

## A possible link between Rocappi and SGML

Jonathan Seybold, September 2018

Most of the people who came to visit Rocappi — or listened to our presentations — didn't really “get” the separation of form and content that we thought so important. Bill Tunnicliffe was one of the exceptions.

At the time, Bill worked for the Lowell, MA *Courier Citizen* newspaper. But, what became his obsession with separation of form and content had nothing to do with the newspaper industry. Bill immediately understood that, in the future, published documents would be stored as binary data on computers. But, sharing documents encoded in a wide variety of different typographic command languages could be a tower of Babel.

He saw Rocappi's separation of form and content as the answer. However, our tags had no inherent meaning. Until we defined them, they were just tags. This was incredibly powerful, but Bill thought that tags to be used for document exchange could not be this generic. They had to be specific: “Head 1”, “Head 2”, “Body copy” etc. etc.

Bill devoted himself to trying to invent such a language. It was a daunting task. You would have to come up with language of tags that could be used to define the structure of a wide range of document types.

**GCA.** In 1967 Norman Scharpf formed the Graphic Computer Communications Association, as a Printing Industries of America (PIA) special interest group. (“Computer” was later dropped from the name and GCCA became GCA.)

We supported Bill in presenting his ideas to Norm. Norm asked Bill to present them at a GCA meeting in Toronto, Canada that fall. This presentation generated a lot of interest. GCA formed a “System X” Committee charged with creating a generic text markup language (GenCode). Bill was the chairman of that group.

Inspired by Bill's ideas, but independent of the System X group, Stanley Rice, a book designer in New York City, published his own thoughts on what such a language should take into account.

**GML.** In 1969, inspired in part by Stanley Rice's thinking, Charles Goldfarb, Ed Mosher and Ray Lorie of IBM created the precursor to GML (Generalized Markup Language). This was similar in concept to GenCode, but more structured and more hierarchical.

Goldfarb presented his early GML work at the GCA Annual Meeting in 1971 and met with the GcnCode committee. But the two groups continued in parallel for some time.

**SGML.** In 1978, the two efforts joined together under the auspices of ANSI to create a single standard (SGML) language from GML. GCA under Norman Scharpf directed the project. Charles Goldfarb led the committee. The core of the working group were people from the GenCode Committee.

The first SGML working draft was published in 1980. GCA published the 1983 draft as a working standard. And, SGML was adopted by the ISO as an international standard in 1985 — almost 20 years after Bill Tunnicliffe had first started working on the concept.

Bill served as the first head of WG8, the ISO committee responsible for SGML. His incredible single-handed efforts were responsible for getting SGML adopted as an ISO standard, despite fierce administrative and political opposition.

GCA, under Norman Scharpf, continued to have a primary role in promoting, defining and evolving SGML. This continued into the Internet era. In 2001 GCA was finally spun off from PIA as a stand-alone entity in recognition of the fact that, by this time, its primary focus was on electronic documents — and XML in particular.

**Why do we care now?** HTML was a partial implementation of SGML. Tim Berners-Lee called it “an application of SGML.” Creating an international standard does not happen over night. It took years of work by a lot of very dedicated people. Charles Goldfarb, Bill Tunnicliffe, and Norman Scharpf were the key players, but many, many others made important contributions.