NBS
PUBLICATIONS

NBSIR 84-2836

Document Interchange Format

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
Institute for Computer Sciences and Technology
Washington, DC 20234

February 1984

Issued April 1984

Prepared for:
**Department of Navy Office Automation and
Communications System Project (DONOACS)**

NBSIR 84-2836

DOCUMENT INTERCHANGE FORMAT

Joan E. Knoerdel
Shirley Ward Watkins

February 1984

Issued April 1984

Prepared for:
Department of Navy Office Automation and
Communications System Project (DONOACS)

ABSTRACT

In the absence of standards, both the private and public sectors have addressed document interchange among different vendors' text processing systems in a number of ways. In an attempt to solve this interchange problem with respect to encoding of control functions for the Department of the Navy, a project was originated by the Office of the Under Secretary of the Navy for Financial Management to determine the formatting requirements of the Department of the Navy and to translate those requirements to a representation that would be supported by text processing system providers. The encoded representation of the formatting control functions has become known as the Document Interchange Format (DIF).

This paper describes the overall approach taken by the DIF and then provides definitions for and implementation details for DIF. The body of the paper is intended for overall understanding and as such is intended for managers and technical staff. Then, there are a number of appendices provided which are specifically written for those implementing DIF.

Key Words: ANSI X3.64; code standards; control functions; control sequence introducer; DIF; document interchange format; document interchange standards; text processing systems; word processors.

TABLE OF CONTENTS

INTRODUCTION

Communications is a vital function of the office; and so it is not surprising that an essential component of text processing systems is communications support. There are a number of elements required to support communication; not only do transmission speeds and line protocols have to be selected, but also the encoding of the information to be exchanged must be agreed upon.

Wide area networks, local area networks, PBX systems and direct hard wiring of text processing systems all represent reasonable methods to connect text processing systems. The Institute for Computer Sciences and Technology (ICST) at the National Bureau of Standards participates with a number of national and international voluntary standards organizations for the development of the computer network protocols. These protocols which are being developed in the framework of the Reference Model of Open Systems Interconnection will provide for the movement of data among open systems and include such protocols as: Internet, Transport, Session, Presentation, Message Transfer, File Transfer, Virtual Terminal and others at the application layer.

While the above network protocol standards ensure the reliable transfer of data it is still highly probable that the data sent will not be correctly interpreted by the receiving system(s). In the realm of text processing, a document is composed of printed characters and control information (used for formatting a document). While standards organizations are addressing the document interchange problem at a number of different levels, today there is no standard which addresses even a majority of the control information required by text processing systems.

In the absence of standards, both the private and public sectors have addressed the document encoding problem in a number of ways ranging from the purchase of black box translators to limiting procurement of text processing systems to a single manufacturer. In an attempt to solve this interchange problem with respect to encoding for the Department of the Navy, a project was originated by the Office of the Under Secretary of the Navy for Financial Management to determine the formatting requirements of the Department of the Navy and to translate these requirements to a representation that would be supported by text processing system providers.

Technical advisory services to agencies is included under ICST's Brooks Act mandated responsibilities. Therefore, when the Department of the Navy approached ICST and requested technical assistance especially with respect to the representation (encoding) of the formatting requirements, ICST agreed to develop such a representation for them. This representation has become known as the Document Interchange Format (DIF).

This paper describes the overall approach taken by the DIF and then provides definitions for and implementation details for DIF. The body of this paper is intended for overall understanding and as such is intended for managers and technical staff; there are a number of appendices provided which are specifically written for those implementing DIF. Appendix A provides key definitions. Appendix B describes the actual DIF specifications for control information. Appendix C provides a listing of functions by class or category. Appendix D indicates any

special rules associated with the issuance of formatting functions. Appendix E lists the control codes permitted in DIF and supported by the American National Standards Institute (ANSI) existing ASCII code table. Appendix F describes the encoding of control functions already standardized by ANSI and ISO and Appendix G provides the encoding of new control functions.

Approach

There are three phases to the Department of the Navy DIF project: define Navy formatting requirements; specify a representation; and provide a demonstration of document interchange using DIF. This paper discusses phase 2, the representation.

A number of manufacturers elected to provide inputs to the Department of the Navy project. From the project initiation in March 1983, seven companies have provided inputs on the technical specification: Data General, Datapoint, Digital Equipment Corporation, Four Phase, Sperry Corporation, Wang and Xerox. In addition, the specification has been discussed with ANSI X3V1 and ANSI X3L2 and their comments have been incorporated to the fullest degree possible. Interactions continue with these two committees.

The approach taken by NBS was to use existing coding standards where possible and to recommend extensions to existing standards in cases where format requirements were not being addressed. Based on ANSI standards X3.98, X3.64, and ISO 6429, 13 functions were found to be already standardized. It was decided to use the framework provided by ANSI standard X3.64 to encode the remaining 29 functions. ANSI X3.64 (which is also published as Federal Information Processing Standard (FIPS) 86) defines a set of control functions that augment the set of control functions in ANSI X3.4-1977 (ASCII), using the principles defined in ANSI X3.41-1974, Code Extension Techniques for use within 7-Bit Coded Character Set of ASCII.

X3.4 defines a basic code table for 7-bit ASCII which is represented as a matrix consisting of 8 columns and 16 rows (see Figure 1). Columns 0 and 1 contain the control (non-printing) characters and represents the C0 (control) set; columns 2 through 7 contain human readable (printing) characters and represent the G0 (graphic) set. While the G0 set is generally implemented by text processing systems as specified by X3.4, there is a certain liberality taken with the C0 set. Therefore, the only allowable C0 characters in DIF are those listed in Appendix E. The means of specifying a particular character from code tables is to indicate the column number followed by a forward slash followed by the row number; therefore, to indicate an "A" from the basic code table, the representation is 4/1.

X3.64 provides for expanding the control functions as defined by the C0 set. The standard describes three controls to achieve this expansion and DIF uses the one known as Control Sequence Introducer (CSI). The general form of the CSI control function is:

CSI   P...P   I...I   F

where:

CSI is a 1/11 followed by 5/11 in the 7-bit world; or the character sequence
"ESC["

P...P are parameters ranging from 0 to a maximum restricted only by implementation;
all characters are taken from 3/0 to 3/9 inclusive.

I...I are intermediate characters used to expand the set of functions defined
by F; characters are from 2/0 to 2/15 inclusive.

F is the function defining character and comes from 4/0 to 7/14 inclusive.

There are two reasons why CSI was selected as the technique for expanding the
control set: the ability to associate a parameter with a control function, and
the ability to consolidate into a special set. The DIF specific functions make
use of an intermediate character. Currently, 2/4 is used as the intermediate.
There are two types of parameters which may be associated with a function:
numeric and selective. A numeric parameter uses characters 3/0 through 3/9,
inclusive and the parameter assumes the value of the corresponding character,
which for the basic ASCII code table is a decimal digit 0 through 9. As an example,
in DIF the Left Margin Set function has an associated numeric parameter such
that the value of the digits of the parameter represent the decimal value of the
left margin setting. Note that a parameter may assume a value greater than 9.

A selective parameter is also taken from 3/0 to 3/9 inclusive; however, even
though represented by a string of digits, it is not interpreted as a number. The
meaning is listed with the definition of the function. As an example, in DIF the
Page Start/End function has an associated selective parameter such that if the
parameter has a value of 1 "start" is indicated, and 0 an "end".

With this introduction to ASCII and the existing standards relevant to the development
of DIF, the next section will describe all the functions defined by DIF. Functions
previously defined by existing standards will be indicated and definitions of all
functions will be provided.

Column / row	0	1	2	3	4	5	6	7
0	NUL	DLE	SP	0	@	P	`	p
1	SOH	DC1	!	1	A	Q	a	q
2	STX	DC2	"	2	B	R	b	r
3	ETX	DC3	#	3	C	S	c	s
4	EOT	DC4	$	4	D	T	d	t
5	ENQ	NAK	%	5	E	U	e	u
6	ACK	SYN	&	6	F	V	f	v
7	BEL	ETB	'	7	G	W	g	w
8	BS	CAN	(	8	H	X	h	x
9	HT	EM	)	9	I	Y	i	y
10	LF	SUB	*	:	J	Z	j	z
11	VT	ESC	+	;	K	[	k	{
12	FF	IS4	,	<	L	\	l		
13	CR	IS3	-	=	M	]	m	}
14	SO	IS2	.	>	N	^	n	~
15	SI	IS1	/	?	O	_	o	DEL

Figure 1   ASCII Code Table

4

DIF FUNCTIONS

The forty-four control functions defined by DIF have been divided into six classes according to their primary purpose. The six classes are Break functions, Document Format functions, Page Format functions, Line Format functions, Rendition functions, and Miscellaneous functions. This section begins with general descriptions of each class providing distinguishing features of the class, special issuance rules, and a listing of each individual function within the class, along with an indication of which functions are already defined in standards. (Appendix F should be used to determine the specific standard in which definitions occur.) Following this introduction, each function will be more extensively defined, again by class (Appendix B should be used for the DIF techincal specification of each function).

Break functions are those which have the side effect of terminating the parameter value for a given function and invoking a new parameter value for that function. They can also invoke the initial parameter value for a given function. The permissible sequence for functions in this class is the formatting function followed by the Break function, without any intervening text. This sequence was adopted to avoid any problems with decisions as to when the new parameter values of the formatting functions should go into effect. The specific Break functions are Hard New Line, Hard Page End, and Hard Page Start. Hard New Line is the only one of the three which is currently standardized.

Document Format functions include those functions which pertain to the appearance or positioning of text throughout the entire document. The functions defined within this class are Justify, Left Align, Line Height, and Pitch. Each of these functions must be issued prior to the first Hard Page Start only, meaning their parameter values may not be changed within the document. All four of these functions have been defined in existing standards

The Page Format funtions specify characteristics pertaining to each page of the document. Included in this class are Top Margin Set, Bottom Margin Set, Footer Margin Set, Header Margin Set, Left Facing Page Header and Footer Left and Right Margins (four functions total), Right Facing Page Header and Footer Left and Right Margins (four functions total), Page Length, Page Width, Page Numbering Set, and Page Orientation. Like the Document Format functions, the parameter values for these functions must be established prior to the first Hard Page Start and may not be changed thereafter for the entire document. Only Page Orientation has been previously standardized.

The Line Format functions apply to the positioning of text on a line. The functions in this class are Center, Decimal Tab Set, Text Tab Set, Left Margin Set, Right Margin Set, Line Spacing, and Temporary Left Margin Set. Unlike the previous two classes, these functions may be issued throughout the document; however, except for Center and Temporary Left Margin Set, their initial parameter values must be specified prior to the first Hard Page Start. Several of the Line Format functions are examples of functions which must be followed by a Break function; i.e., Decimal Tab Set, Text Tab Set, Left Margin Set, Right Margin Set, and Line Spacing. Of the functions in this class, Center and Text Tab Set already have standard definitions.

The Rendition functions apply at the word and character level; they deal with the appearance or imaging of individual words and characters. This class covers Emphasis, Overstrike, Subscript, Superscript, and Underscore. In each case, the function is applied to the text immediately following the activation of the function and remains in effect until explicitly turned off. All five of the Rendition functions are standard functions.

The last class of functions is reserved for those which are not general enough in their functionality and issuance to fit into any of the other five classes; thus the name Miscellaneous functions. The following functions fall within this category: Document Start/End, Footer Start/End, Header Start/End, Hard Hyphen, Soft Hyphen, Hard Space, Perform Decimal Tab, Perform Text Tab, and Page Numbering Initial Value. As mentioned, the permissible occurrence of these functions varies; Document Start/End may occur once only as the first and last control functions of the document, Footer Start/End and Header Start/End must be issued only once prior to the first Hard Page Start, and Page Numbering Initial Value must also be issued a single time before the first Hard Page Start. The remaining five functions in this class may occur wherever appropriate in the body of the document after the first Hard Page Start. An ASCII character was chosen as the representation for Perform Text Tab; however, none of the others currently have standard representations.

The following section is devoted to a brief description of each of the forty-four functions and any relevant interactions among the functions.

Break Functions

The three functions in this class are formatting functions which have a dual role. The primary purpose of a Hard New Line is to indicate a mandatory new line. Hard Page Start indicates that any following text must begin on a mandatory new sheet of paper, while Hard Page End indicates the point at which text should stop being printed on the current sheet of paper. Note that there may be an arbitrary number of physical pages between a Hard Page Start and a Hard Page End. The significance of "hard" in each case is that the function must be performed and may not be altered by the system.

The secondary purpose of the Break functions is their interaction with other formatting functions. As mentioned earlier, each of the Break functions can be used to terminate the parameter value(s) for a given function and invoke a new value(s) for that function. More than one formatting function may occur before the Break function, but no text is permitted to be intermixed with the formatting and Break functions.

An example of such an interaction is the issuance of a Temporary Left Margin Set function. The left margin might be set at 12, for instance, and the following paragraph is a quote which is to be indented five spaces. Immediately following the last word of the paragraph, a Temporary Left Margin Set (TLM) function with a parameter value of 17 should be issued followed by the Hard New Line function. The left margin is immediately set (but only temporarily due to the nature of TLM) to 17 for the quoted paragraph. That value remains in effect until the occurrence of the next Hard New Line, at which time the left margin returns to 12.

Document Format Functions

These functions affect the document as a whole, excluding any headers or footers. Their values must be established at the beginning of the document and may not change thereafter. Justify is the process of adjusting lines of text by expansion and compression of white space and by word fill so that the lines are vertically aligned along the left and right margins. If Justify is not in effect, the left margin is aligned while the right margin is ragged. The method of expansion and compression is at the discretion of the receiving system.

Left Align is the vertical alignment of the left margin. This is not meant to imply that all text must be aligned on the same column since the left margin may change; it only implies that the left margin may not be ragged. Therefore, left alignment must always be in effect.

Line Height is used to indicate the height of the characters, or potential characters for blank lines, on a line. Most users think of the line height in terms of lines per inch, which is how the parameter value for this function is specified. The only value allowed for Line Height in DIF is six lines per inch, including both printed and blank lines. Currently, the number of printed lines versus blank lines per inch (i.e., single, double, or triple spacing) is indicated by the Line Spacing function. The reader should refer to the Line Formatting Functions section for a description of that function. Since DIF users have a requirement for triple spacing as well as single and double spacing, the standards committees for coding, ANSI X3L2, and ISO TC97/SC2, has been requested to add a parameter value for two lines per inch to Line Height, implying triple spacing. Values for single and double spacing already exist for that function. Since the functionality of Line Spacing could be subsumed by Line Height, it is likely that the Line Spacing function will be deleted if the requested parameter value is ultimately defined.

The last function in this class is Pitch, which indicates the number of characters per inch. In DIF, only monospaced characters are permitted; the two choices are 10 or 12 characters per inch.

Page Format Functions

Generally, the Page Format functions are applied to every page of the document. The exceptions to that rule are the header and footer left and right margin setting functions for the left and right facing pages and Page Numbering Set. The reason the header/footer left and right margin setting functions are in the exception category is very subtle; that is, these functions cannot be applied to every page of the document since the left facing page functions apply to left facing pages only and the right facing page functions apply to right facing pages only. Note that for centered headers and footers regardless of the facing page the Center function is to be used. However, the header or footer left and right margins for the appropriate facing page must still be specified as the points of reference for the centering to occur.

The other exception, Page Numbering Set, is used to invoke the process of page numbering and is allowed to change one time within a document. It must be initialized prior to the first Hard Page Start if page numbering is to occur anywhere within the document. A separate control function from the Miscellaneous class is used to specify the number which is to be printed on the first page where

numbering is to begin; that function is Page Numbering Initial Value. In explanation, suppose the first page is a title page and page numbering is desired to begin on the second page. Then Page Numbering Set would be initialized to OFF and Page Numbering Initial Value to 2. After the Hard Page End for the title page but before the Hard Page Start for the second page, Page Numbering Set would be set to ON, resulting in page numbers being printed on the document at printout time from page two to the end.

The remaining functions in this class are intended to address every page of a document. Four of them denote the vertical margins on each page; they are Top Margin Set, Bottom Margin Set, Header Margin Set, and Footer Margin Set. The other three, Page Orientation, Page Length, and Page Width, deal with the dimensions of a page.

The topmost vertical margin is designated by Top Margin Set. This function's parameter value indicates the line on which the first printed line is to occur; i.e., the header, if any, otherwise the first line of the body text. The Bottom Margin Set establishes the bottommost line on which the body text is to be printed. Both of these functions use the top edge of the paper, which is line 1, as their point of reference.

The other two margin setting functions in this class, Header Margin Set and Footer Margin Set, are not relative to the top edge; rather, they are relative to the header or footer and the body text. Header Margin Set establishes the number of blank lines between the last line of the header and the first line of the body of the document on a page, while Footer Margin Set establishes the amount of space between the last line of body text and the first line of the footer.

The functions dealing with the page dimensions are used to restrict the overall length and width of a page. Page Orientation refers to whether the paper on which the text is to be presented should be aligned in the printer such that the length of the page is greater than the width (portrait) or the width is greater than the length (landscape). It was decided that for the DIF specification only portrait orientation would be supported. Page Length specifies the maximum number of lines to be printed per page including headers, footers and body text. The maximum setting for a portrait oriented 8-1/2 inch x 11 inch page is 66. Page Width indicates the maximum number of characters per line. The maximum setting is 85 for 10 pitch and 102 for 12 pitch.

Line Format Functions

Within the class of Line Format functions, eight of the nine functions affect the horizontal positioning of text along a line. This includes Left Margin Set, Right Margin Set, Temporary Left Margin Set, Center, Decimal Tab Set, and Text Tab Set. The ninth function, Line Spacing, determines whether the text will be single-, double-, or triple-spaced. As mentioned in the description of Line Height, this function may be eliminated if the ANSI and ISO standards committees on coding agree to define a parameter value for two lines per inch for Line Height.

As expected, the purpose of Left Margin Set and Right Margin Set is to establish the leftmost and rightmost columns, respectively, for the following lines in the body area text. The parameter values remain in effect until explicitly changed.

Temporary Left Margin Set serves to temporarily override the parameter value of Left Margin Set. This function is terminated by the occurrence of a Hard New Line, Hard Page End, or Document End. It is conceivable that the temporary left margin could be to the left of the left margin; however, the result of doing so is device dependent.

Decimal Tab Set and Text Tab Set indicate the tabstop positions for their respective type of tab. Each function accepts multiple values for its parameter value. The issuance of one of these functions has a side effect of clearing out all previous values for that function, so any previous values remaining in effect must be explicitly restated. The functionality of a decimal tab is to align a column of figures along the decimal point. The invocation of a decimal tab is performed by Perform Decimal Tab and a text tab is performed by Perform Text Tab, both classified as Miscellaneous functions.

All of the Line Format functions described up to this point can be changed at any point within a document, and any number of these functions can appear in a sequence with no intervening text. It is mandatory that the final control function in such a sequence be one of the Break functions. The new parameter values of the changed functions are then in effect for any following lines. For example, if a table with columns of textual information is to be embedded in a report, it may be necessary to change the left margin and the text tabstop for the table. After the last character of the paragraph, a Left Margin Set function (with the new parameter value) and a Text Tab Set function (with the tabstops for the table as its parameter values) should be issued. Then, Hard New Line should follow, invoking the new parameter values for those two functions, and inputing of the actual table can proceed.

The functionality of Center is straightforward; it delimits a line of text which is to be centered between two specified points. Those specified points are indicated by the left and right margin setting functions. Center can be applied to headers, footers or body text. Unlike the other functions in this class, center must follow a Hard New Line unless the line being centered is the first line of the document. This restriction is applicable to body text only. For headers and footers, the Center delimiters encapsulate the header/footer delimiters with header/footer text.

Rendition Functions

Rendition in this application can be thought of as the process of applying an imaging attribute to a character or a word. In all cases, the function must be explicitly activated and deactivated. One or more Rendition functions can be applied at the same time to characters or words, although something like overstriking super and subscripted alphanumerics doesn't seem to be very meaningful in a typical office document.

Emphasis is a function which provides a mechanism for indicating that text is to be highlighted in some way. It could be done by bolding, italicizing, underscoring, or blinking, etc. The form of highlighting is at the discretion of the receiving device.

Overstrike could be used to construct composite characters, such as foreign language characters with diacritical marks, or to indicate the replacement of existing characters, such as deleted text in a legal document. However, for DIF it is intended that Overstrike be used only for the latter. The overstrike character is, again, at the discretion of the receiving device. This does not preclude using for the overstrike character the same character as the underlying one, resulting in a bold-like effect.

Underscore is a continuous single ruling between start and stop delimiters. Therefore, spaces, text tabs, and super- and subscripts, as well as characters and words will be underscored. In the case of superscripts and subscripts, the ruling remains on the baseline.

Superscript and subscript cause the active position to be moved to the character position of an imaginary line with a partial vertical offset. For superscript, that vertical offset is upward, and for subscript it is downward. Each function acts as a toggle for the other: subscript returns a superscript to the baseline and vice versa.

Miscellaneous Functions

The Miscellaneous functions provide a variety of functionality. Three serve as delimiters and are grouped in pairs, three are control functions which could be considered printable characters, two perform the action established by other functions, and the last one serves solely as an initialization value.

The three delimiters are Document Start/End, which bound an entire document with all its control information, Header Start/End, which delimit the document header if any, and Footer Start/End, which delimit the document footer if any. The latter two do not have to be included with the document if there is no header or footer. The only permissible characters within the header and footer delimiters are printable ASCII characters and Hard New Line. The rationale for that restriction was that the combined total for headers and footers is limited to 150 characters; thus, more control functions permitted in the header/footer would mean fewer text characters.

Hard Hyphen and Soft Hyphen are control functions which are actually presented as text. The two differ in that the first type of hyphen is mandatory and may not be eliminated by the system nor may a line be broken immediately following it; while the second is discretionary. For the purpose of adjusting line lengths (wordwrapping) most text processors define an area at the end of the lines such that if a word falls within that area a decision must be made as to whether the word remains on the line, moves to the next line, or should be hyphenated. In the case where the word should be hyphenated, either the system or the operator inserts a soft hyphen. During subsequent wordwrapping by a receiving device, or even the same device, that word might not fall at the end of a line any longer for various reasons; the soft hyphen, then, would not be presented by the imaging device.

The Hard Space, too, could be considered as text. It is similar to a regular space, except that it cannot be used as a line break location and is not subject to compression or expansion during the justification process. By way of example, some style guides may require a date to be on one line. Therefore, January 1, 1984, would require hard spaces rather than regular spaces to preserve its format.

Perform Decimal Tab and Perform Text Tab were discussed in conjunction with Decimal Tab Set and Text Tab Set under Line Formatl Functions. Both can appear anywhere in the body of the document to indicate that the active position be moved to the location of the next parameter value established by their respective tab set function.

The last function, Page Numbering Initial Value, was also described earlier, under Page Format Functions. This control function merely indicates the initial value to be printed on the page where page numbering is invoked. The page numbers begin incrementing once they are invoked by Page Numbering Set.

It should be noted that no default values may be implied or assumed for DIF. All parameter values must be explicitly established at the appropriate time as spelled out through this section for those values to be in effect.

CONCLUSION

This paper has discussed the representation for control functions required by the Department of the Navy for text processing systems. It is important to stress that it is not the intent of this effort that manufacturers of text processing systems change their internal representation of these functions. Rather, manufacturers will provide a common format for an agreed subset of text processing functions for interchange only. Thus, a document must be processed by a "filter" program, developed by the manufacturers, which will do the mapping of the DIF control functions from their internal representations to DIF representations on export and the reverse on import.

APPENDIX A

Definition of Terms

Document Related Terms

document: a string of zero or more characters bounded by a Document Start function and a Document End function. In DIF, the maximum document size is 120 pages (i.e., that amount of text which fills 120 8-1/2 inch x 11 inch sheets of paper with a one-inch margin at the top, botton, left side and right side).

footer: A line or lines of text placed at the bottom of a page in a document as an aid to the reader. It usually appears on all pages of the document.

header: a line or lines of characters placed at the top of a page of a document as an aid to the reader. It usually appears on all pages of the document.

left facing page: When printing a document using both sides of the paper and binding the document on the left side (i.e., book style), the page which appears on the back side of the paper and, thus is on the left side when the document is opened up is the left facing page.

right facing page: When printing a document using both sides of the paper and binding the document on the left side (i.e., book style), the page which appears on the front side of the paper and, thus is on the right side when the document is opened up is the right facing page.

text body area: that area available for presenting text, exclusive of headers and footers. It is established by format functions between the Document Start function and the first Hard Page Start function. (Note: the left and right margin values can be changed within the document.)

Classes of Functions

break: functions which have as a side effect the action of terminating the parameter values for given functions and invoking new parameter values for those functions; or, invoking the initial parameter values for given functions.

document format: functions which pertain to the appearance/positioning of text in the entire document; they must be established prior to the first Hard Page Start function and their values cannot change throughout the document.

line format: functions which apply to the appearance and/or positioning of text on a line; the parameter values can change from one line to the next.

page format: functions which pertain to pages; their values must be established prior to the first Hard Page Start function and the values cannot change throughout the document.

rendition: functions which affect the appearance or positioning of words or characters.

Control Functions

BOTTOM MARGIN SET - BMS

BMS is a format effector that establishes the bottommost line of of text body area.

CENTER

Center is a format effector which indicates that a line of print is to be positioned so that its mid-point is aligned with a specific point or line.

DECIMAL TAB SET - DTS

DTS is a format effector which causes a horizontal decimal tabulation stop to be set at each position corresponding to a parameter value.

DOCUMENT START/END - DSE

DSE is a control function which is used to delimit the start and end of a document depending on the selective parameter:

1   Document Start - DS

DS is a control function that indicates the beginning of a document.

0   Document End - DE

DE is a control function that indicates the ending of a document.

EMPHASIS

Emphasis is a format effector which specifies increased intensity of all printable characters between start and stop delimiters.

FOOTER MARGIN SET - FMS

FMS is a format effector that establishes the number of blank lines between the bottom-most line of text body area and the top-most line of the footer text.

FOOTER START/END - FSE

FSE is a control function which is used to delimit the start and end of the footer text depending on the selective parameter:

1   Footer Start - FTS

FTS indicates that the following characters, until the occurrence of a subsequent FTE, are to be treated as the footer of the document.

0   Footer End - FTE

FTE indicates the end of a string of zero or more characters which are to be treated as the footer of the document.

HARD HYPHEN - HH

HH is a connective graphic character that always prints as a hyphen. The system may no break the line immediately after this function. It is not subject to deletion by the system.

HARD NEW LINE - HNL

HNL is a format effector which indicates a mandatory new line that may not be altered by the system.

HARD SPACE - HSP

HSP is like SPACE (2/0), except that HSP cannot be used as a line break and is not subject to compression or expansion when the parameter value of JUSTIFY is equal to ON.

HEADER MARGIN SET - HMS

HMS is a format effector that establishes the number of blank lines between the bottom-most line of the header text and the topmost line of the text body area.

HEADER START/END - HSE

HSE is a control function which is used to delimit the start and end of the header text depending on the selective parameter:

1   Header Start - HS

HS indicates that the following characters, until the occurrence of a subsequent HE, are to be treated as the header of the document.

0   Header End - HE

HE indicates the end of a string of zero or more characters which are to be treated as the header of the document.

JUSTIFY

Justify is a format effector which indicates the process of composing and adjusting lines of text by using word fill and/or inter-word spacing so that the lines are vertically aligned on the left and right margins is to be performed.

LEFT ALIGN

Left Align is a format effector which indicates the setting of lines so that their left end lines up vertically on the left side.

LEFT FACING PAGE FOOTER LEFT MARGIN - LPFL

LPFL is a control function which establishes the leftmost column of the footer on the left facing page.

LEFT FACING PAGE FOOTER RIGHT MARGIN - LPFR

LPFR is a control function which establishes the rightmost column of the footer on the left facing page.

LEFT FACING PAGE HEADER LEFT MARGIN - LPHL

LPHL is a control function which establishes the leftmost column of the header on the left facing page.

LEFT FACING PAGE HEADER RIGHT MARGIN - LPHR

LPHR is a control function which establishes the rightmost column of the header on the left facing page.

LEFT MARGIN SET - LMS

LMS is a format effector which establishes the left most column of the text body area for the current and following lines in the document. LMS takes effect immediately following the next subsequent Break function.

LINE HEIGHT (Select Vertical Spacing)

Line Height is a format effector which specifies the line spacing for subsequent text (i.e., the number of lines per inch).

LINE SPACING - LSP

LSP is a format effector which indicates the presentation of blank lines, if any, between lines of display text.

OVERSTRIKE

Overstrike is a format effector which specifies the superimposing of one character on top of another character in either a constructive (building a composite character) or destructive (replacement of an existing character) manner.

PAGE LENGTH - PL

PL is a control function that indicates the maximum number of lines (both printed and blank) allowed per physical page, including header, footer, and text body area.

PAGE NUMBERING INITIAL VALUE - PNIV

PNIV is a control function that indicates the initial value at which page numbering is to begin when printed.

PAGE NUMBERING SET - PNS

PNS is a control function that indicates whether or not page numbers are to be printed on the pages of the document at printout time (either on a hardcopy or softcopy device).

PAGE ORIENTATION (Page Format Selection)

Page Orientation is a control function which designates the format of the page to be invoked by a subsequent Hard Page Start function.

PAGE START/END -PSE

PSE is a format effector which is used to delimit the start and end of pages depending on the selective parameter:

1   Hard Page Start - HPS

     HPS indicates the point at which the following text is to begin on a mandatory new page (i.e., sheet of paper)

0   Hard Page End - HPE

     HPE indicates the point at which text should stop being printed on the current page (i.e., sheet of paper) and any following text must begin on a mandatory new page (i.e., sheet of paper).

PAGE WIDTH - PW

PW is a control function that indicates the maximum number of character positions allowed per line.

PERFORM DECIMAL TAB - PDT

PDT is a format effector that advances the active position to the next decimal tab stop.

PITCH (Select Horizontal Spacing)

Pitch is a format effector which specifies the character spacing for subsequent text (i.e., the number of characters of a monospaced font that will fit horizontally in one inch).

RIGHT FACING PAGE FOOTER LEFT MARGIN - RPFL

RPFL is a control function which establishes the leftmost column of the footer on the right facing page.

RIGHT FACING PAGE FOOTER RIGHT MARGIN - RPFR

RPFR is a control function which establishes the rightmost column of the footer on the right facing page.

RIGHT FACING PAGE HEADER LEFT MARGIN - RPHL

RPHL is a control function which establishes the leftmost column of the header on the right facing page.

RIGHT FACING PAGE HEADER RIGHT MARGIN - RPHR

RPHR is a control function which establishes the rightmost column of the header on the right facing page.

RIGHT MARGIN SET - RMS

RMS is a format effector which establishes the right hand most column of the text body area for the following lines in the document. RMS takes effect immediately following the next subsequent Break function.

SOFT HYPHEN - SH

SH is a connective graphic character that indicates the position within a word where the formatting program may hyphenate that word if it will not fit on the current line. The hyphen is not presented unless the work is actually hyphenated at that point at printout time (or display time).

SUBSCRIPT (Partial Line Down)

Subscript is a format effector which causes the active position to be moved to the corresponding character position of an imaginary line with a partial vertical offset. This offset should be sufficient to image following characters as subscripts until the first following occurrence of SUPERSCRIPT in the data stream or, if the immediately preceding character is imaged as a superscript to restore subsequent imaging of characters to the active line.

SUPERSCRIPT (Partial Line UP)

Superscript is a format effector which causes the active position to be moved to the corresponding character position of an imaginary line with a partial vertical offset. This offset should be sufficient to image following characters as superscripts until the first following occurrence of SUBSCRIPT in the data stream or, if the immediately preceding character is imaged as a subscript to restore subsequent imaging of characters to the active line.

TEXT TAB SET (Horizontal Tabulation Set Absolute)

Text Tab Set is a control function which indicates the positions along a line where tap stops are desired.

TEMPORARY LEFT MARGIN SET - TLM

TMS is a format effector that temporarily sets the left margin position to the column specified by the parameter value until the next occurrence of Hard New Line, Hard Page End, or Document End.

TOP MARGIN SET - TMS

TMS is a format effector that establishes the topmost line of the header text, if any. If no header is specified, this is also the topmost line of the text body area.

UNDERSCORE

Underscore is a format effector which specifies a continuous ruling (single) between start and stop delimiters.

APPENDIX B

DIF Specification of Control Functions

CENTER

1. In DIF, that specified point is the mid-point between the left and right margins. Therefore, the left and right margin positions must be established prior to issuance of a Center function.

2. This function is Terminated by a JFY parameter for Center OFF.

3. This function is intended to be used to center headers and footers as well as text in the text body area.

4. If the amount of text before a Center OFF function is greater than the space allowed between the Left and Right Marigns or if a Perform Text/Decimal Tab function is included with the text to be centered, the result will be determined by the receiving system and is implementation-dependent.

5. This function shall follow a Hard New Line, except if the line being centered is the first line in the document.

DECIMAL TAB SET - DTS

1. DTS may occur anywhere in a document as one of the format commands that shall be followed by a Break function.

2. See Perform Decimal Tab (PDT).

3. The intention of a decimal tabulation is that the text is aligned vertically along the decimal point.

DOCUMENT START/END - DSE

1   Document Start - DS

    1. A document shall start with a DS.

0   Document End - DE

    1. A document shall end with a DE.

EMPHASIS

1. In DIF, either a bold function or a double strike function may be used for emphasis. Both will map into the Emphasis function in the exported DIF data stream. The mapping of the imported DIF data stream is implementation specific.

HARD HYPHEN - HH

1.   For example, HH is used to form compound words, such as X-ray.

2.   See SH.

HARD NEW LINE - HNL

1.   This function terminates any TLM function which might be in effect. The active cursor position goes to the position indicated by the parameter value of the most recent LMS function.

HARD SPACE - HSP

1.   In contrast, a soft space (2/0) is presented by the system as a non-printing character position. Soft space is subject to expansion or compression when JUSTIFY = ON is in effect, and it may be used as a line break.

JUSTIFY

1.   This function must be issued before the first Hard Page Start Function if the value is ON.

2.   The value of the Justify function cannot be changed within the document (i.e., after the first Hard Page Start function).

LEFT ALIGN

1.   In DIF, the Left Align function is always in effect (i.e., it is not affected by turning JUSTIFY OFF).

2.   Left Align is not in effect when the parameter value for Center is ON.

LINE HEIGHT (Select Vertical Spacing)

1.   In DIF, the height of each line must be 1/6 of an inch, making six lines per inch. However, a line is permitted to be blank.

2.   Also, see Line Spacing function.

LINE SPACING - LSP

1.   DIF allows three parameter values for this function: single, double, and triple. Single line spacing has no blank lines between lines of displayed text; double spacing has one blank line between two lines of displayed text; and triple spacing has two blank lines between two lines of displayed text.

2.   Any blank lines which occur as a result of this function must be included in the total number of lines allowed per page for the Page Length function.

3.   This fuction is not applied to Headers and/or Footers, only body text.

OVERSTRIKE

1.	This function is defined by start and stop delimiters.

2.	In DIF, overstrike is not intended to be used with diacritical marks to form composite characters. Instead it is intended to be used in the destructive manner (e.g., to cross out deleted text in a legal document).

PAGE LENGTH - PL

1.	The maximum setting for North American letter size paper is 66 lines at 6 lines per inch.

2.	The value for this control function shall be specified prior to the first Hard Page Start and may not occur subsequent to the first HPS.

3.	The combined total of TMS, lines of Header, HMS, lines of text body area, FMS, lines of Footer, and blank lines between the bottom-most line of Footer to the bottom-most edge of the paper shall not exceed the parameter value for PL.

PAGE NUMBERING INITIAL VALUE - PNIV

1.	PNIV is invoked by PNS.

2.	PNIV shall be issued prior to the first HPS. It need not be issued if page numbers will never be printed in the document

3.	The parameter value of PNIV is the number desired to be printed on the page on which page numbering is turned on. This value shall not begin incrementing until turned on by PNS.

PAGE NUMBERING SET - PNS

1.	The beginning page number is specified by PNIV.

2.	PNS shall be immediately followed by (i.e., no intervening text) the Break function Hard Page Start for the page on which page numbering is to begin.

3.	Incrementation of the PNIV parameter value does not occur until page numbering is turned on by PNS.

4.	See Header and Footer section for the positioning of page numbers.

PAGE ORIENTATION (Page Format Selection)

1.	In DIF, only vertical North American letter page format (portrait orientation) is supported.

2.	The page orientation in DIF may not change; therefore, it is not necessary to issue a Page Orientation function prior to each Hard Page Start function.

PAGE START/END - PSE

1   Hard Page Start - HPS

    1.   HPS shall be issued for the first page of a document following any functions pertaining to the entire document which cannot be changed within the document. Thereafter, HPS is typically introduced by the operator to force a new page for formatting purposes, such as the start of a new chapter, or the beginning of a block of text to appear together on a page, where upon it shall be preceded by a HARD PAGE END to end the previous page.

    2.   If the active position is 1) already at the beginning of a new page and 2) arrived there from a previous HARD PAGE START with no intervening text, advancing to a new page is required; otherwise advancing to a new page is not required.

0   Hard Page End - HPE

    1.   See HPS.

    2.   The amount of text between an HPS and HPE may require more than one physical sheet of paper.

    3.   Format effectors (LMS, RMS, HTSA, DTS, LSP and PNS) may occur after HPE, but must be followed immediately by an HPS.

    4.   HPE must be followed by either an HPS or a Document End.

    5.   Since the effect of HPE is similar to that of Form Feed (ASCII 0/12), a question has been raised about why we have chosen to specify a coding other than 0/12.

PAGE WIDTH - PW

1.   This control function is directly related to the pitch control function and Page Orientation function. In DIF, the maximum parameter value for PW is 85 for 10 pitch and 102 for 12 pitch when using North American Letter size paper in portrait orientation.

2.   The parameter value of this function shall be specified prior to the first HPS and shall not change within the body of the document.

3.   Any future extensions of DIF that support additional page orientations will require additional parameter values for PW (up to a maximum of 132).

PERFORM DECIMAL TAB - PDT

1.   The decimal tab stops are established by DTS.

PITCH (Select Horizontal Spacing)

1. DIF permits both 10 and 12 characters per inch. However, the value of the Pitch function may not be changed after the first Hard Page Start function (i.e., may not occur in the body of the document).

SOFT HYPHEN - SH

1. The X3L2 Ad Hoc Group has agreed that SH should be coded as CSI 2/4 6/1.

2. Breakable hyphen has been deleted from consideration since most word processors do not support it.

SUBSCRIPT (Partial Line Down)

1. Any interactions between SUBSCRIPT and vertical format effectors other than SUPERSCRIPT are not defined in this document.

SUPERSCRIPT (Partial Line UP)

1. Any interactions between SUPERSCRIPT and vertical format effectors other than SUBSCRIPT are not defined in this document.

TEXT TAB SET (HTSA)

1. A clear of previous text tabs settings is implied when this code is encountered. Therefore, all desired text tab settings must be explicitly stated with each occurrence of Text Tab Setting function when any changes are made to previous text tab settings.

2. A text tab will be performed by the ASCII character Horizontal Tab (HT) (0/9).

UNDERSCORE

1. This function is applied to everything between the start and stop delimiters, including SPACE, TAB (text), SUPERSCRIPT, and SUBSCRIPT.

HEADERS AND FOOTERS

In DIF, the following rules and restrictions shall apply:

1. Allowable characters between the header and footer delimiters are limited to ASCII 2/0 through 7/14 inclusive, and Hard New Line.

2. DIF has a maximum of 150 characters including both printable (graphic) and control characters for the combined total of header and footer (i.e., not 150 characters for each). There is no requirement if a document has both a header and a footer that the 150 characters must be divided equally between them.

3. Headers and footers, whether on the Right Facing Page or the Left Facing Page, are to be centered using the Center function. The proper sequence shall be

   Center (ON) Header (Start) HEADER TEXT Header (End) Center (OFF)

   or

   Center (ON) Footer (Start) FOOTER TEXT Footer (End) Center (OFF)

   This implies that, if the headers/footers are to be centered, they must be centered on the Left Facing Page and the Right Facing Page. The mid-point of the header is to be aligned on the mid-point between the Left or Right Facing Page Header Left Margin and the Left or Right Facing Page Header Right Margin. The analogous relationship is true for the footer and the Left or Right Facing Page Footer Margin functions.

4. In DIF, page numbers are treated as the (or part of the) header or footer. A page number location is transmitted as ### (ASCII 2/3) placed at the desired location between either a Header Start/End or Footer Start/End function. The three #s are replaced with the actual page number(s) at print out time, with leading zeros suppressed. Currently, the highest number permitted to be printed is 999 due to the 3-digit limitation.

FOOTER START/END - FSE

1   Footer Start - FTS

0   Footer End - FTE

      1.   The footer cannot be changed in DIF after the first HPS.

HEADER START/END - HSE

1   Header Start - HS

0   Header End - HE

      1.   The header cannot be changed in DIF after the first HPS.

LEFT FACING PAGE FOOTER LEFT MARGIN - LPFL

1.   See LMS function for restrictions and valid settings.

LEFT FACING PAGE FOOTER RIGHT MARGIN - LPFR

1.   See RMS function for restrictions and valid settings.

LEFT FACING PAGE HEADER LEFT MARGIN - LPHL

1.   See LMS function for the valid settings.

LEFT FACING PAGE HEADER RIGHT MARGIN - LPHR

1.   Ree RMS Function for restrictions and valid settings.

RIGHT FACING PAGE FOOTER LEFT MARGIN - RPFL

1.   See LMS function for the valid settings.

RIGHT FACING PAGE FOOTER RIGHT MARGIN - RPFR

1.   See RMS function for restrictions and valid settings.

RIGHT FACING PAGE HEADER LEFT MARGIN - RPHL

1.   See LMS function for the valid settings.

RIGHT FACING PAGE HEADER RIGHT MARGIN - RPHR

1.   See RMS function for restrictions and valid settings.

MARGINS

All parameter values for margin functions must be specified (at least initial values) prior to the first Hard Page Start. (LMS and RMS may change within the document.)

All horizontal margins are relative to the left edge of the paper, with the left edge being at position 1. Top Margin Set and Bottom Margin Set are relative to the top edge of the paper, with the top edge being line 1.

The margin is the unprinted area that frames the text body area on all four sides: top, bottom, left, and right. It is the space between the text and the edge of the paper.

NOTE: DIF has two additional margins which are not relative to the top or left edges of the paper: header margin and footer margin. Instead, they are relative to the header text/text body area and the footer text/text body area, respectively. Also see Left and Right Facing Page Margin Set functions.

BOTTOM MARGIN SET - BMS

1.   There is no default value.

2.   The BMS shall be greater than or equal to the TMS.

FOOTER MARGIN SET - FMS

1.   For example, if BMS equals 56, plus two lines of Footer Text and FMS equals 2, the the first line of the footer will be on line 59.

2.   FMS shall occur only before the first Hard Page Start.

3.   If the document does not contain a footer, the FMS need not be transmitted.

4.   The default value is 0.

5.   FMS restrictions are less than or equal to 65 for 6 lines per inch when using portrait oriented North American letter size paper.

HEADER MARGIN SET - HMS

1.   For example, if TMS equals 7, plus two lines of Header Text, and HMS equals 2, then the first line of text body area is on line 11.

2.   HMS shall occur only before the first Hard Page Start.

3.   If the document does not contain a header, the HMS need not be transmitted.

4.   The default value is 0.

5.   HMS restrictions are less than or equal to 65 for 6 lines per inch when using portrait oriented North American letter size paper.

LEFT MARGIN SET - LMS

1.   Valid left margin settings are 1 to 85 for 10 pitch and 1 to 102 for 12 pitch when using portrait oriented North American letter size paper.

2.   An LMS shall occur before the first HARD PAGE START (HPS) of a document. Thereafter LMS may occur anywhere in a document as one of the format commands that shall precede a break command (HARD NEW LINE, HARD PAGE START, HARD PAGE END).

3.   Future revisions of this standard may include Landscape orientation, in which case the maximum setting will be 132.

RIGHT MARGIN SET - RMS

1.   Valid right margin settings are 1 to 85 for 10 pitch and 1 to 102 for 12 pitch when using portrait oriented North American Letter size paper and shall be greater than or equal to the left margin setting.

2.   Future revisions of this standard may include landscape orientation; in which case the maximum setting will be 132.

TEMPORARY LEFT MARGIN SET - TLM

1.   TLM takes effect immediately and stays in effect until terminated. TLM functions may be nested (i.e., no terminator function between them); in which case, the value of each subsequent TLM overrides the previous value. Upon termination of TLM, the left margin position is re-established as the parameter value of the most recently issued LMS.

2.   Valid TLM parameter values are 1 to 102 for 12 pitch and 1 to 85 for 10 pitch when using portrait oriented North American letter size paper.

3.   The parameter value of TLM may be greater than, less than, or equal to the most recent parameter value of LMS; however, the result of having a value less than LMS is not defined by this standard.

4.   TLM may occur anywhere in the document, including in the middle of text.

TOP MARGIN SET - TMS

1.   The default value is 1.

2.   TMS shall occur only before the first Hard Page Start.

APPENDIX C

Classes of Functions

Break functions

    Hard New Line
    Hard Page End
    Hard Page Start

Line Format functions

    Left Margin Set
    Right Margin Set
    Text Tab Set (Horizontal Tabulation Set Absolute)
    Decimal Tab Set
    Center
    Temporary Left Margin Set
    Line Spacing

Document Format Functions

    Justify
    Left Align
    Pitch
    Line Height

Page Format functions

    Top Margin Set
    Bottom Margin Set
    Header Margin Set
    Footer Margin Set
    Page Length
    Page Width
    Page Numbering Set
    Page Orientation
    Left Facing Page Header Left Margin
    Left Facing Page Header Right Margin
                    Footer Left Margin
                     Footer Right Margin
    Right Facing Page Header Left Margin
                    Header Right Margin
                    Footer Left Margin
                    Footer Right Margin

Rendition functions:

    Emphasis
    Overstrike
    Underscore
    Superscript
    Subscript

Miscellaneous functions:

    Page Numbering Initial Value
    Header Start/End
    Footer Start/End
    Hard Space
    Hard Hyphen
    Soft Hyphen
    Perform Decimal Tab
    Perform Text Tab (Horizontal Tabulation)
    Document End
    Document Start

APPENDIX D

Rules Regarding Issuance of Control Functions

Control functions which may be issued once only; issuance must follow the Document Start function and precede the first Hard Page Start function:

Justify
Page Orientation
Line Height
Pitch
Top Margin Set
Bottom Margin Set
*Header Margin Set
*Footer Margin Set
Page Length
Page Width
Page Numbering Initial Value (if page numbering is desired within the document)
*Left Facing Page Header Left Margin
Header Right Margin
Footer Left margin
Footer Right Margin

*Right Facing Page Header Left Margin
Header Right Margin
Footer Left Margin
Footer Right Margin
Left Align

*Header Start/End
*Footer Start/End

Control functions for which parameter values may change throughout the document; the initial value must be specified prior to the first Hard Page Start:

Left Margin Set
Right Margin Set
Line Spacing
Text Tab Set (Horizontal Tabulation Set Absolute)
Decimal Tab Set
Page Numbering Set (may change once only)

Control functions which may occur throughout the document, following the first Hard Page Start:

Temporary Left Margin Set
Hard Space
Hard Hyphen
Soft Hyphen
Hard New Line
Perform Decimal Tab
Horizontal Tab (Perform Text Tab)
Hard Page Start
Hard Page End
Center
Underscore
Emphasis
Overstrike
Subscript
Superscript
Perform Text Tab

Control functions which delimit the document:

Document Start
Document End

The following functions must be followed immediately by a function from the Break function class; i.e, no intervening text is permitted. However, one or more of these functions may be specified prior to the Break function. The specified functions are then invoked by the Break function.

Left Margin Set
Right Margin Set
Line Spacing
Text Tab Set
Decimal Tab Set
Page Numbering Set (Must precede a Hard Page Start function ONLY)

*These functions may be omitted if there are no headers and/or footers in the documents.

APPENDIX E

Permissible C0 Controls

Horizontal Tab (HT)   0/9
(Perform Text Tab function
in DIF)

Escape   1/11
(Used in combination with
5/11 to achieve CSI)

Encoding for the 13 Previously Standardized Functions

DIF Function Name	Standardized Function Name	Standard Defined In	Encoding
Center (on)	JFY-JUSTIFY	X3.64	**CSI 3/6 2/0 4/6
Center (off)	JFY-JUSTIFY	X3.64	CSI 3/0 2/0 4/6
Text Tab Set	HTSA-HORIZONTAL TABULATION ABSOLUTE	ISO 6429	***CSI Pn;Pn;...;Pn 2/0 4/14
Justify (on)	JFY-JUSTIFY	X3.64	CSI 3/2 2/0 4/6
Justify (off)	JFY-JUSTIFY	X3.64	CSI 3/0 2/0 4/6
Underscore (on)	SGR-SELECT	X3.64	CSI 3/4 6/13
Underscore (off)	GRAPHIC RENDITION	X3.64	CSI 3/2 3/4 6/13
Emphasis (on)	SGR-SELECT	X3.64	CSI 3/1 6/13
Emphasis (off)	GRAPHIC RENDITION	X3.64	CSI 3/2 3/2 6/13
Overstrike (on)	SGR-SELECT	ISO 6429	CSI 3/9 6/13
Overstrike (off)	GRAPHIC RENDITION	ISO 6429	CSI 3/2 3/9 6/13
Left Align (on)	JFY-JUSTIFY	X3.64	CSI 3/5 2/0 4/6
Subscript (on)	PLD-PARTIAL LINE DOWN	X3.64	1/11 4/11
Subscript (off)	PLU-PARTIAL LINE UP	X3.64	1/11 4/12
Superscript (on)	PLU-PARTIAL LINE UP	X3.64	1/11 4/12
Superscript (off)	PLD-PARTIAL LINE DOWN	X3.64	1/11 4/11
Page Orientation (Portrait)	PFS-PAGE FORMAT SELECTION	X3.98	CSI 3/4 2/0 4/10
Line Height	SVS-SELECT VERTICAL SPACING	X3.98	CSI 3/0 2/0 4/12
Pitch (10 chars/in)	SHS-SELECT HORIZONTAL SPACING	X3.98	CSI 3/0 2/0 4/11
Pitch (12 chars/in)	SHS-SELECT HORIZONTAL SPACING	X3.98	CSI 3/1 2/0 4/11
Hard New Line	NEL-NEXT LINE	X3.64	1/11 4/5

**CSI is 1/11 5/11
***P is a parameter character in a control sequence. P is from 3/0 to 3/15, inclusive.
Pn is a numeric parameter in a control sequence. Ps is a selective parameter
in a control sequence. Within a control sequence containing a variable number
of selective (or numeric) parameters, they may be separated from each other
by 3/11. Pn and Ps are strings of zero or more characters from 3/0 through 3/9,
inclusive. See ANSI X3.64 for more details.

APPENDIX G

29 New Control Functions

Description	CSI	Parameters	I	F
Left Margin Set	1/11 5/11	Pn=1 to 132	2/4	4/0
Right Margin Set	1/11 5/11	Pn=1 to 132	2/4	4/1
Top Margin Set	1/11 5/11	Pn=0 to 66	2/4	4/2
Bottom Margin Set	1/11 5/11	Pn=0 to 66	2/4	4/3
Header Margin Set	1/11 5/11	Pn=0 to 66	2/4	4/4
Footer Margin Set	1/11 5/11	Pn=0 to 66	2/4	4/5
Temporary Left Margin Set	1/11 5/11	Pn=1 to 132	2/4	4/6
Decimal Tab Set	1/11 5/11	Pn=1 to 132;...	2/4	4/7
Hard Space	1/11 5/11		2/4	4/8
Hard Hyphen	1/11 5/11		2/4	4/9
Hard Page Start/End	1/11 5/11	Ps=1(start)=0 end	2/4	4/10
Document Start/End	1/11 5/11	Ps=1(start=0 end	2/4	4/11
Page Length	1/11 5/11	Pn=1 to 66 2/4 4/12		
Page width	1/11 5/11	Pn=1 to 132	2/4	4/13
Page Numbering Initial Value	1/11 5/11	Pn=0 to 999	2/4	4/14
Page Numbering Set	1/11 5/11	Ps=1(on)=0(off)	2/4	4/15
Left Facing Page				
Header Left Margin	1/11 5/11	Pn=1 to 132	2/4	5/0
Header Right Margin	1/11 5/11	Pn=1 to 132	2/4	5/1
Footer Left Margin	1/11 5/11	Pn=1 to 132	2/4	5/3
Footer Right Margin	1/11 5/11	Pn=1 to 132	2/4	5/4
Right Facing Page				
Header Left Margin	1/11 5/11	Pn=1 to 132	2/4	5/6
Header Right Margin	1/11 5/11	Pn=1 to 132	2/4	5/7
Footer Left Margin	1/11 5/11	Pn=1 to 132	2/4	5/9
Footer Right Margin	1/11 5/11	Pn=1 to 132	2/4	5/10
Line Spacing	1/11 5/11	Ps=0(single)=1(double) =2(triple)	2/4	5/12
Header Start/End	1/11 5/11	Ps=0(end)=1(start)	2/4	5/13
Footer Start/End	1/11 5/11	Ps=0(end)=1(start)	2/4	5/14
Perform Decimal Tab	1/11 5/11		2/4	6/0
Soft Hyphen	1/11 5/11		2/4	6/1

Bibliography

1. American National Standards Institute, <u>Additional Controls for Use with American National Standard Code for Information Interchange</u>, ANSI X3.64, 1979 July 18.

2. American National Standards Institute, <u>Code Extension Techniques for Use with the 7-Bit Coded Character Set of American National Standard Code for Information Interchange</u>, ANSI X3.41, 1974 May 14.

3. American National Standards Institute, <u>Code for Information Interchange</u>. ANSI X3.4, 1977 June 9.

4. American National Standards Institute, <u>Text Information Interchange in Page Image Format (PIF)</u>, ANSI X3.98, 1983.

5. Knoerdel, Joan E., <u>A Survey of Standardization Efforts of Coded Character Sets for Text Processing</u>, National Bureau of Standards' Special Publication 500-81, September 1981.

6. Knoerdel, Joan E., and Roy Pierce, "Description of Text Structures Defined for Office Document Interchange", to be published in <u>Journal of Telecommunication Networks</u>, Winter Issue, 1983.

7. International Organization for Standardization, <u>Additional Control Functions for Character-Imaging Devices</u>, ISO IS 6429, May 1983.

8. International Organization for Standardization (ISO), <u>Open Systems Interconnection Basic Reference Model</u>, IS 7498, October 1983.

9. National Bureau of Standards, <u>Additional Controls for Use with American National Standard Code for Information Interchange</u>, Federal Information Processing Standards Publication 86, January 29, 1981.

NBS-114A (REV. 2-80)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBSIR 84-2836	2. Performing Organ. Report No.	3. Publication Date February 1984

4. TITLE AND SUBTITLE

Document Interchange Format

5. AUTHOR(S)

Joan E. Knoerdel and Shirley Ward Watkins

6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)	7. Contract/Grant No.
NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234	8. Type of Report & Period Covered

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP)

National Bureau of Standards
Institute for Computer Sciences and Technology
Systems and Network Architecture Division

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)

In the absence of standards, both the private and public sectors have addressed document interchange among different vendors' text processing systems in a number of ways. In an attempt to solve this interchange problem with respect to encoding of control functions for the Department of the Navy, a project was originated by the Office of the Under Secretary of the Navy for Financial Management to determine the formatting requirements of the Department of the Navy and to translate those requirements to a representation that would be supported by text processing system providers. The encoded representation of the formatting control functions has become known as the Document Interchange Format (DIF).

This paper describes the overall approach taken by the DIF and then provides definitions for and implementation details for DIF. The body of the paper is intended for overall understanding and as such is intended for managers and technical staff. Then, there are a number of appendices provided which are specifically written for those implementing DIF.

12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)

ANSI X3.64; code standards; control functions; control sequence introducer; DIF; document interchange format; document interchange standards; text processing systems; word processors.

13. AVAILABILITY	14. NO. OF PRINTED PAGES
☒ Unlimited	40
☐ For Official Distribution. Do Not Release to NTIS	
☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.	15. Price
☐ Order From National Technical Information Service (NTIS), Springfield, VA. 22161	

USCOMM-DC 6043-P80