SGML: Erik Naggum's Brief Description

```
Article: 7689 of comp.text.sgml
Newsgroups: comp.text.sgml
Date: 07 Feb 1995 11:02:46 UT
From: Erik Naggum <erik@naggum.no>
Organization: Naggum Software; +47 2295 0313
Message-ID: <19950207T110247Z.enag@naggum.no>
References: (7688) <3gs6gt$5nl@carbon.cudenver.edu>
Subject: Re: SGML-- What is it and how can we use it
```

```
[Daniel E. Cogswell]
```

> I need to know exactly what SGML is. Is it a "standard" or is it
> something else. Where can I get the "textbook" on it?

SGML is a standard, in the formal sense -- hammered out and agreed upon by an international community working under the auspices of the International Organization for Standardization (ISO). its formal, full name is ISO 8879 Information processing -- Text and office systems -- Standard Generalized Markup Language (SGML), first edition published 1986-10-15, first amendment published 1988-07-01. it's available from your friendly standards office (who will demand tons of money for it).

to describe exactly what SGML is is very difficult. it's a language that can be used to build the infrastructure for interchange of and longevity for information. by way of analogy, one could describe it as "SGML and the Art of Information Maintenance -- An Inquiry into the Value of Information" (with apologies to Robert Pirsig). that is, a way of life once you have realized that the information we create take on a life of its own and it can die if we don't care for and feed it properly. in ancient times, you had to burn down a major library to destroy information, but you got to be remembered for it. today, you need only upgrade to the latest version of a particular software product, change a printer, use patented software in the compression of the data, etc, to destroy many orders of magnitude more information, but the history books have yet to notice that the previous generation was the last to leave permanent traces of its tools.

SGML is still mostly viewed as a format used in publishing printed documents and multimedia CD-ROMS and such. publishing was the original purpose of the standard, but it was soon apparent that it had far greater potential (sometimes referred to as "Sounds Great, Maybe Later"). outside of the publishing industry, understood suitably widely, SGML is thus regarded as a possible means to save the information that mankind generates and stores in perishable, proprietary, un(der)documented formats. e.g., during the time it takes to write and produce a dictionary, the computer industry will go through at least two major revolutions. in an industry where "three seconds is a long time", the things it helps build: oil rigs, cities, laws, "cultural heritage", standards, all have lifespans of several billion seconds. (trivia: 3 billion seconds is a little over 95 years.)

however, in the current trend that has lasted for a couple hundred million seconds, the only things that matter are products, "compatibility", and using computers to mimic paper and display media. this trend will pass. then where will we be? scared people will stick to their old data and realize that those annoying visionaries in the 80's and 90's were right.

lots of people are perfectly happy to do he same work over again, work that others have done before them and yet others will do after them, even when there are no external demands on the material to require this. vast hordes of managers are perfectly happy to waste as many dollars as they waste seconds with this scheme. whole nations are built by employing such people to push papers around. clearly, one standard can't stop such waste. but it can make life easier for those who refuse to subscribe to this wanton abuse of resources.

SGML is an idea, a philosophy, a language to express grand visions for both present and future. if it is any good, it will of necessity be used by those who have much more limited visions. like the literary quality of pocketbooks, cartoons and MTV, HTML fills this role. like the readers of pocketbooks and cartoons, most HTML fans scorn its heritage. some do see the connection, and continue to care about the grander schemes, though.

unfortunately, there aren't any really good "textbooks" on SGML. you won't read about any grand schemes, any infrastructure or systems building, any issues of the longevity of information. you can read about the details of the language in Charles F. Goldfarb: The SGML Handbook; Oxford University Press, 1990. ISBN 0-19-853737-9. you can read various accounts of some aspects of its use in other books with "practical" and "guide" in their titles, the younger the better. <plug> if you look for a textbook on SGML suitable for university level computer science, I hope to get it out late this year or early next. I believe SGML will become irrelevant unless its core concepts are well understood and supported by programmers and computer scientists at all levels. the lack of tools, books, and even interest in academia (where the C, Unix, the Internet, WWW, etc, started), is a death knell ringing for SGML. I intend to reverse that. this means a change of focus away from publishing and information products to information sources, which is where I think SGML has always belonged. </plug>

somebody else can cover the practical issues. thanks for listening.

#<Erik>

```
--
miracle of miracles.  look what the Net dragged in.


------------------------------

End of comp.text.sgml digest Vol 7 #4
************************************
```