PROCESSING PAPERS FOR JACoW CONFERENCES

John Poole, CERN, Geneva, Switzerland and
JP Scientific (Nantwich) Ltd, Nantwich, Cheshire CW5 6DU, UK

Abstract

Authors submit their papers to a conference for publication in the proceedings. If this is a JACoW conference, then the conference organisers are responsible for the production of files suitable for publication on JACoW from the files submitted by the corresponding author. In general the initial production of raw PDF files is carried out by a team of people who process and where necessary, repair the author's files. This paper will discuss the boundary conditions which are applicable to this file processing and suggest some basic rules.

INTRODUCTION

The processing of papers is often done by a core team of experts assisted by a few trainees. At big conferences this team is brought in at the expense of the conference and it is equipped with fully installed computers, printers and networking etc. This obviously involves a considerable cost to the organisers and this expense is a factor in deciding how to organise the process. JACoW imposes some constraints on the files which it will publish and these must be met or publication will be refused. Further conditions are applied by the conference organisers who may wish to impose certain constraints on the format of the PDF files containing the papers.

CONSTRAINTS

JACoW

JACoW specifies a number of technical constraints which have to be met for publication. These are:

- Paper size
- Page layout (Text centred on JACoW paper size pages, each page to carry the conference name and a page number etc.)
- Performance: each page to display in less than 5/N seconds, where N is the processor speed in GHz
- Hidden fields: Title, author and keywords set inside the PDF file
- PDF file opening parameters: Open with fit-width and thumbnails
- Type3 fonts: to be avoided as far as possible
- Acrobat version
- PDF parameters like correct compression, no hyperlinks etc.

Many of these requirements are not determined by the author: only page(paper) size, the layout of the page, performance and avoidance of type 3 fonts are determined at the authoring stage. From JACoW's point of view, these are the only constraints which have to be met and therefore are the minimum set of requirements that should be checked when processing a paper.

Scripting

If one is going to use the JACoW Proceedings Script Package, then the files should have the correct page layout (sufficient margins top and bottom to fit the header and footer).

JACoW User

The JACoW users want to be able to locate papers of interest and to be able to display and print them easily. The conference concerned should be identifiable from a printed copy of a paper.

Conference

Each conference can set its own additional requirements for the papers. Most conferences will insist on some of the styles to be correctly used, for example, the title should be in upper case, two-column layout etc. This list of conference requirements may be quite long and can have a large impact on the speed of paper processing.

It has usually been the aim to provide feedback to authors following processing in order to improve standards and this is traditionally done using the dotting board. This means that the conference should ensure that there are adequate resources to carry out the processing and provide feedback and help to authors within the timescale of the conference.

JACOW TEMPLATES

The JACoW templates are available to help authors to deliver papers which conform to the JACoW requirements. They also define styles which authors can use to develop their own documents which will comply with the requirements. It is therefore not obligatory that authors use the templates and even those who do, often do not use the styles correctly.

EDITOR OR PROCESSOR ?

In the pre-electronic era, papers submitted to accelerator conferences were submitted as camera-ready and there was therefore no modification apart from scaling. Since the electronic era (1995) electronic files have been processed in order to produce PDF files for publication. In principle, the author submits a PostScript file which is processed with Acrobat Distiller (which applies the JACoW parameters) to produce a PDF file which can be modified to produce the final result on JACoW papersize. This is the ideal situation and is rarely achieved. In most cases some modification of the PDF file is required and this may be done using the PitStop plug-in or Acrobat tools or it may be necessary to return to the source files and modify them.

This methodology is in contrast to an **editorial** process in which the technical content, as well as the layout, is modified or changes are suggested. It is therefore semantically incorrect to refer to the JACoW team that deals with paper contributions as editors.

Because most of the experts (and probably the trainees as well) who work on paper processing have considerable expertise as editors, they notice inconsistencies which, whilst they do not contravene requirements, would not come up to their personal standards if they were submitting a paper themselves.

PROCESSING GUIDELINES

Each conference should produce its own set of guidelines for processing. In general these guidelines go beyond the basic JACoW requirements and they have a strong impact on the time it will take to process papers. I question the value in making these guidelines more than minimal. There is a saying in English that "you cannot make a silk purse from a sow's ear" which means that if the basic materials are not good, then the final product cannot be any better. The JACoW website is there to provide information to the accelerator community and the user is concerned with the scientific content of the papers, not their appearance. Therefore if one cannot correct the content in the papers in cases like the examples below (taken from recent international conferences), then I do not believe that it is worth wasting time on trivial formatting errors.

In the first example below, the text in italics is difficult to understand. The first paragraph is not good English but one can imagine what the author is trying to say. The parts in italics are more difficult.

In example 2, again the majority of the text is not written well and is difficult to understand. The part in italics simply does not make sense.

These examples only concern language problems which, if corrected, would certainly improve the information flow from the author to the reader. However, having format errors like initial capitals in a table caption or in a subsection title probably pass un-noticed by most readers. If the papers are refereed then scientific and language problems can be resolved as part of this process. Paper processing cannot fix the problem of the sow's ear - poor English and dubious scientific content will remain unchanged, no matter how perfect the formatting is.

Example 1

The coupling of RF-components is for new accelerator developments attractive to reduce the costs for RF-amplifier and using more compact devices.
Examples of coupled systems are already present or planned. In the development of the FAIR Proton Injector at GSI coupled CH-DTL cavities are planned and already under construction [1]. The example of an existent system is a coupled RFQ drift tube combination for medical application development at the HICAT (Heavy Ion Cancer Therapy) center in Heidelberg, Germany. *At HICAT center a 4-Rod-RFQ with a 2 gap rebuncher is merged by Institute for Applied Physics, University Frankfurt [2].*
Coupled systems, in this case RFQ and IH-DTL (Fig. 1) with the same resonance frequency can be driven in 0 and π-mode. A switch between the 0 and π-mode needs an extra drift. *In case of the FRANZ-combination are shown the investigation for the 0-mode and the drift between RFQ and IH-DTL is 60 mm [3].*

Example 2

Proton beam-induced cell death is identified as apoptosis [6]. However mechanisms of proton beam-induced cell death is not fully clear. At the present, invivo experiment is more and more important and an essential factor. In Korea, mechanism study of proton beam-induced apoptosis has stayed at the in-vitro level. *Because it is not enough as the Korea has machine for a research and in vivo experimental devices*. In this research, we composed in-vivo experiment device adding ridge filter type modulator, range shifter, collimator, bolus etc. at LEPT (Low Energy Proton Therapy) beam line of MC-50 cyclotron (Korea Institute of Radiological and Medical Sciences). In this paper, we confirmed that size of mouse tumor decreases with in-vivo experiment device by proton beam. And we confirmed the possibility of the developed in-vivo experiment device for its application to in-vivo experiments in the field of biomedical sciences.

THE IMPACT OF PROCESSING CRITERIA

The basic requirements for JACoW publication must be met, so processing must check page layout, paper size, fonts and performance. Any further processing requirements are at the discretion of the conference organisers. There are a number of environmental factors which affect the time it takes to process papers, like the hardware and networking but the level of detail which people fix during

processing has a major impact. Conference organisers invest heavily to bring a team of people to the conference to process papers as well as providing the informatics infrastructure for them to work with. The rate at which papers can be processed is therefore an important driving factor in the cost to the conference.

Experience has shown that a team comprising a core of experts can process an average of about 35 papers per person per day (IPAC11 was much lower than this, closer to 20 papers per person per day), if the processing criteria are not too restrictive and the informatics infrastructure performs well. This level of performance will fall dramatically if the criteria are too particular.

It is essential that the whole team processes papers in the same way and applies the same criteria. These criteria must therefore be defined even before deciding on the number of people, computer and meals etc. that will be required and certainly before processing starts.

I would classify the criteria into several categories:

- Must have for JACoW
- Must have for the conference
- Would be nice to have for the conference
- In an ideal world would have

The JACoW criteria have already been mentioned above and they also include a number of further requirements like no page numbering on the original. *Must haves for the conference* might include having the title in uppercase, author list correctly formatted, no section numbering, a restriction on the number of pages, one or two columns and so on. *Would be nice* might include sub-sections with initial caps, captions correctly formatted, references correctly formatted and so on. *In a ideal world* the paper would follow the template to the letter so that missing full stops (periods) in captions are inserted, figure captions have a colon after the number followed by a single space etc.

The *must haves* are the normal baseline and correspond to the normal processing rate. In such a scenario, if the person processing the file needed to open the source file to fix something, then he/she would also take care of the *Would be nice* errors.

The instructions to processors normally say that if it is not necessary to open the source file, then only the *must haves* need to be fixed. The first problem for someone working on a paper is deciding whether or not it is necessary to open the source file to fix something. There are many ways to fix problems and if the file can be made acceptable using PitStop and/or the Acrobat TextTouchUp tool, then the temptation to fix trivial problems is not there.

On average it takes around 15 minutes to process a paper but a problem-free paper can be processed in two or three minutes. Applying very strict criteria will at least double the time it takes to process a paper, pushing the time that some less-experienced processors take, close to one hour. This should be taken into account when setting the time limit for processing beyond which the paper should be red-dotted. In principle, the decision to red-dot a paper or not should be taken early on the basis of an estimate of how long it will take to process the paper. Increasing the number of format problems to be fixed in the papers will lead to an increase of the time spent during the initial 'expert' phase at conferences like IPAC by around 20% (estimated from the number of green dot papers at IPAC11). Having editors at the conference, equipping them and housing them etc. leads to a cost of around 100 Euros per paper. However, this financial argument is less important than the simple argument about scientific content versus aesthetic appearance.

PROPOSED CRITERIA

For smaller conferences the level of detail to be fixed in papers is not such an issue because a couple of experts can easily handle 400 papers during a one-week conference and there are usually more than just 2 working full-time on processing. The issue arises for conferences on the scale of IPAC and NAPAC where one aims to process all papers before the end of the conference and perhaps even have pre-press publication.

I propose the following outline structure for the processing criteria at large conferences.

1. All JACoW requirements (papersize, margins, fonts, display performance and Acrobat version)
2. All conference requirements (number of pages, format etc.)
3. Optional conference requirements

This would appear to be quite straight forward but the difficulty is in the implementation. Traditionally, at IPAC and its predecessors it was a rule that non-essential corrections would not be applied if it was not necessary to open the source file. This decision depends on the experience and skills of the person processing the paper - many things can be fixed using only PitStop and TextTouchUp by an expert. Of course, certain things cannot be corrected in the PDF or it is inefficient not to open the source file to fix them. If it is required by the conference that these errors are fixed during processing, the source file will have to be used, it will take longer and a yellow dot will be assigned.

It would be useful to give clear guidelines in writing to the team before they start processing concerning what they are expected to fix and under what circumstances. The appendix lists some examples of how some faults can be fixed and this might serve as a basis for some training of the team which would lead to more equality in the way files are processed. My preferred set of guidelines which is very close to that used for IPAC and its predecessors, is listed below. Application of these guidelines and the use of the same processing methods by all of the team will result in a more homogeneous set of PDF files.

- Must fix:
  - Papersize, margins, fonts, display performance (assuming that the Acrobat version is determined by the software installation).

  - Remove page and section numbering.
  - Uppercase title (and/or force lower case where it should be like MHz or GeV).
  - Font sizes.
  - Line spacing if it is less than single (it should be exactly 12pt).

- Only fix if the source file is open:
  - Section and subsection formatting (caps or initial caps etc.)
  - Caption, figure and table alignment (no paragraph indent, centred/justified as appropriate)
  - Obvious typos (mis-spelled titles etc.)
  - Replace 'Fig. caption text' by 'Figure: Caption text' and likewise for Tab. caption text. Leave Fig. and Tab. in the text.
  - Reference formatting

- Fix if it can be done in less than 2-3 minutes:
  - Author and institute list order or layout.
  - Missing full stops (periods).
  - Space around headings etc. if it can be done without exceeding the page limit.
  - Equation number alignment.
  - Initials should be on the same line as the name
  - Quantities and their units should be together on one line (e.g. in 9 GeV)

These criteria are based on the JACoW template and should therefore evolve as the template evolves. The underlying principle is that the PDF files must meet JACoW requirements and present a reasonably homogeneous appearance.

CONCLUSION

Publication through the JACoW system is not the same as publication in a scientific journal: it provides a channel for publication of technical information in the field of accelerators with easy access to state-of-the-art information. Accelerator conferences have always tried to publish their proceedings rapidly and the papers submitted are considered to be 'camera-ready'. These fundamental principles should be kept in mind when defining what authors are required to do in terms of layout and formatting of their papers.

The use of a well chosen set of guidelines, such as those proposed in this paper, by the team responsible for producing the initial PDF files will facilitate efficient processing. The choice of guidelines which supplement the JACoW technical requirements, should be made by the scientific management of the conference, not JACoW. I also think that it would be useful to have a training/briefing session with the team before the start of processing to share techniques and to explain the guidelines so that everyone in the team will approach the task in the same way and then they will produce a more homogeneous set of PDF files. This method will lead to optimised costs and rapid feedback to the authors as well as rapid publication. The latter is highly appreciated by authors and JACoW users alike and is one of the many benefits of electronic publication. In my view there is little point in correcting any more than the *Must fix* items because it will not add scientific value to the paper although it may be aesthetically more pleasing to look at.

APPENDIX

SOME PROBLEMS THAT CAN BE FIXED WITHOUT OPENING THE SOURCE FILE

Mixed Case Titles

One of the most common problems which has a high visual impact is that the title is not set in upper case. This can be fixed either by using the Acrobat TextTouchUp tool or by generating a dummy paper which contains just the correct title (from a WORD or LaTeX skeleton) and then making a PostScript file which is distilled. The title can be cut and pasted from the dummy PDF to replace the incorrect title using PitStop. This is relatively quick and ensures the embedding of the correct fonts in the final PDF. This method also allows the vertical re-alignment of blocks of text to make space for the title which may occupy more space once it is in uppercase.

Page and Section Numbering

It is simple to remove the numbers using PitStop but one has to be sure that section numbers are not referenced in the text. If they are, it will probably be necessary to open the source file.

Footnotes

Badly aligned footnotes can be moved using PitStop as long as they are single column and sufficient space can be made available. A missing separator line above the footnote can easily be inserted by a copy and paste from a dummy PDF file (using the JACoW template, for example).

Minor Spelling Errors

The TextTouchUp tool is well suited to correcting many spelling errors, although it may be necessary to adjust the alignment afterwards using PitStop.

Localised Use of Incorrect Fonts

The TextTouchUp tool can often be used to fix this and it is possible to force embedding of the font in the PDF file.

Figure Caption and References Alignment

Figures can easily be moved/scaled with PitStop. It is also possible to split each line of a badly formatted caption into words and adjust the spacing from centred to justified - this is OK for one or two shortish captions but is tedious for longer ones and is better done in the source file. It is often possible to make the references appear acceptable by moving text with PitStop.

Hyperlinks

The appearance of hyperlinks can be removed by using PitStop to remove the underline and TextTouchUp tool to change the colour.

Page Offsets

In some cases (usually LaTeXpapers) the text has a vertical offset on the page. This can easily be fixed using a PitStop Action List (Up 9mm, Down 18mm, Left 3mm etc.) and it is useful to have these available to everyone during processing.

A4 Paper Printed on US Letter and vice-versa

In theory this can be fixed in PitStop by scaling and shifting the text but is more often done by opening the source file and using the correct generic printer. An Action List should be developed to aid in this.

SOME PROBLEMS THAT REQUIRE THE SOURCE FILE TO BE OPENED

Missing or Bad PostScript File

This was the case for around 15% of the papers at IPAC11 and in these cases, there is no option but to open the source file.

Captions

This was the most common problem area at IPAC11 and it can be very time consuming to fix these problems because the layout of the paper changes in an unacceptable way.

General Format Problems

Errors like incorrect spacing between column or wrong column widths have to be fixed in the source file and often lead to further layout problems.

All Type 3 Fonts

This is quite a common problem for WORD papers, LaTeX seems to be better than it used to be and it is rare to find Type 3 fonts now in a file generated from LaTeX.