Looseleaf Publishing

A TopLeaf White Paper

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

July 1999

Section 1
TopLeaf Methodology

Introduction

TopLeaf is a powerful and complex publishing system. This White Paper deals with the looseleaf publishing methods provided in TopLeaf Version 5.1. TopLeaf is Y2K compliant.

The more general aspects of TopLeaf publishing are the subject of another White Paper *TopLeaf Publishing.*

Using TopLeaf

TopLeaf V5.1 runs on Windows95 or NT4 or later. It is implemented in two parts, the TopLeaf Graphical User Interface, and the TopLeaf Applications Program Interface.

The TopLeaf GUI gives a menu-driven interface which enables an operator to make use of all the TopLeaf functions, which it calls via the TopLeaf API.

The TopLeaf API contains the whole processing engine (including the looseleaf processing dealt with in this White Paper). The API is implemented as a DLL file which may be linked into third party applications to allow them to call TopLeaf functions directly.

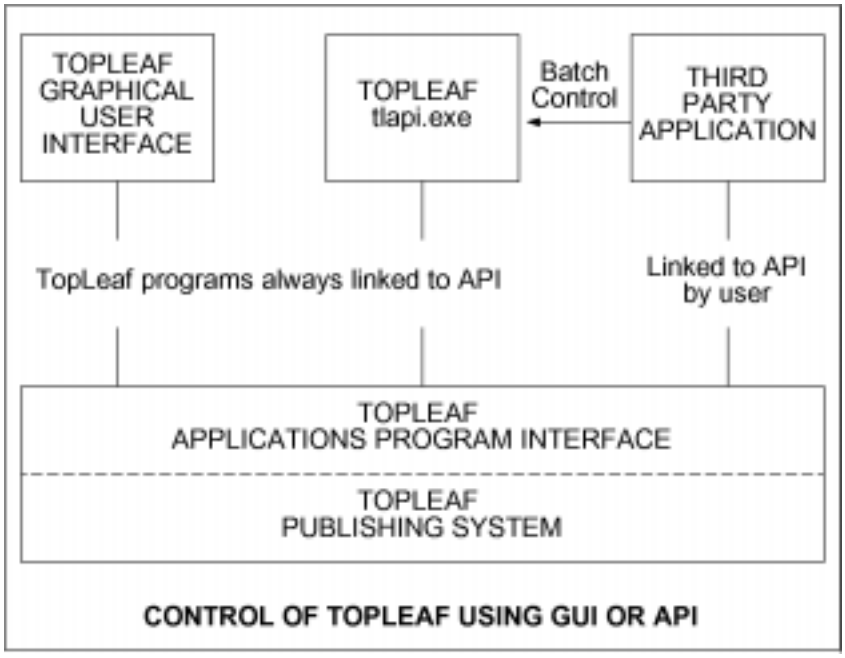
In addition to the GUI and API, a utility (Tlapi.exe) is provided to allow command line calls to be made to the API by a third party application without the need for linking it with the DLL.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au



Characteristics of looseleaf publications

Where a publication needs to be kept up to date by regular changes to its contents, looseleaf methods enable these changes to be incorporated in a timely and controllable manner, and published at minimum cost.

In a looseleaf publication successive leaves are not bound together as in an ordinary book but are held in a ring binder, enabling leaves to be replaced or removed, or new leaves added, as required. Minimum cost is achieved if leaves are replaced only when their content changes.

Looseleaf publications mostly contain reference material subject to continual change. Some are subscription services (e.g. legal) sold to users who must have access to current information in their field. Others are more specific, such as technical documentation of maintenance procedures for aircraft, where the material relates to individual models and so is tied logically to the sale and maintenance of particular equipment.

The economics of regular subscription services requires strict cost control of the number of replacement and new leaves sent to subscribers at each

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

update. As an update proceeds, editorial and production controllers need to be supplied with exact and timely statistics on the current number of leaves in the update, to help decide on the amount of material to be included.

The economics of maintenance documentation is different, being essentially part of the production cost, and often placing more importance on convenience of use and accuracy.

All looseleaf publications require strict discipline to be applied to consistency of style. They typically have a long life span, and new pages made today may have to be inserted among pages made years ago and still look the same. Eventually style changes may have to be made, and the system should be designed to allow them to be made as economically as possible.

For economy and accuracy, looseleaf systems should produce the output of an update automatically, ready for despatch to the subscriber or user, with clear instructions on how to merge the update material into the publication.

TopLeaf input data for looseleaf systems

General requirements

The looseleaf system should impose no special constraints on the structure or storage of the input data for a publication.

Ideally, input data structure and storage should be common for all possible uses to which the data will be put. It is now standard practice for a publication to be published electronically on the internet and also as a paper product, in which case both versions should be made from the one data source. It is assumed that the data structure is defined by a Document Type Definition (DTD) and marked by SGML or XML tags, in such detail that the data is sufficiently well defined for the purposes of each required output.

Another trend in publishing systems is to make a logical distinction between the actual stream of data applied as input to the programs and the manner in which this data is stored. This distinction seems trivial where the data stream results from the sequential reading of a file or group of files. Here the publication's source data may be held in a file which is processed into the required output format. If this file is edited with prescribed changes, the next update is made by processing the edited file.

The distinction is not trivial when a publication's input data stream is extracted in a specified sequence of records from a Data Repository — a set of databases holding text and other data, structured for more general purposes

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

than just publishing. Here the sequencer dynamically builds the publication's input data stream; editing of records is part of the normal database updating, and is totally independent of any of the publication data streams generated from the Repository.

The move towards Data Repositories requires a dramatic change in looseleaf processing. The essence of looseleaf processing is that once a leaf is published it is thereafter regarded as a separate entity from adjacent leaves. If at the next update the input data for the published leaf has been changed, the effect of these changes is not allowed to flow on to the next leaf. If the amount of data is reduced, a replacement leaf is produced to represent the new status of the leaf in the publication, and this leaf will end with data that flows naturally to the start of the next published leaf which remains unchanged. If on the other hand the edit adds extra material, the published leaf may have to be replaced by two or more leaves, and the last of these replacement leaves will end with data that flows naturally to the start of the next original leaf. To achieve this, the program needs to have knowledge of the boundaries in the input data stream where the published leaf starts and where it ends.

This is commonly achieved by the insertion of marker tags into the data stream by the program as it makes the leaves during an update; these markers then show the leaf boundaries when the update is declared to be the current published version of the publication. The marked input is then subjected to editing in preparation for the next update, and the markers are present in the input data stream when the edited material is made into leaves for that update.

The above scheme is easy to implement when the input data stream is stored as a file or sequence of files — the leaf processing has only to write a copy of the input files unchanged except for insertion of marker tags, then return this to the system for use in the next editing.

It is not however a suitable method when the input data stream is dynamically generated from a Data Repository. Generally managers of Data Repositories will not compromise the integrity of the Repository by insertion of such foreign material as marker tags into their records. In any case, the insertion of multiple markers when the records are used in several different publications would rapidly become unworkable.

TopLeaf eliminates this problem. Incorporating one of the most significant advances in looseleaf processing, TopLeaf now holds the necessary status information on the currently published release of a publication to be able to reconstruct leaf boundaries in an input data stream, without there being any marks at all in the data stream. This means that when a data stream is

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

dynamically made from the Data Repository, there is no back flow of information from the publishing processing to the Repository. Therefore all Repository manipulation, including editing of records, is totally uninfluenced by the looseleaf processing, and there is no marking of leaf boundaries in the Repository records.

Using TopLeaf with a Data Repository

The following figure ‘‘TopLeaf with Data Repository’’ outlines how TopLeaf may be used by a Third Party Application program to maintain a number of looseleaf publications where the input data is held and edited in a generalized Data Repository.

Each publication is to be moved through the normal sequence of updates with update parcels produced for each update intended to be distributed to the publication’s subscribers.

It is assumed that data may be shared among publications, but without any interaction of leaf boundaries from one publication to another. TopLeaf treats each publication separately, and has no awareness of any data sharing.

It is assumed that all data editing is performed totally within the control of the Data Repository. For a typesetting run, TopLeaf is sent the input data stream assembled for the publication using the data extracted from the current state of the Repository.

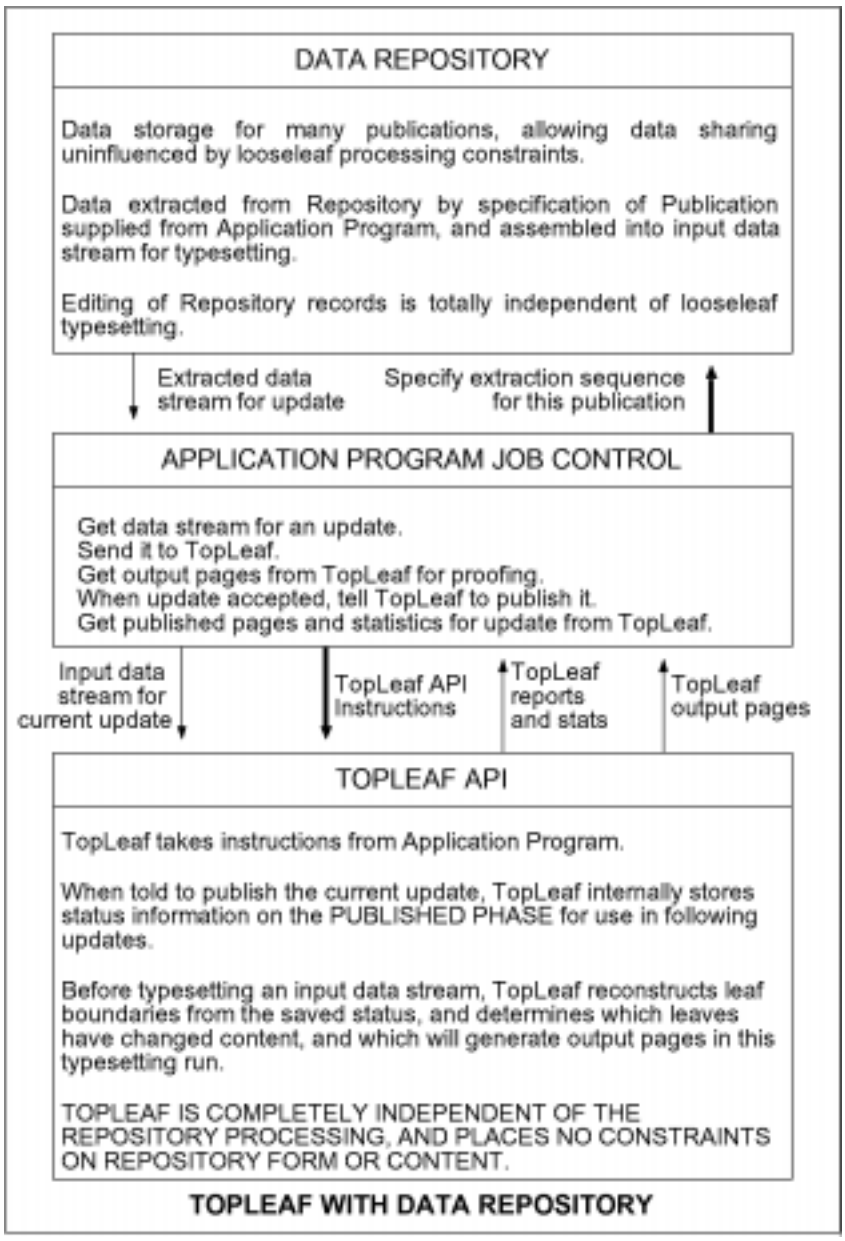
It is also assumed that there is no back flow of information from the TopLeaf processing to the contents of the data held in the Repository. This means that TopLeaf is to be run in its *markerless input* mode, the innovation that makes it possible for the looseleaf typesetting to run without interacting with the Repository.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au



TOPLEAF WITH DATA REPOSITORY

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

TopLeaf looseleaf output media

The traditional output medium in looseleaf publishing is a parcel of paper leaves containing changed and new material for an update, destined to be sent to holders of the main ring-bound volume, with detailed instructions how obsolete leaves are to be replaced by the leaves in the update parcel. TopLeaf generates this parcel, including the filing instructions, automatically as the product of the update. Strictly speaking, what TopLeaf produces is a set of print masters generated from PostScript page files; the actual parcels for subscribers are made in a normal copying or printing run.

TopLeaf produces, as a byproduct of an update's processing, SGML coded fragments with indexing and table of contents information, ready to be combined with similar output from other TopLeaf runs to make consolidated indexes for groups of associated publications.

TopLeaf supports a paperless distribution mode for the update parcels and also the full current version of the publication, using normal electronic publishing methods on the internet. For this, the TopLeaf PostScript output files are converted to PDF files, for transmission across the internet to the subscriber.

In addition, TopLeaf can produce its output files in HTML form, for inclusion on the publisher's web site, for direct browsing by subscribers.

These methods confer an invaluable benefit on both publisher and subscriber, in getting the latest information to the subscriber more quickly, by cutting out the delay due to printing and despatch of the paper parcels.

Internet distribution holds great promise for the future of looseleaf publishing. Even now, most looseleaf users would already have the internet connection needed, and be using the freely available Adobe Acrobat viewer to display PDF files and print them on their office printers.

In addition to the HTML direct browsing option, the subscriber may be given two ways of getting pages from the publisher's web site:

1. Print the downloaded PDF parcel to a laser printer, and update the ring-bound volumes exactly as if the paper parcel had been sent by the publisher, or
2. Hold the whole volume as a PDF file, and allow this to be updated automatically by each update's PDF parcel (using an update program supplied within the parcel). The subscriber may then read the up-to-date volume on his computer screen, and print out for reference pages of immediate interest. This method has not only the advantages

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

of promptness and economy, but also removes a bugbear of traditional looseleaf publishing — the occasional reluctance of subscribers to update the ring-bound volume promptly and accurately when the new update parcel arrives.

TopLeaf also supports a variant of the above paperless distribution, in which the PDF files are distributed by CD. This is particularly appropriate for new subscriptions.

TopLeaf internal file system

TopLeaf maintains an internal file system to hold working data and status and control information for all publications it is set up to process.

The internal file system follows the natural structure of paper publications. For ease of administration, TopLeaf presents a user-defined grouping of the work called the Department level. The user may set up any number of departments, each to hold a number of publications. The naming of departments is arbitrary, but would normally be meaningful, e.g. Trades, Technical, Science. See figure ‘‘TopLeaf Internal File System’’.

To bring a new Publication into the system, the user assigns it to a Department (say Trades) and then chooses a suitable name (e.g. Plumbing) for it. Publications vary widely in size, from say a book of ten chapters to one that may first have to be divided into several volumes. Mostly, it is convenient to divide publications into smaller working units, such as Volumes, and from there into Partitions (e.g. chapters or groups of chapters).

TopLeaf allows a publication to be divided into any number of levels. For smaller publications three levels are usually appropriate (Department, Publication, Partition) and for larger publications four levels (Department, Publication, Volume, Partition).

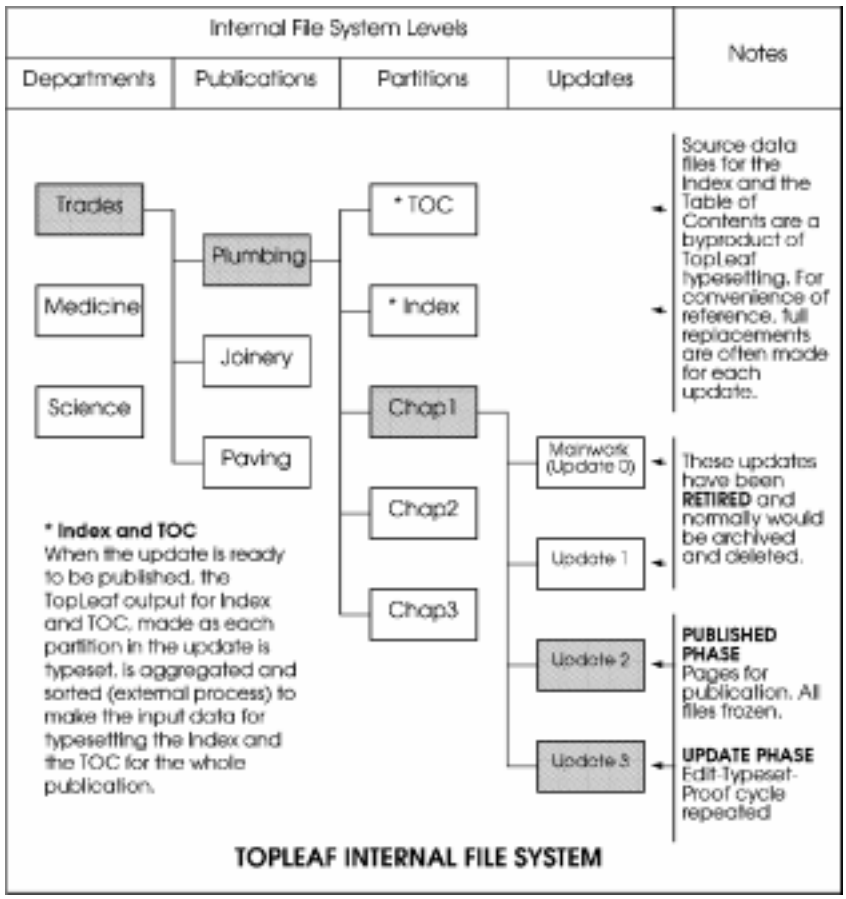
The significance of the partition is that it has real data associated with it, and is typeset as a unit, while the higher levels specify typesetting style and other processing control information.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au



Two more steps are needed to make a working publication. The first is to assign to the publication a set of control instructions to specify its typesetting style and how this is to be interpreted from the SGML tags and the DTD. These controls apply to the whole publication, passing down through any intermediate levels that may be present (such as volumes) to the partitions where they are finally applied to the data. In most publications some volumes and some partitions require a different style from the rest of the publication (e.g. indexes or tables of contents). To accommodate this, control information may be set up within any of the individual volumes or partitions to override any style set up at a higher level. Setting up style and other control information for a looseleaf publication is a significant task.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

The second step is to assign the data to its partition i.e. to enter data directly in the partition or (more commonly) to transfer externally held data into the partition's data area within the TopLeaf system. In practice a publication will usually be built one partition at a time, by creating a new partition within the publication, then transferring its data from the source into the new partition. If the control files have been installed, all the TopLeaf functions (typesetting, page viewing, proofing, data editing, statistics) are now available. All reference to the partition is by its name, making it unnecessary for an editor or operator to become familiar with the actual file system structure.

The above outline holds good regardless of whether TopLeaf is run in the GUI mode, or in API mode under control of a third party application.

In GUI mode however, the system is self-contained, all data and associated files being held totally within the internal file system, for the life of the publication. Editing of data files for an update is an external process, but is called from the GUI, and the edited files are immediately returned after an editing session to the internal file system for proofing and further editing.

In the API mode the third party application is responsible for control of the current update data, including editing. The current data for a partition is sent to TopLeaf for the looseleaf typesetting process for the update. Output leaves for printing and status information for the update are usually sent back to the application upon demand, and there is no need for any of this update data to be held other than temporarily within the TopLeaf internal file system. TopLeaf does however hold in its internal file system all the currently published status information and other data it needs for appropriate typesetting of the current update. This includes the status information it needs to reconstruct the leaf boundary positions in the current update's edited data.

Making the mainwork

The first stage in the real life of the partition is to typeset the data into leaves for the *mainwork* (or Update 0) which will be sent in its entirety to the subscriber. This follows normal book design except that starting page numbers are allocated to each partition so that it has a block of numbers it can expand into without clashing with the next. In addition, partitions themselves may be divided into a number of sections, each with its pages numbered sequentially, but with page number gaps between one section and the next. Gapping is discussed below.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

Administration of a new publication over its lifespan

Producing the mainwork (Update 0)

Each partition of a new publication is initialised independently of the others. The data may be transferred from files elsewhere in the system, or directly entered by keyboard. In either case the data has to be checked for correctly parsed markup (parsing is not a function of the looseleaf system but is carried out by the parser built in with the user's selected author/editing tools). Next the partition should be typeset and the output viewed and proofed. After the first typesetting, page gapping can conveniently be specified. The cycle of editing, typesetting and proofing should then continue until a satisfactory output is achieved. The finalised partition is now ready to be authorised for publication. The mainwork (Update 0) differs from all subsequent releases in that the whole of the material is printed and sent to the subscriber.

The first event in the lifespan of the partition is the editorial and production decision to authorise publication. TopLeaf freezes the input data files and the output leaf files. Update 0 moves into the *Published phase*, which is the source of published pages to be sent to the subscribers. The frozen input data is also the proper data for making other forms of output, e.g. for display on a web site, if the various forms of output are to be synchronised.

Producing Update 1

The next event is the editorial decision to initiate the preparation of Update 1. TopLeaf first makes a copy of the frozen files under the name Update 1. Editing and typesetting and proofing are then permitted for these files. The new release (Update 1) is said to be in the Update Phase.

The Concept of Phases

From here on there are always two active phases for the partitions: the Published Phase, the source of pages to be sent to the subscriber, and the Update Phase, containing the release at present in preparation (work in progress).

Editing and typesetting of the Update Phase follow the special procedures for looseleaf. The original boundaries between successive leaves are maintained throughout the update. Editing may change some or all of the material between boundaries. When the data is typeset in the Update Phase, the typesetting works on the leaves (i.e. the material between the leaf boundary markers) separately; if the material has changed, the new output may be a

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

smaller leaf, or the leaf may be expanded by the addition of new data so that it makes extra pages. The chief conventional looseleaf principle is that if the input is different, it is to be included in the output of the release.

Retiring a Release

This applies chiefly to the GUI mode of working.

The next events are a repetition of the first two: the decision to accept the data files of the Update Phase (and the corresponding output pages) as the Published Phase for Update 1, and the initialisation of the next Update Phase (for Update 2) from a copy of the now frozen Update 1 files. The move of Update Phase Update 1 to Published Phase Update 1 is preceded by retiring Update 0 from the Published Phase. The retiring Update should at this stage be archived with all control files current at the time. It may then be removed from the active system storage. Note that considerations of timelined data for electronic publishing may necessitate retaining all Updates on the system — adequate storage must then be available.

This pattern is followed for the whole of the publication's life. At any stage, there is a release designated the Published Phase which is the official source of pages to be sent to subscribers, either as a looseleaf release to be inserted into the subscriber's binders, or as a full set of pages to be set up in binders for new subscriptions. There is also the next release in preparation (the Update Phase) which originated from the current Published Phase and is aware of all changes made since then.

Maintaining a consistent appearance

A thorny question in looseleaf administration is maintenance of style consistency. Theoretically, if the control files are never changed, TopLeaf will continue to make page after page in correct style, and generally every effort should be made to avoid changes to the control files. However, after a publication has been in existence for some years, it may be desirable to give it a face lift. If, as is usual, a complete reprint is thought to be too expensive, piecemeal changes from the old to the new may be made if the publication is set up appropriately in the first place. The most important provision is to employ page gapping between partitions and between sections when the partition is first established. If this is done, the new style can be used in an update on those sections which have leaves that must appear in the update. This produces a complete set in the new style whenever a change is demanded. The cost of changing the style is spread over a period and under the publisher's control.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

Numbering pages

Looseleaf publications require special conventions for numbering their pages. When one of a sequence of leaves is amended for an update and expands into two leaves, the additional leaf needs a folio that falls between its parent and the following leaf. A common method uses a stroke notation. If in the sequence of pages

15-16, 17-18, 19-20 (three leaves)

the middle leaf is expanded into three, the resulting pages (now five leaves) are numbered

15-16, 17-18, 18/1-18/2, 18/3-18/4, 19-20

If at a later update the 18/1-18/2 leaf expands, the numbering is extended by another stroke level to

18/1-18/2, 18/2/1-18/2/2

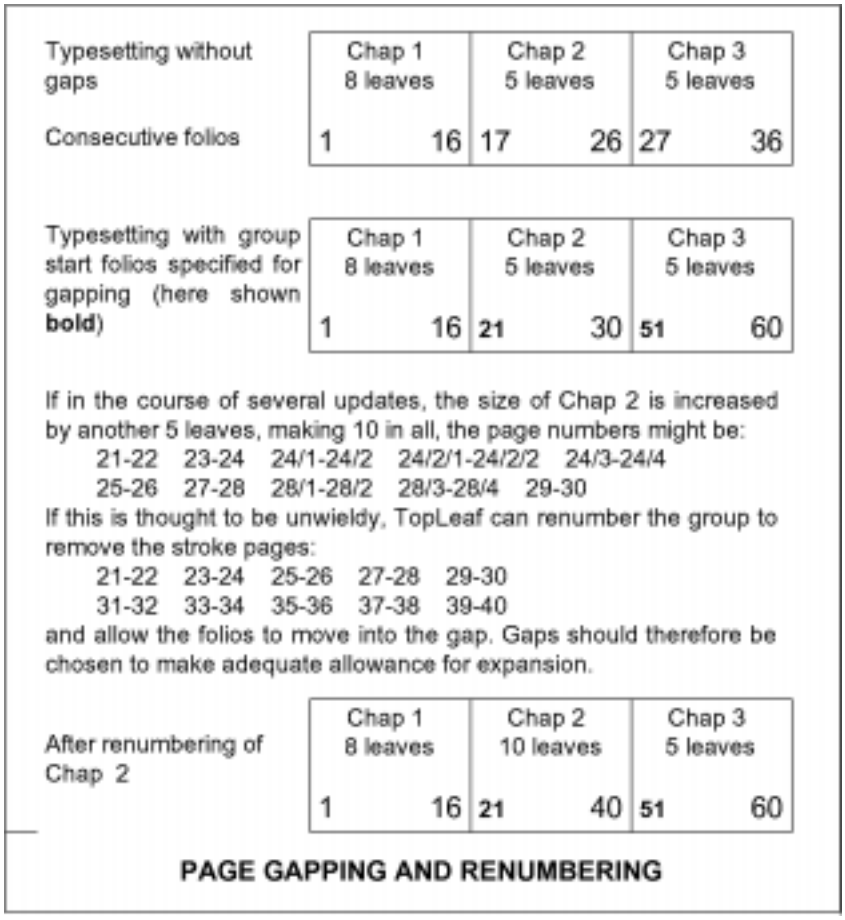
This scheme is usually regarded as satisfactorily intuitive in specifying the proper sequence of the pages in the book. Note however that it cannot deal with all insertions; for example, if at a further update the leaf 18/1-18/2 is expanded into two, there is no acceptable pair of numbers in this scheme lying between 18/2 and 18/2/1. A looseleaf publication requires design not only of the folio numbering scheme but also procedures to be followed in these special cases.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au



To confine stroke page difficulties to local areas in the book, editors may specify page number gaps between sections of the book, to give a reserve of numbers that a section may expand into if required. Such sections must start with a new leaf and finish at the end of a leaf. The mainwork might therefore have a page sequence

...55-56 101-102...165 166 201-202...

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

To achieve this, the input data (control) at the end of the 55-56 leaf must be marked to indicate that a new leaf should be started and that its start number is 101. Page numbering follows this instruction, numbering in sequence until a similar gapping instruction is found at the end of the leaf numbered 165-166, after which following leaves are numbered, as instructed, 201-202...

Link lines

To counter the problem that the reader may assume that pages are missing it is customary to introduce a 'link line' on the last page of a section, saying, for example on page 56:

[The next page is 101]

Link line text and style information are not present in the input data, but are handled automatically by TopLeaf through the control files.

Renumbering

During the course of updating a section, new material may be added into a leaf, resulting in new leaves with stroke page numbers. After many changes the stroke page numbering may become cumbersome and unsightly. To fix this, TopLeaf can consolidate a section, replacing stroke pages by a normal sequence, making use of the unused numbers in the gaps either side. The same automatic renumbering facility may also be used to relocate the section bodily. Renumbering is to be used with care, balancing appearance against cost, for it forces all the renumbered leaves into the current update.

Conversion of existing looseleaf publications to TopLeaf

What is required of the conversion?

Before transfer of an existing looseleaf publication to a highly automated system like TopLeaf is undertaken, much expense and trouble will be saved if a detailed specification is prepared, showing what is required of the new system.

The most difficult conversion is from an existing system in which manual intervention has been possible in page content of previous releases, and where the take-over by the new system is required to be seamless, with no detectible change in the appearance of the pages in the first update from the new system. This method is at first sight very attractive, but may have hidden

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

pitfalls which should be searched for before any conversion work starts. See ‘‘Matching existing pages’’ below.

Planning the conversion

To convert an existing looseleaf publication to TopLeaf may require a change in the markup method employed in the existing data files, and design of the new TopLeaf publication to make pages sufficiently similar to those made by the existing system. For a successful conversion it is important that a detailed design study of both these processes should be undertaken before the conversion is started.

Conversion of the data

If the input data for the existing system is marked in SGML or XML, it is likely no conversion is needed.

However, the most likely markup for older systems is not SGML but a proprietary coding designed for a specific typesetting program or photosetter. Usually these codes may be used as a guide to the document structure, allowing perhaps the major part of conversion to SGML (and generation of the DTD) to be done by program.

TopLeaf SGML style mapping

Once the DTD is available, and the data is marked appropriately, the next stage is exactly the same as if the publication were being set up in TopLeaf from scratch — the detailed design of the pages. This requires the design of the physical page layout, such as specification of fonts and sizes, column measures and depths. It also requires mapping of the SGML tags in the data to relate these typesetting styles to the various structural elements in the data.

Matching existing pages

If a seamless change from the old to the TopLeaf system is required, the detailed TopLeaf typesetting must be identical with that of the existing system. This is not always possible to achieve, especially with older systems using fonts no longer available, and where manual adjustments have been made to individual pages to make them non-standard. It is common experience that in the past compositors have indulged in many ingenious solutions to make text fit into pages (e.g. space padding, and extra-deep columns). Generally a precise page formatter like TopLeaf which follows strict rules cannot (and should not) match such idiosyncratic pages.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

These problems go away if a completely new publication is made in the new system, but this approach is usually prohibitively expensive.

The first requirement in the design of the typesetting conversion process is to determine the seriousness of the above problem. If the problem is acute, it may be more economical in the long run to plan the updates so that a complete replacement of the existing publication is made, not all at once, but gradually, spreading the replacement by sections, as they need to be updated.

Summary

When converting from older systems, it is prudent to examine the consistency of typesetting of the pages of the existing publications. If the pages are not reproducible exactly by the new (automatic) system, it may be less costly in the long run to convert to the style of the new system in manageable sections over a year or so, rather than spend time and effort trying to bend the new system to match non-standard legacy pages.

TopLeaf statistics output

TopLeaf keeps a record for each leaf of the Release Number at which the leaf was last issued as part of an update. Whenever leaf pages are printed, they carry this Release Number. All pages printed as the output of an update carry the current Release Number. If a complete print is made of the partition, each page carries the Release Number at which it was last printed as part of an update.

This information is also provided during the update as a ‘‘live pages list’’, specifying how many pages are contained in the update, providing an important editorial control tool for management of the content and timing of the update.

The same information is also provided for distribution to the subscribers when the output pages are released; here it is organised into ‘‘filing’’ instructions, a list that tells the subscriber which pages are to be removed and which are to be inserted.

TopLeaf identifies Releases numerically in its internal processing, but individual publications may be customised to print a defined character string (usually a date) on the output pages, the live pages list and the filing instructions.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

Section 2

Page management with TopLeaf

Introduction

This section describes some special TopLeaf facilities which have been developed to allow individual annotations to be applied to

specified leaves and groups of leaves. These are mostly to satisfy the needs of document managers using the TopLeaf API.

The significance of the leaf

In most looseleaf processing, the chief significance of a leaf is as an identifiable unit which confines and contains changes to content from one update to the next, so that these changes do not ripple down to successive leaves, as would happen in normal book setting. The usual purpose is to reduce the number of new leaves needed in a given update to bring the whole ring-bound book up to date.

When an update parcel is sent to a subscriber, the leaves are usually present in the update because of some change to their content. If detail of the changes is to be indicated, TopLeaf can mark changed areas with margin rules and a variety of special margin symbols.

For some applications, particularly in the publishing of technical documentation for maintenance of equipment, there is need for a meaningful grouping of leaves, and when changes are made, the need to include the whole amended group in the update distribution. This grouping is simple to achieve in TopLeaf. In these circumstances, there may be need to maintain a history of the changes to the leaves, and for annotations to be made to the leaves to show the history or to give special messages that apply to particular leaves.

The current TopLeaf has been enhanced to allow the storage of historical variables within the status information held on a leaf-by-leaf basis. An example of this might be specifying the date upon which the changed leaf's material is to become effective, with possibly a list of the effective dates of previous changes.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

The effect of this on the system is that leaves or groups of leaves may be treated as individual entities, and their individual update histories maintained and selectively annotated on the output pages.

The single page leaf

TopLeaf provides the option of treating the leaf as a single page, instead of the normal two.

This is essential for true looseleaf processing for publications which by design are printed only on one side of the sheet.

Folios use the sequence 1 2 3 ... for successive leaves. When stroke pages are needed, the numbering is as expected:

e.g. 1 2 2/1 3 3/1 3/2 3/3 4

but, as with the usual two-page leaf, other numbering schemes can be accommodated.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

SGML/XML Tagging Strategy and Looseleaf

This section applies to designing markup for new or legacy publications.

As far as the input files are concerned, the terms SGML and XML are virtually interchangeable. This is because the extra disciplines imposed by the XML standard, such as tag balancing and attribute quoting, are good markup style and should always be adhered to. The differences are more apparent in the DTD design, but that is outside the scope of this section.

The essence of tagging for publication is that the requirements of all desired delivery formats (e.g. looseleaf, CD-Rom, HTML) must be taken into account during the design phase. Most tags, such as levels, headings, paragraphs and the normal inclusions and effects, are required by all output versions.

The tags unique to electronic delivery are mostly associated with link activation.

Standard hyperlinks, popups and applets can all be specified fairly simply. Hardcopy tagging normally involves page-based formatting (such as running heads), and constructs for improving the appearance of the text, such as binding certain items together on a page.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

However, a good tag design can greatly reduce the amount of specialised tagging. Firstly, some functions can be combined. for example, the same tag which italicises a book reference can also trigger a hyperlink or popup to further information on the book concerned. Secondly, most typesetting functions can be inferred without the need for explicit markup. Thus a <head> can automatically bind to the <para> below, eliminating unsightly splitting of material. The TopLeaf system provides in-built support for most such requirements, by allowing sophisticated typesetting constructs to be invoked by simple tags.

As to the question of which material to tag, and which to include as attributes, the general rule is as follows: any material which appears in the body of a page should be keyed directly. Attributes can be used to modify the appearance of text or to contain non-printing information. Printed text, even running heads etc, should *never* be keyed as attributes. Footnotes should be

keyed (appropriately tagged) directly within the originating text, never at the end of a section. This allows for the style of footnoting to be changed as required, or even between delivery modes (e.g. normal page footnotes for print, popups for CD-Rom, and hyperlinks for HTML).

This section has only touched on some of the more general aspects of typeset and multiple mode tagging design. A complete specification requires detailed knowledge of the data, the end products, house styles, and other specialised requirements. Turn-Key Systems is able to provide complete SGML or XML solutions.

SGML/XML Typesetting

TopLeaf produces looseleaf output directly from SGML or XML input data plus DTD. No pre-filtering or translation is required and no intermediate files are produced. This eliminates problems caused by tinkering with proprietary format files ‘‘just to get the page looking right’’, or even worse by systems which translate material into and out of SGML on the fly.

To achieve the desired printing effects, three control files are required. The Style file specifies the overall page layout — height, width, number of columns, plus the location of headers, page numbers etc. The Macro file allows typesetting constructs of arbitrary complexity to be specified as simple instructions. Finally, the Mapping file specifies how the SGML tags should invoke particular typesetting actions.

The mappings can be context sensitive. For example, a <para> within a quote can have a different format from a normal paragraph. The mappings are also sensitive to location, so the first paragraph after a heading can be set to the margin, while other paragraphs have a first line indent. Finally, the mappings can be influenced by attributes, allowing particular styles to be invoked explicitly by the markup.

Combined with the page and release management facilities outlined in previous sections, TopLeaf can provide hard copy in almost any format. Yet because the input remains pure SGML, the ability to output the same material as (say) HTML is not compromised in any way.

Bulk automatic conversion to SGML or XML

Established publishers may have vast amounts of legacy data which needs to be transferred into marked up computer files for continuity of publishing.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

The aim is to devise a procedure that uses as much automatic processing as possible, but the problems are generally so complex and variable that each publication or group of publications has to be separately assessed and appropriate methods devised for it.

First the data structure of the material should be analyzed, and the SGML markup scheme designed as required for the final files after transfer from the existing state.

Next the legacy material has to be analyzed for the presence of distinguishing items that may indicate the data structure as an aid to deciding the required format markup at these places in the text. Old legacy material may not be very helpful in this respect. On the other hand, magnetic tape files of material set by a computer typesetting system might allow a fully automatic program conversion of essential structure.

For printed material, it may be possible to build the files by scanning pages, but this is often a slow process with an unacceptably high character error rate, and re-keying may be cheaper in the long run, especially for the older material. It is also necessary to recapture the data structure as well as the text. For scanning, this means an additional stage to add markup by keying the files made by the scanner, whereas it would usually be possible to add the markup as part of the re-keying process.

The above glosses over some real difficulties of designing effective automatic methods of building the new SGML files. One that shows up in many ways is the (almost inevitable) gradual change of style and printing methods over the life span of the older data.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au

Special looseleaf strategies for converting existing publications

The simplest way for TopLeaf to take over an existing publication is for it to produce an entirely new book, disregarding any existing leaf structure in the input data.

If the publication is already looseleaf, it may be required for TopLeaf to take it over in its existing state and continue issuing releases without any break in service to the subscribers. To do this, TopLeaf needs to be given a file of the input data in which the existing boundaries between leaves are identified.

Special care will be needed in design of the control files if an exact match of typesetting is required for all pages. This is particularly true if the original publication used fonts that it is no longer possible to match, or if compositors

over the years have not observed strict style rules and have allowed deep pages and other arbitrary page fitting methods. It is important to analyze the original publication to estimate the severity of these difficulties. It is undesirable to force TopLeaf to try to match irregular pages, and may often be cheaper to reprint some or all sections rather than to try.

Turn-Key Systems Pty Ltd
ACN 002 858 068

Telephone: +61 2 9906 1577
Fax: +61 2 9906 1342

Suite 203, 511 Pacific Highway,
Crows Nest NSW 2065
PO Box 1216,
Crows Nest NSW 1585
Australia

Email: mailbox@turnkey.com.au
Web site: www.turnkey.com.au