Abstracts

Les Cahiers GUTenberg
Contents of Recent Issues

Numéro 21 — juin 1995
Congrès 95 GUTenberg:
Une nouvelle vision des documents
La Grande Motte, 1 et 2 juin 1995

Editor's note: Thematic issue entitled "A new vision of documents", the proceedings of GUTenberg'95, a meeting held June 1–2, 1995 at La Grande Motte (just outside Montpellier) by the French-speaking TeX Users Group. The issue contains 166 pages in all, with 11 articles.

ALAIN COUSQUER, Éditorial : un congrès, pour quoi faire? [Editorial: Why have a conference?]; pp. 3–4

Cousquer outlines the pros and cons of TeX meetings, starting with reasons for the 4-year gap since the last major GUTenberg meeting: difficulties in getting things organised, the costs, finding a theme that would attract a wide range of people. And while focused, thematic day-long meetings have proven quite successful — the interest, number of participants, and the high calibre of the discussions — there's nothing quite like a longer meeting for people to meet and exchange views and experiences over a few days' time.

The theme for the 1995 meeting — a new vision for documents — was introduced with a paper by Fernand Baudin (whose presentation is not in the issue) and then carried further by the subsequent eleven presentations.

MICHEL GOOSSENS and MICHÈLE JOUHET, Les commandes graphiques en LaTeX 2_ε [Graphics commands in LaTeX 2_ε]; pp. 5–29

Authors' abstract: "LaTeX 2_ε provides a generalized driver-independent interface for the inclusion of external graphic material, as well as for scaling and rotation of LaTeX boxes. These features are not in the LaTeX 2_ε kernel, but are implemented in the graphics or graphicx extension packages. These packages rely on features that are not in TeX itself, but which must be supplied by the "driver" used to print the dvi file.[1] Inside a `picture` environment

---

[1] Unfortunately not all drivers support the same features, and even the internal method of accessing these extensions varies between drivers. Consequently all these packages take options such as "`dvips`" to specify which driver is being used. [footnote in original]

LaTeX 2_ε also offer[s] better support to draw Bezier curves with the `\qbezier` command."

MICHEL GOOSSENS and MICHÈLE JOUHET, Utiliser la couleur avec LaTeX 2_ε [Using colour with LaTeX 2_ε]; pp. 30–52

Authors' abstract: "The first part of this article reviews some basic principles underlying the use of colour in publishing. We explain how to use colour in a functional way to help focus attention, explain relationships, guide the viewer/reader through the presented information. Simple rules for optimizing communication using colour codes are discussed. The colour interface of LaTeX 2_ε is the subject of the second part of the article, where we describe commands to write text in colour and to generate coloured boxes or backgrounds, and the various colour models to define names of available colours. As with the graphics package it [is] the DVI drivers that are responsible for actually generating the colours on paper or on screen." [includes 9 figures in colour]

SARRA BEN LAGHA and MOHAMED BEN AHMED, Intégration de graphiques dans du texte [Integrating graphics with text]; pp. 53–67

Publishing technical documents is a complex operation, requiring a variety of skills and tools. For document exchange in a diverse environment to take place, the documents have to be standardized. Their various components (text, graphics, etc.) must each conform to one or more standards, while the subsequent integration of these components must then also meet certain standards. Here we're talking about documents that address multiple standards, on multiple platforms. This article reviews the basic ideas behind structured SGML documents and the CGM graphics standard, before presenting two methods for integrating graphics into structured SGML documents.

Keywords: document, graphics, integration, standards, SDIF, SGML, DTD, CGM, GKS.
[from the Résumé]

JEAN-JACQUES GIRARDOT and BERNARD AMADE, Une expérience de production automatisée de documents : approche, problèmes et solutions [An experiment in automated document production: approach, problems and solutions]; pp. 68–74

This article discusses two experiments in production automation of documents linked to financial applications.

The main difficulty was to find a way to combine a "logical" vision of documents — and the constraints that existing applications impose on the

production of such an abstract form — with both the manual addition of information by users and the final form of the documents.   [from the Résumé]

JACQUES ANDRÉ, JEAN-DANIEL FEKETE and HÉLÈNE RICHY, Traitement mixte image/texte de documents anciens [Combined image/text handling of old documents]; pp. 75–85

We are interested in "texts with graphics" which are documents (for example, old manuscripts) studied for both their content and their form. Putting them into electronic form requires manipulating them both as textual objects and as images. The idea of a "hotspot" [*zone sensible*] for images is a first step but proves inadequate for the electronic manipulation of old documents in a professional manner.   [from the Résumé]

DENIS MÉGEVAND, Personnaliser ses lettres avec LaTeX 2$_\varepsilon$ [Customising letters with LaTeX 2$_\varepsilon$]; pp. 86–95

The lettre style [an adaptation and extension of letter] can be used for letters and personal faxes, and can be adapted for the European style and format of envelopes with windows. Various languages can be used for the text; once selected, various default values (format for the date, and even the correct words for the various elements of a letter — enclosure, salutation, closing, etc.) are invoked. The style file also allows the user to customize a letterhead and the other various "standard" fields found in letters and faxes.   [from the Résumé]

DANIEL TAUPIN, Faire un ouvrage portable en LaTeX 2$_\varepsilon$ [Making a large job portable in LaTeX 2$_\varepsilon$]; pp. 96–106

Author's abstract: "Writing a big book in LaTeX 2$_\varepsilon$ with many figures raises several problems of 'TeX capacity' and also some portability problems between the various kinds of printers and drivers. We solved that challenge by inserting figures in PCX code, and using several well-known and more [reliable] MS-DOS tools which enable enlarging and shrinking these figures, [while] minimising distortions and information loss."

DANIEL TAUPIN, ROSS MITCHELL, and ANDREAS EGLER, MusiXTeX : L'écriture de la musique polyphonique ou instrumentale avec TeX [MusiXTeX: Writing polyphonic or instrumental music with TeX]; pp. 107–113

Authors' abstract: "MusiXTeX is a new music typesetting package derived from MusicTeX, but it provides more beautiful scores than MusicTeX did. While MusicTeX was a single-pass package, MusiXTeX is a three-pass system: the first pass per-

forms a rough TeXing which reports the spacings of each music section, the second pass is a computation of the best note spacings, and the third one is the final TeXing process.

The quality of single notes is the same as in MusicTeX, but slurs are much more beautiful, and notes are regularly spaced instead of being irregularly spaced with glue."

MICHEL BEIGBEDER and JEAN-JACQUES GIRARDOT, La conversion de midifiles en MusicTeX [Converting midifiles into MusicTeX]; pp. 114–126

The typesetting of musical scores is quite different from that of text, and thus requires sufficiently sophisticated programs — and even TeX in its regular form is not up to the task. However, LaTeX style files now provide the functionality specific to musicologists' needs, by providing a bridge between an already widely used format for notation — *midifiles*, a de facto standard of the Computer Music Association — and MusicTeX.

Keywords: midifile, MusicTeX, musical parts

HÉLÈNE RICHY, CHRYSTÈLE HÉRAULT and JACQUES ANDRÉ, Notion de "feuille de style" (résumé étendu) [The idea of a "style sheet" (expanded summary)]; pp. 127–134

The structured document concept makes it possible to distinguish between the content of a document and its logical structure (DTD in SGML). The idea of the graphical structure of a document can now be added. As we show in this article, the specifications of such a graphical structure are far from standard: various proposals based on HTML in particular exist. Beyond some apparent deviations, the idea of style sheet is beginning to be accepted and should make document exchange easier.

Keywords: style sheet, structured document, formatting, HTML, Grif

[from the Résumé]

R.D. HERSCH and V. OSTROMOUKHOV, Introduction à la génération d'images en demi-tons [Introducing the generation of half-tone images]; pp. 135–165

A detailed technical discussion of half-tone images, after a brief introduction, the 31-page article moves through eight sections on the various methods and results available to the user, and is completed with a short bibliography. Each section includes schematic diagrams, equations for various calculations, and sample output for each method, each one a further refinement of the last. Not an article for the general reader, but likely very

interesting and helpful to those dealing with halftone images in their documents.

Numéro 22 — septembre 1995
Ligatures & caractères contextuels

Editor's note: Thematic issue entitled "Ligatures and context-dependent characters". It is a compilation of papers originating from several sources in 1992: a typography course given by the Didot project[2] at the École supérieure Estienne (in Paris), and several meetings held in Lurs (Upper Provence) by the "Rencontres de Lure", an international typographic association founded in 1952.

The 135-page issue contains 8 articles.

JACQUES ANDRÉ and JEAN-LOUIS ESTÈVE, Éditorial : de Didot à GUTenberg [Editorial: From Didot to GUTenberg]; pp. iii–vi

Along with his co-editor from the École Estienne, Jacques André introduces the word play on "Didot" [see footnote] and then elaborates on the Didot project, which tries to bring together the highly aesthetic world of typography and the highly technical world of the computer. About a dozen Didot sessions took place between 1989 and 1993, resulting in two books and a teaching program on computer-based typography. This *Cahier* issue suffered production delays since it seemed only appropriate to ensure the highest typographic quality in a publication about high-quality typography; and this was after lengthy editorial work had been completed. The issue was prepared using LaTeX 2ε and Adobe-Garamond as the main font; the resulting PostScript files were sent via the Internet to the Louis-Jean printing house for offset printing on Arcole paper. An electronic version is available at http://www.univ-rennes1.fr/pub/GUTenberg/publications/publis.html.

JACQUES ANDRÉ, Introduction : vous avez dit "ligature"? [Introduction: you said "ligature"?]; pp. 1–4

While Jérôme Peignot's article "Petit traité de la ligature"[3] serves as the base text on the subject — until photocomposition, that is — this issue of the *Cahiers* brings the reader up to date on new, or revitalized and re-tooled ligatures. The author outlines a number of types of ligatures, which are not confined to usual character combinations (e.g. "fi" or ffi"), but include also special characters (e.g.

& and @) or logotypes. In short, "ffi" is just the tip of the iceberg.

ADOLF WILD, La typographie de la *Bible* de Gutenberg [The typography of the Gutenberg *Bible*]; pp. 5–16

Author's abstract: "The author examines Gutenberg's *Bible* from a typographical point of view, looking both at page structure (in the light of Rosarivò's grid constants) and, especially, at line structure, where it can be seen that in contrast with present-day methods of justification, based on variable inter-word spacing, Gutenberg balanced his lines using a variety of letter combinations: logotypes, ligatures, abbreviations, etc."

RENÉ PONOT, Le *Didot* a-t-il besoin de ligatures ? [Does *Didot* need ligatures?]; pp. 17–42

Author's abstract: "Having first reviewed the history of ligatures, the author then describes the origins of the Didot typeface and its principal features. He concludes that this typeface has no need for ligatures."

GÉRARD BLANCHARD, Nœuds & esperluettes : Actualité et pérennité d'un signe [Logotypes and ampersands: A symbol for today and forever]; pp. 43–59

In May 1992 a Didot seminar on "logotypes, ligatures and context-dependent symbols" was held in Paris at the École Estienne. In the course of this meeting, Gérard Blanchard presented a paper with an unexpected topic: the ampersand, because, to his eyes, this symbol — or rather, this "type" — was the touchstone for the sense of calligraphy in typography. Starting from a Tschichold text,[4] this article gives the background to his ideas.   [from the opening summary]

JACQUES ANDRÉ, Ligatures & informatique [Ligatures & computers]; pp. 61–85

Author's abstract: "In digital typography a ligature consists in replacing n glyphs by a single glyph. This raises problems at the keyboard level: how can two letters such as c and t be specified to be replaced by the ct ligature? How does the user know if the font in current use contains the necessary ligature? Is it possible to design a system in which the user does not have to deal with such problems? The present paper looks at such problems from the point of view of coding and font metrics.

---

[2] "**DI**dacticiel de **D**essin assisté par **O**rdinateur de caractères **T**ypographiques", or "**DI**gitization and **D**esign **O**f **T**ypefaces".

[3] *Communication et langages*, 73.3 (1987), pp. 20–35.

---

[4] *Formen Wandlungen der Etzeichen*; translation into French by René Grasset forthcoming, from the École Estienne.

The final section provides a brief introduction to script ligatures and the related problems of handwriting simulation and recognition."

Yannis Haralambous, Tour du monde des ligatures [A world tour of ligatures]; pp. 87–99

The term "ligature" signifies a number of different things and is often the source of confusion. In the TEX world, the situation is exacerbated by the fact that D.E. Knuth has given this name to a portion of the font metrics information.

To clarify this very intriguing subject, we are going to try to classify the different types of ligatures that one finds in natural languages and to relate these findings to TEX. The three types of ligatures are: linguistic, esthetic, and contextual.   [from the introduction]

Thierry Gouttenègre, Ligatures & bâtardes bourguignonnes : Du xvᵉ au xxᵉ siècle [Ligatures and the Burgundian Schwabacher fonts: from the 15th to the 20th century]; pp. 101–106

Author's abstract: "The author outlines some ligature-related problems which he had to deal with when designing a typeface based on a fifteenth-century book hand."

François Boltana, Ligatures & calligraphie assistée par ordinateur [Ligatures and computer-assisted calligraphy]; pp. 107–124

Author's abstract: "Three calligraphic typefaces (Champion, Messager and Aurore) provide their designer with the basis of a discussion of the possibilities created by the use of computer tools [to] recreate ligatures. Such a development would not only bring us closer to the rich calligraphic texture of [18th century] written forms, it would also give new life to the concept of contextual typography created by Gutenberg. For though Gutenberg's technique is no longer appropriate to modern conditions, the spirit in which he copied calligraphic forms is more than ever topical."

Gérard Blanchard, Postface ; pp. 125–132

Blanchard discusses the wherefores and whys of ligatures themselves, calling them *cosa mentale* — something like a "state of mind". From flourishes to minimalist strokes, the ordering of elements, and thus of ligatures survives. Ligatures tie things up neatly — they join two words or phrases into a single entity; logotypes, another variation on ligatures, are like a company logo — the final seal on a package for delivery. Both pass from one language to the next, understood and untranslated.

The author then discusses each of the 8 articles in the issue, bringing in much additional information

on other papers and presentations, providing a very useful and informed view of what has been detailed by each of the other authors.

Cahiers GUTenberg, Sommaire des derniers numéros parus [Summary of recent issues]; pp. 133–135

A listing of the contents of issues 16 through 22 of the *Cahiers*.

Numéro 23 — avril 1996
FAQ : questions souvent posées

Editor's note: Thematic issue entitled "FAQ : frequently asked questions". This is a French translation (by various members of GUTenberg's board). The issue has six preliminary pages (including the editorial) and 120 pages of contents (including a listing of contents of previous *Cahier* issues, and information about joining GUTenberg).

GUTenberg Board Members, Editorial ; pp. iii–vi

The general notion behind a FAQ is given — that while there are many sources of information available to users, there is often much repetition of certain questions and their answers over time. However, it's not enough to simply compile the frequently asked questions; one also has to ensure that the answers are indeed complete and correct. As well, while the main source of both questions and FAQs is the Internet, having a paper copy available for the wider audience is also important — and a big job. Bobby Bodenheimer is well known as the originator of the (LA)TEX FAQ, a large portion of which was published in *Cahier* 13 (1992). The UKTUG group also published the FAQ, and their revised and updated edition[5] is the source of the current French version, which augments the original text with issues pertinent to French-language texts, additional questions deemed important, and French sources and/or translations of information. As well, an index has been added.

The editorial closes with a translation of the *Baskerville* editorial by Robin Fairbairns, which mentions that translations in French (voilà!) as well as Czech and Russian, are in progress, and that Elsevier Science had graciously underwritten publication of copies for attendees at the TUG'95 meeting in Florida.

In short — the FAQ is being viewed by many groups as being of great assistance and significance to all users.

---

[5] *Baskerville* 5(6) (1996), pp. 1–29.

UK TeX Users Group, Questions souvent posées
sur (LA)TeX [(LA)TeX Frequently Asked Questions];
pp. 1–114

The entire issue is given over to some 129
questions, divided up into sections A through S, and
is completed with a 7-page index. The issue can be
viewed simply as a French translation of the FAQ,
with additions to directly address French-language
issues; it can also be viewed as an exceptional
resource for TeX- and computer-related terminology
in French. While the GUTenberg team occasionally
maintained terms such as "bitmap" (they question
whether anyone would know what to make of
"plan de bits"!), they made every effort to provide
official/correct French terminology. Thus, not only
are the "contents" a mine of information but also the
"form" of the contents are equally valuable. And for
a linguist, form and content are always at the heart
of the issue ;-)

Cahiers GUTenberg, Sommaire des derniers
numéros parus [Summary of recent issues];
pp. 115–117

A listing of the contents of issues 16 through 23
of the *Cahiers*.

Numéro 24 — juin 1996
TEI : Text Encoding Initiative

Editor's note: This thematic issue contains 8
articles.

François Role, Éditorial : TEI — Text Encoding
Initiative ; pp. 1–3

While the *Cahiers* have often addressed the use
of TeX in the mathematical sciences, articles on such
topics as multilingualism, literature, and the social
sciences have also appeared in its pages. GUTenberg
and its *Cahiers* are thus open to discussions on
any and all electronic document codings, and not
uniquely those of TeX.

The TEI project grew out of a conference held
in Poughkeepsie, NY, in 1987. The purpose: to fa-
cilitate the electronic exchange of documents within
the scientific community; in 1994, the "Guidelines
for Electronic Text Encoding and Interchange" were
published (and now contain some 1300 pages). The
TEI recommendations rely on SGML, and in fact,
they contain several DTDs, along with detailed com-
mentary on the use of tags and attributes.

This issue of the *Cahiers* begins with an
introduction to the subject, and then essentially
contains a translation of a large and comprehensive
document called "TEI Lite", written by two of
TEI's founders, Lou Burnard and C.M. Sperberg-
McQueen. The subsequent articles provide short

introductions to various applications of the TEI
principles to a variety of research projects, all
currently underway in France.

The editorial concludes with a listing of Web
sites where TEI tools can be obtained:
```
http://www.sil.org/sgml/publicSW.
    html#parsers
http://aix1.uottawa.ca/~dmeggins/SGMLSpm/
    sgmlspm.html
http://ftp.ifi.uio.no/pub/SGML/demo/
http://www.oclc.org:5046/oclc/research/
    panorama/panorama.html
http://www.sq.com/products/panorama/
    pan-free.html
http:/ftp.lri.fr/LRI/soft/ihm/tei2latex-0.
    1.tar.gz
```

Nancy Ide and Jean Véronis, Présentation de
la TEI : *Text Encoding Initiative* [Introducing
TEI: The Text Encoding Initiative]; pp. 4–10

This is a short overview of the TEI project, its
history and development, the purposes and princi-
ples behind the project. It includes an outline of
the core TEI DTD, as well as a list of project groups
which have adopted the TEI recommendations. The
piece concludes with information on where to get
more complete documentation, some of it being the
English originals of what appears in this issue of
the *Cahiers*. There is a bibliography and then a
translation of the Table of Contents of the TEI
Guidelines document.

A Web site with the TEI documentation
is: `http://www-tei.uic.edu/orgs/tei`. As well,
electronic copies can be ftp'd from: `ftp-tei.uic.
edu` (pub/tei), `info.ex.ac.uk` (pub/SGML/TEI),
`TEI.IPC.Chiba-u.ac.jp` (TEI/P3), or `ftp.ifi.
uio.no` (pub/SGML/TEI). And finally, paper copy is
also available by ordering from `http://www-tei.
uic.edu/orgs/tei/info/p3order.html`. [Recall:
it's 1,300 pages long, hence a cost of $75.00 or £50.]

Jacques André, Balises, structures et TEI [Tags,
structures and the TEI]; pp. 11–22

By way of introducing the Text Encoding Ini-
tiative (TEI), a summary is provided of this type
of coding, which features the use of parenthetical
tags in the text, thereby making it possible to mark
one or more eventual structural outputs. The differ-
ence between such "internal" coding and "external"
coding (as formatted by a text editor for screen
display), and those [codings] found in commercial
products such as Word must particularly be empha-
sized. This structured tagging is meaningless unless
one uses tools designed to interpret it, making it

possible to revise the coding as well as editing the documents.   [from the Resumé]

LOU BURNARD and C.M. SPERBERG-MCQUEEN, La TEI simplifiée : une introduction au codage des textes électroniques en vue de leur échange [TEI made easy: an introduction to electronic documents for the purposes of interchange]; pp. 23–151

Authors' abstract: "This document is a French translation [by François Role] of the English document: 'TEI Lite: An Introduction to Text Encoding for Interchange'. These Text Encoding Initiative (TEI) Guidelines are addressed to anyone who wants to interchange information stored in an electronic form. They emphasize the interchange of textual information, but other forms of information such as images and sounds are also addressed. References of the genuine English documents are listed below."

FRANÇOIS ROLE, Le codage informatique des apparats critiques : évaluation des recommandations de la *Text Encoding Initiative* [Computer coding of the critical apparatus: An evaluation of the TEI recommendations]; pp. 153–165

Following a quick definition of the notion of critical apparatus and a detailed presentation of the TEI recommendations for this type of document [critical editions], we will examine some possible extensions to the TEI to better address the particular characteristics of certain types of texts (modern manuscripts, texts whose physical characteristics are of interest).   [from the Resumé]

NANCY IDE and JEAN VÉRONIS, Une application de la TEI aux industries de la langue : le *Corpus Encoding Standard* [A TEI application in the field of languages: The *Corpus Encoding Standard*]; pp. 166–169

The Text Encoding Initiative covers an extremely wide range of documents (prose, poetry, theater, dictionaries, lexicons, and so on), each intended for a variety of final uses (electronic publication, literary and historical analysis, hypertext, etc.). The tags available, as well as the Document Type Definition (DTD), are therefore both too rich and too poor for any one application. The tags and the DTD, however, can be made modular.

The Corpus Encoding Standard (CES) is one attempt at extending and customizing tags and the DTD to address the encoding of linguistic corpora, which are becoming larger and larger as current research projects (e.g. MULTEXT, EAGLES) expand their areas of interest and depth of focus. This brief introduction to the CES outlines the main principles

being considered, and then provides a list of Web sites for further information.

NANCY IDE and JEAN VÉRONIS, Codage TEI des dictionnaires électroniques [TEI Coding of dictionaries]; pp. 170–176

The purpose of the Dictionary Working Group is to establish a set of TEI standards for coding dictionary entries; other components, such as headers and footers, preliminary matter, and so on, are the same as for other types of documents. Preliminary guidelines have been distributed to research groups around the world, to test and improve them. From an initial test base of modern western dictionaries, it is hoped that the work can expand to include massive works such as the *Oxford English Dictionary* and the *Trésor de la Langue Française*. The basic TEI principles appear to be robust enough to deal with the specific case of dictionaries, although the possible need for new tags and attributes should not be discounted.

PATRICE BONHOMME, FLORENCE BRUNESEAUX and LAURENT ROMARY, Codage, documentation et diffusion de ressources textuelles [Coding, documentation and distribution of textual sources]; pp. 177–180

In the social sciences community, including computational linguistics, electronic text sources often form the basis of research. The diversity of such sources, their content and structure, complicates processing of their data. This article briefly presents some solutions which TEI (an SGML application) can offer; as well, some examples produced by Project Dialogue [the authors are part of this project] show the results of processing TEI-encoded data.

Keywords: corpus, textual sources, server, alignment, SGML

[from the Resumé]

STÉPHANE HARIÉ, ÉLISABETH MURISASCO, JACQUES LE MAITRE and JEAN VÉRONIS, SgmlQL : un langage de requêtes pour la manipulation de documents SGML [SgmlQL: A query language for manipulating SGML documents]; pp. 181–184

SgmlQL is a programming language developed by the European MULTEXT project, which allows the manipulation of SGML documents, and in particular, of TEI documents. Based on SQL, it makes possible complex manipulation of documents: extracting portions of an SGML document according to specified criteria; tests, counts, and other calculations on SGML elements; constructions of new elements, based on query results.

The article then provides some examples, discusses some current implementations, and refers the reader to Web sites for further information:

```
http://www.lpl.univ-aix.fr/projects/
    multext/MtSgmlQL
http://www.lpl.univ-aix.fr/projects/
    multext/
```

Sommaire des derniers *Cahiers GUTenberg* parus [Summary of recent issues]; pp. 185–187

A listing of the contents of issues 16 through 23 of the *Cahiers*.

[Compiled by Christina Thiele]

Articles from *Cahiers* issues can be found in Post-Script format at the following site:

```
http://www.univ-rennes1.fr/pub/GUTenberg/
    publicationsPS
```