Genf, dem Sitz der Standard-Organisation, neu verhandelt. Die meisten Beteiligten halten einen Kompromiß aus beiden Zeichensätzen, mit einem deutlichen Übergewicht von Unicode, für wahrscheinlich.

Wenn es nicht zu einer Einigung kommt, soll der Unicode-Zeichensatz, Version 1.0, im Herbst als Buch veröffentlicht werden. Seit Ende vergangenen Jahres verfügen die Konsortiumsmitglieder über eine Vorabfassung. So will etwa der Software-Riese Microsoft sein Erfolgsprogramm „Windows“ dem Unicode anpassen, eine Demo-Version wurde vorletzte Woche bei einer Entwicklerkonferenz vorgeführt.

Die Unicode-Befürworter sind jedenfalls für alle Eventualitäten gewappnet. Außer im Unicode-Konsortium sind Firmen wie IBM, Apple, Microsoft und Xerox, sicher ist sicher, auch noch im ISO-Komitee vertreten.

Mayas

Kopf ab

Eine Öko-Katastrophe, ausgelöst durch wiederholte Angriffskriege, führte zum Untergang der Mayas.

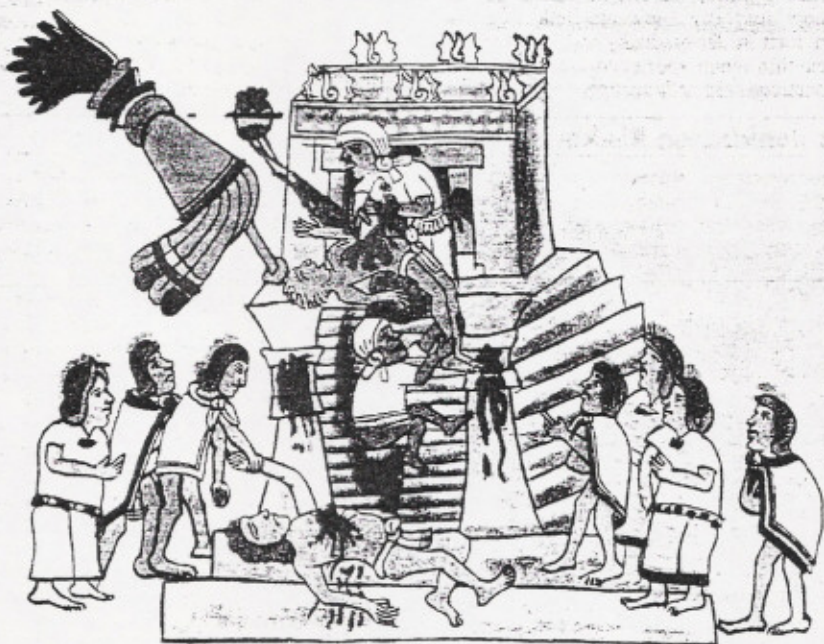
Die Archäologen erklimmen die Spitze einer zerfallenen Pyramide; dort hoben sie einen brunnenförmigen Schacht aus. Durch den steinernen Schlund drangen sie, bedrängt von rutschendem Geröll, tief in das Innere der Tempelruine.

Die schweißtreibende Wühlerei, mitten im dampfenden Urwald von Guatemala, wurde zum Abstieg in eine dunkle Vergangenheit. In einem von Einsturz bedrohten Gewölbe, nahe dem Pyrami-

um „die Schlüsselfigur beim Untergang der Maya-Zivilisation“.

Der einflussreiche Maya-Herrscher, erläutert der Anthropologe von der Vanderbilt University in Nashville (US-Staat Tennessee), habe eine „katastrophale Kette von Ereignissen ausgelöst, die dazu führten, daß das Maya-Reich in ein Dutzend sich bekriegender Kleinstaaten zerbrach“. Innerhalb von nur 50 Jahren sei die gesamte Indio-Gesellschaft „militarisiert“ worden, sagt Demarest, was schließlich die „Ökologie der Region zerstört“ habe.

Demarests Erkenntnisse könnten womöglich erklären, weshalb die zwischen dem Golf von Mexiko und der Karibischen See angesiedelte Hochkultur irgendwann nach dem Jahre 900 ihr Zerfalls-Datum erreichte. Die vorherr-



Menschenopfer bei den Mayas*: „Blut als Bindemittel der Gesellschaft“

den-Zentrum, stießen die Forscher auf ein Gerippe, das mit funkelnden Edelsteinen bestückt war: Ein mosaikartiger Kopfschmuck, zusammengesetzt aus Muscheln, Perlmutter und Jade, krönte den bleichen Schädel.

Der prächtige Kopfputz entsprach der besonderen Bedeutung des Toten. Die Altertumsforscher hatten, nach zweijähriger Suche, die Grabkammer eines einst mächtigen Maya-Herrschers entdeckt – nach Ansicht der Wissenschaftler stellen die fürstlichen Gebeine, die im Frühjahr in der verlassenen Maya-Stadt Dos Pilas ausgegraben wurden, einen sensationellen Fund dar: Bei dem Toten handele es sich, wie der Ausgrabungsleiter Arthur Demarest annimmt,

* Mexikanische Zeichnung aus dem 17. Jahrhundert.

schende Forscher-Meinung ist, daß die Mayas seinerzeit aus den tropischen Regenwäldern des heutigen Guatemala und Südmexiko verschwanden.

Eine Bestätigung ihrer neuen Untergangs-Deutung versprechen sich die Altertumsforscher nun von der Dechiffrierung der verschiedenen Grabbeilagen, die neben den Gebeinen des Maya-Fürsten in der „archäologischen Schatzkammer“ (*Newsweek*) verstreut lagen – darunter scharfe Obsidian-Messer, mit denen rituelle Handlungen vollzogen wurden, und mit Inschriften versehene Tonkrüge.

Auf den Grabgefäßen, vermuten die Wissenschaftler, sei gewiß auch der Name des unheilträchtigen Potentaten vermerkt, den die Forscher einstweilen als „Herrscher Zwei“ der Maya-Stadt Dos Pilas bezeichnen.