ELECTRONIC FORMS IN THE HEWLETT PACKARD ENVIRONMENT

Barry Gillespie
Indigo Software Ltd.
560 Rochester Street
Ottawa, Ontario K1S 5K2
Canada

1. INTRODUCTION

This paper will discuss the use of electronic forms in the Hewlett Packard environment. An "electronic form" is a means of printing formatted business data combined with graphic elements **without using pre-printed forms.** The "Hewlett Packard environment" includes the HP Vectra PC, HP3000, HP9000-300, HP9000-800 and HP1000 with one of the family of HP LaserJet printers. These printers, (LaserJet Plus, LaserJet 500 Plus, LaserJet Series II and the LaserJet 2000), will be referred to by their collective name "the LaserJet". The original basic LaserJet is not included because it lacks sufficient memory, vector graphics capabilities, macros and font variety to make it useful for anything but the simplest of forms applications. This paper does not directly address the HP2680 and HP2688 high end laser printers, though many of the ideas presented here are applicable.

The use of electronic forms will be examined from several perspectives:

1) What are the benefits of electronic forms compared to using pre-printed forms?

2) What tools are required for electronic forms design?

3) How can the capabilities of the HP LaserJet be used to optimum advantage?

4) Are there tasks that cannot be accomplished with pre-printed forms that are now possible with electronic forms?

2. BENEFITS OF ELECTRONIC FORMS

There are many benefits that can be derived from using electronic forms instead of pre-printed forms. Some of the benefits produce direct cost savings. Other benefits are related more to operational flexibility, efficiency and productivity.

The first and most obvious benefit of using electronic forms is **reduced consumables costs** as compared to using pre-printed forms. This is especially true with the LaserJet 2000 which has consumable costs of only $0.01/page. In addition, duplex printing can cut paper costs in half. Printing data on both sides of pre-printed forms with an impact printer is not feasible.

No investment in pre-printed forms inventory is required. First, there is the cost of the pre-printed forms themselves. Payment must be made immediately for forms that sit on the shelf, sometimes for months. With electronic forms **there is no forms inventory or forms storage costs.** Finally, there is no need for staff to manage the shipping, receiving and handling of the forms inventory.

With interactive graphics forms design software **the lead time to put new forms into production is greatly reduced.** This is often of critical importance in today's rapidly changing business environment. Designing a form takes much less time if you use an interactive design tool. The savings that a draftsman can gain by using Computer Aided Design (CAD) tools has been well documented. Automated forms design provides similar benefits to the forms designer. These time and cost savings are even more dramatic when simple revisions must be made. Instead of going through four to six weeks of re-drafting and reprinting, the new version of an electronic form can be available in minutes.

The fact that operational forms are constantly changing introduces another area of cost savings. **Obsolete forms never have to be thrown out.** In most organizations, between five and twenty percent of all pre-printed forms end up in the garbage.

Electronic forms are **much easier to distribute.** They are usually simple binary files that can be distributed over normal communications links very easily. All branches of a business, whether "down the street" or "down under", can be updated simultaneously when the new forms are ready to go into production. The old version of the form can be deleted or archived. There is no possibility of users inadvertently using outdated forms.

No operator intervention is required to change forms. All form changes are controlled automatically and only plain, white stock paper is ever loaded into the printer. This means that expensive hardware resources are never sitting idle waiting for physical operator intervention. It also means there may be no need for an operator at all.

Since the operator is not required to change the forms, there is no chance they will mount the wrong form by mistake. **This reduces errors and wastage.** In addition, the scheduling of which print job runs next becomes much more flexible, since the forms mounting requirement has been removed.

When the financial and operational value of all these benefits is calculated, the cost justification for buying laser printers and the required electronic forms software becomes obvious. Financial payback in a year or less is not uncommon.

3. ELECTRONIC FORMS SYSTEMS - REQUIRED CAPABILITIES

3.1 FORMS DESIGN

There are currently two different types of tools being used to design electronic forms: code based and interactive forms design systems.

Code based tools require the designer to create the form specification in a specialized language. The code-based approach has several disadvantages. The first and most obvious is that the designer cannot see the results of his work as he goes. He must visualize it as he generates the form design, and hope that the result matches what is wanted. Often, several rounds of trial and error are required before the desired result is achieved.

Code based design tools require a skilled coder in order to be used successfully, especially if the form is at all complicated. Specifications of several hundred lines are not uncommon. Making changes to these long, complex coded forms specifications is difficult and time consuming.

Interactive form design programs solve all of these problems. With a good WYSIWYG (What You See Is What You Get) interface, the designer can see the form take shape on the computer screen as it is drawn. Wordy commands do not have to be coded. Often all that is required is a few clicks of a mouse button. Forms can be created much more quickly, with fewer test passes required to achieve the desired results.

A skilled coder is not required. The forms designer who formerly used a drafting table can quickly learn to use a well designed interactive graphics design tool. The analogy to a draftsperson moving from a drafting table to a CAD system is obvious. Corresponding increases in productivity will follow.

Many products that have been designed for Desk Top Publishing are also being presented as forms design tools. This is misleading. A product designed to handle the text intensive task of typesetting a newsletter, a long report or manual is ill-equipped to design forms. A form is usually largely a graphic entity. A good interactive forms design tool will be much more like a CAD system, with added capability of handling the much more complex text and font requirements of many common forms.

Most forms are set in a sans serif font like the Helv font available with several font cartridges and downloadable fonts available for the HP LaserJet family of printers. An interactive forms design tool must be able to represent the space taken by text accurately on the screen. Seeing the text on the screen in the actual font at the proper point size, boldness, etc., is also desirable.

A string of text with a mixture of regular, bold and italics, perhaps in different point sizes, should be stored as a single logical text item. This is important because during the life of the form words in the string may change. If the string is stored in pieces, much shuffling around will be required to implement a simple change.

Four different types of text alignment are required: flush left, flush right, centered and justified. An interactive forms design system must be able to handle all four types of text alignment -- even when the text contains multiple fonts in different point sizes. The capability of controlling the white space between lines is also essential.

A facility for handling raster images is required. Logos, signatures and small pieces of "artwork" occur on almost all forms. The ability to handle such images is essential. They can either be drawn with a paint program or scanned using a device like the HP ScanJet.

Users often ask why they cannot just scan in an entire form. A letter size (8.5 by 11 inch) form scanned in at 300 dots per inch (the resolution of both the ScanJet and the LaserJet) takes about 1 megabyte of memory. Even when raster compression techniques are applied, the scanned form will take up to thirty times the memory required to represent the same form in PCL (the printer control language used by the LaserJet). Every time the form is printed, this large raster image must be transmitted to the printer. The speed requirements of electronic forms applications usually make the production use of scanned forms impractical.

The use of scanned forms also does not allow for easy revisions. The form is stored as a single complex raster image. Individual items on the form (words, lines, boxes, etc.) cannot be modified without re-drawing the form and re-scanning.

All forms are usually drawn on a grid, typically six lines per inch vertically and ten characters per inch horizontally. A forms design product should allow the user to place all items on a built-in grid. This makes corner and T junction alignment foolproof. Items that must appear directly under one another or beside one another can be easily aligned. A single set of grid values is not sufficient. Complex forms often have multiple sections set using different grids. The fine details of text placement often require use of a much finer grid. The capability for the user to change the built-in grid settings at any time during the design of the form is desirable.

The forms design system requires high level graphics facilities like those found in a CAD product, especially if complicated forms are to be drawn. Multiple layers of zoom, pan, group moves, copies and deletions all fall under this feature category. The ability to do global changes makes experimentation with alternative form designs much easier (e.g., change all text strings in a selected area from 8 point Helv to 10 point Helv).

The power and graphics capabilities of the HP Vectra ES/12, RS/16 and RS/20 make them ideal computers for the design of forms. Some arguments could be made for using an HP9000/300 machine, but the extra cost would be hard to justify. The 12 to 20 Mhz processors in the PCs will typically be more than adequate.

Up to this point only the design of the form template, the fixed portion of the form, has been discussed. A complete electronic forms system should

include the capability to print the electronic forms merged with a data file to fill in the variable data fields.

The design system must allow the user to specify where on the form the variable data fields should be printed. It should allow for fields printed in different fonts, so that important pieces of information can be highlighted by bolding or by using a larger point size. It must allow for right justification of numeric data for decimal alignment. The ability to word wrap and justify paragraphs of text data is also desirable.

3.2 ELECTRONIC FORMS PRINTING CAPABILITIES.

Several methods of merging data files with the electronic forms are necessary.

The simplest form of data file merge is where the data file is simply printed overlaid with the electronic form. The data file must be formatted correctly such that the data fields are printed in the correct locations on the form. This method is exactly analogous to printing the data file onto a pre-printed form.

Simple sequential processing should be supported. The first piece of data will go into the first field on the form, the second piece of data into the second field on the form, etc. Each data value can be delimited by a carriage return/line feed, or by some special delimiter character. If this delimiter is a comma, the Basic Data Interchange Format that most PC database packages use would be supported. The sequence for filling the fields on the form can be determined by simply scanning the form for data fields top to bottom and left to right.

This method has been used by the HP Boise division to produce the catalog of third party software products that support the LaserJet. Each page of the catalog is in reality a form, a collection of data formatted on the page along with lines, boxes and fixed text labels. The advantage of this method is that the actual camera-ready production of the hundreds of catalog pages can be delayed until the last moment. Full attention can be paid to making sure the data stored in the product database is accurate. As a final step, the camera-ready pages of the catalog are printed on the LaserJet by merging the data in the database with the electronic form.

In a data processing environment, that is common on minicomputers such as the HP3000, the data files are often in a fixed record format. For example: Columns 1 to 10 contain the Purchase Order Number, 11 to 40 contain Vendor Name, etc.. The capability to extract the data fields from their specified location in the data record and then print them in the correct position on the electronic form is desirable. Support for multiple different record formats in a single data file is also desirable for more complex forms.

The electronic form system should be able to print bar codes as well as regular text. Bar codes can be required both as part of the form or for printing variable data fields. Bar codes can be treated as a special type of

font. Bar codes are composed of a simple collection of vertical lines of varying widths. The LaserJet's vector line drawing capabilities can be used. The electronic forms product must be able to handle the different formatting rules for each bar code format supported, including how to generate the standard start and stop characters and the check digits required. Data that appears on bar coded labels for shipping, receiving and inventory applications can be merged with label "forms" allowing labels to be printed, on demand, using LaserJets right on the factory or warehouse floor.

In some applications it is desirable to include graphics images as part of the data to be printed on the form. An example might be a real estate listing which includes a picture of the property listed. The capability to include graphics images as part of the data is desirable for advanced electronic forms systems.

4. TAKING ADVANTAGE OF THE CAPABILITIES OF THE HP LASERJET

The LaserJet has several features that make it an excellent electronic forms printer. It also has some limitations that should be taken into account when designing forms.

4.1 GRAPHICS CAPABILITIES

PCL (Printer Control Language) is the language of the LaserJet printer controller. A system that takes full advantage of the PCL orthogonal line drawing and box drawing capabilities can greatly reduce the size of the printer data required to print the form. PCL's shaded box capability (with multiple levels of gray scale and other shading patterns) allows the form designer to use shaded areas on the form with only a few bytes of printer data.

It has often been stated that a megabyte of printer memory is required to do full page graphics on a LaserJet. If the entire form is printed as a large raster image this would be true. A forms design program that takes advantage of the vector graphics capabilities of the LaserJet can greatly reduce printer memory required. For example, the first page of a 1040 tax form can be printed with **less than 20K of printer memory** (not including soft fonts).

Problems can arise if a form has graphic elements that include arcs (rounded corner boxes, circles, etc.) or diagonals, since the LaserJet's vector graphics capabilities do not include diagonal lines or arcs. These types of objects must be printed using raster graphics. A large number of these objects can significantly increase the size of the PCL file needed to print the form, and reduce the speed of printing. Keeping the radius of the corner arcs small and using only short diagonals will minimize the impact of this problem.

4.2 FONTS

Three types of fonts are supported on the LaserJet; built-in fonts (i.e. fonts that are permanently available in the printer), cartridge fonts, and "soft"

fonts (i.e., fonts that initially are stored on disk and can be loaded into the printer). A good forms design tool should be able to handle all font types and allow the user to mix and match fonts from all three sources. The ability to use fonts from two or three different cartridges simultaneously is also required, in order to take full advantage of the LaserJet Series II and LaserJet 2000.

Most forms are designed using sans serif Helv fonts. In all of the printers in the LaserJet line only one built-in Helv font is available (LaserJet 2000 only). This shortcoming will hopefully be corrected in future generations of printers from HP. Fortunately, the choice of fonts available with font cartridges and soft fonts is excellent. The best choice from the list of available font cartridges would seem to be the HP92286Z cartridge. This cartridge includes 8, 10, 12 and 14 point versions of Helv as well as TmsRmsn (a serif font that can be used on forms with lengthy paragraphs of text).

Helv 6 point is common on many forms, especially on more complex forms. Some forms may even need 4 point text. Some forms require large 18 or even 24 point text for titles and section delimiters. These fonts can be easily obtained as downloadable soft fonts from HP's library or from the wealth of third party fonts and font generation programs available.

A forms design system should allow for the use of any downloadable font in standard HP format. It should be able to represent them on the screen correctly at design time and download them automatically at print time. This automatic download should be optional, allowing the user to load the fonts at the beginning of the day and not pay the price for constantly reloading fonts every time a form is printed.

Choosing the right combination of cartridge and soft fonts may take some careful planning. There is a way to avoid having to use downloaded fonts at all. HP and third party vendors will make custom font cartridges on demand, putting all of the fonts required on one (or more) cartridges.

4.3 USE OF MACROS

Probably the most important feature of the LaserJet for electronic forms applications is its macro capability. A macro is a collection of PCL commands downloaded in the printer memory, ready for activation at any time. Applications can be developed where all the forms required can be loaded into the printer as macros at the beginning of the day. At production print time the application program need only send the data. In this way maximum throughput can be achieved.

One problem that may occur with this method is that the total size of all the required form macros may be too large because of the amount of repeated raster data. These raster images could be logos, signatures, etc.. If the same 10Kb logo occurs on 30 different forms it will take up 300Kb of printer memory. Raster graphics processing is also one of the slower processes in the LaserJet.

This problem can be solved by converting the raster images into a font or set of fonts. Splitting them into groups by size, so that the font "cell" size can fairly closely match the size of the images is advisable. There are several utility programs available today that perform this conversion process quickly and easily. The converted logo font(s) can be loaded with the regular soft fonts. Each logo is now stored only once in printer memory. Font processing is also somewhat faster than raster image processing. The forms will take less printer memory and will also print more quickly.

It is clear from these discussions that no simple answer can be given to the common question, "How much memory do I need in my laser printer?". The answer will depend on a careful analysis of how many forms are required, how complex the forms are, how many different downloaded fonts are used, etc.. The best plan is to build the forms first and determine the memory required -- leaving room for future expansion. Many applications can run successfully with the base 0.5 Mb memory of the LaserJet Series II. Other applications may utilize the fully loaded 5.5 Mb memory of the LaserJet 2000.

4.4 USE OF COMPILED FORMS

Any Hewlett Packard Vectra PC, HP3000, HP9000-300 or HP9000-800 computer can be used in conjunction with electronic forms. Which computer is used will usually depend more on where the data resides than on anything else. The forms are created by the design software on the PC. Before the forms are moved to the production print environment they should be compiled. The electronic form compiler converts the form from the design file format to PCL. The compiled form also contains information about the data fields. The compiled form is a simple binary file that can be transferred to any other computer using standard binary file transfer utilities.

Forms compilation has two major benefits.

A pre-compiled form prints more quickly. A good analogy would be the difference in execution speed of executing a program from source code or from object code.

The second advantage is more subtle but of great importance to large organizations. Compiled forms allow corporate MIS and forms design groups to maintain control of the content and format of forms used in the company. The corporate form design and utilization standards can be maintained. The end users simply use the compiled forms on their production computer. They cannot change them. End users should never need or have access to the original forms design files.

The LaserJet is an ideal forms printer. Its special capabilities make it faster in a production environment than comparable printers with the same "rated" speed. It can be attached to any of the HP computers and used to produce electronic forms merged with data on demand.

5. ELECTRONIC FORMS EXTEND FORMS PROCESSING

Electronic forms allow for the processing of data in several ways that are not possible with pre-printed forms. Multiple part forms no longer need to have data fields in the same positions on all parts of the form. Forms with header and repeating detail pages can be created. Forms where all pages are not printed for a particular set of data are feasible. Finally, the template of the form can vary, depending on the data itself. The best way to explain each of these four extensions to traditional forms processing is by examples.

5.1 MULTI-PART FORMS

The separate parts of a multiple part form are actually variations on the same form. Each part has a slightly different format. For example, a three part purchase order has the three separate parts labelled: accounting copy; receiving copy; and vendor copy. The vendor copy has a special terms and conditions paragraph. The receiver's copy does not include the price and order quantity information. The degree of variation possible on pre-printed forms is constrained by the fact that the actual position of the data fields must be the same on all parts. With "electronic forms" this constraint is removed. Data fields can appear in different positions on different parts. Data fields can be printed in different fonts or even suppressed altogether.

5.2 REPEATING PAGES

Forms like invoices, purchase orders, etc. that have a large block of header data at the top followed by multiple lines of detail data at the bottom are inherently wasteful if a large number of detail lines are required. Examine what happens using current pre-printed forms technology when an invoice with one hundred items is produced. The first page contains the required header information on the top half of the form. The bottom half of the form will contain perhaps fifteen lines of detail item information. On the second page the top half is left essentially blank, except for a page number. This is wasted paper. Only another fifteen lines of detail information can be printed. This process continues until all one hundred details have been spread over seven pages.

With electronic forms only four pages are required. The first page is identical to the first page in the pre-printed form example. The second page, however, can have a different format. All the header data fields are removed. Only detail line item data appears on the second and subsequent pages. The number of detail pages used is controlled by the amount of data. In our example, the second and subsequent pages might contain thirty detail lines each instead of fifteen. The same one hundred lines of detail data will be spread over only four pages instead of seven. Continuous pre-printed forms cannot be used because the number of detail pages varies depending on the amount of data.

5.3 MULTI-PAGE FORMS

Many data processing operations require multiple page forms. In some cases, not all pages need to be filled out. A simple example of this would be an employee application form that has a page for the employee's armed services record. Only employees who actually served in the armed service have to fill it out. In all other cases this page would be left blank, wasting paper and filing space. With electronic forms this page can simply be skipped. The data file for the process can contain commands to alter the normal flow of processing and skip the unused page except when it is needed.

A more complex example is found in the insurance industry. Many policies have thirty or more related forms associated with them. Typically there is a header page of general policy information and many optional pages. Which optional forms are used for a particular client depends on which policy options are chosen. A forms processing program can have all the required forms loaded into the memory of the LaserJet as macros and then print only the forms required for each individual policy. This procedure can be controlled by allowing page selection commands to be imbedded in the data file. These commands will skip to the required page in printer memory before sending the data for that page. This type of multiple page processing cannot be accomplished using pre-printed forms.

5.4 DYNAMIC FORMS

The final example of sophisticated electronic forms processing is what might be called "dynamic" forms. It is similar to the previous multiple page example. In this case each section of data may not occupy a full page. An auto insurance accident report is a good example. Typically it will have several distinct sections; where the accident occurred, when it occurred, the weather conditions, and sections for each vehicle damaged and each person injured.

With pre-printed forms a limited number of vehicles and injured people can be entered on the form. Additional entries for these sections must be appended to the form in attached lists at the back of the policy.

A dynamic forms processor will create each page of the form based on the amount of data for each section. Sections that do not apply will be omitted. Each potential section of the form can be stored as a separate macro. Instead of placing elements of the section template in fixed positions on the page as is done with full page form macros, each section will be compiled using relative positioning. The last PCL command for each section will put the page cursor at a fixed relative reference point, so that the next section will join with the current section correctly.

Before placing a new section and its corresponding data on the page the dynamic forms processor will first determine if there is enough vertical space. If space is not available the program will eject the current page and start a new page.

Using this approach for the accident claim report will allow all vehicles involved in the accident and all people injured in the accident to be printed together sequentially. The number of pages for an accident report will vary. All data will be recorded in the required format. The dynamic form is an improved solution for this type of application.

The examples given here are some of the many ways forms applications can take full advantage of the text and graphics capabilities of the LaserJet. These same ideas can be applied to many other data processing situations in the Hewlett Packard environment.

6. CONCLUSIONS

The concept of using electronic forms to replace pre-printed forms is still relatively new. The potential is enormous. Every large business and government organization uses hundreds and often thousands of forms to control the flow of information within the business and with other institutions. The concepts presented here can substantially improve this process.

Very significant cost savings can be realized. The ability to respond to changes is much improved. The use of forms in a distributed organization is optimized. Operational flexibility is improved, with reduced errors.

In many organizations the major benefit will be that applications that cannot be implemented using pre-printed forms become feasible with electronic forms.