STANDARD ST.8

Version 1.0

STANDARD RECORDING OF INTERNATIONAL PATENT CLASSIFICATION (IPC) SYMBOLS ON MACHINE-READABLE RECORDS

Revision adopted by the SCIT Standards and Documentation Working Group at its fifth session on November 11, 2004

Editorial Note

by the International Bureau (December 2004)

The Standards and Documentation Working Group (SDWG) of the Standing Committee on Information Technologies (SCIT) adopted this revision of Standard ST.8 at its fifth session on November 11, 2004. This revision of Standard ST.8 incorporates changes made necessary by the IPC reform initiative.

Industrial property offices are asked to implement this new version of Standard ST.8 for all patent documents with a publication date from January 1, 2006, onwards. For patent documents published prior to that date, the previous version of the Standard should continue to be used.

The previous version of Standard ST.8, valid until December 31, 2005, is reproduced as an Annex to the new Standard ST.8.

TABLE OF CONTENTS

Introduction

1. This recording convention provides that symbols of the International Patent Classification (IPC) should be presented on machine-readable records for the exchange of information in machine-readable form in a fixed-length field in 50 positions, each part of the Int. Cl. symbol being recorded in specific positions and in the manner prescribed.

2. The examples given are intended to clarify the text and should not be considered as comprehensive.

Recording

3. For the recording of IPC symbols on machine-readable records a field of 50 positions should be assigned for each symbol, the 50 positions of the field to be used as follows:

POSITION(S)	CONTENT	VALUES
1	Section	A,…,H
2,3	Class	01,…,99
4	Subclass	A,…,Z
5 to 8	Main Group (right aligned)	1,…,9999, blank
9	Separating character	/ (“Slash”)
10 to 15	Subgroup (left aligned)	00,…,999999, blank
16 to 19	For future use	4 blanks
20 to 27	Version indicator	YYYYMMDD date format
28	Classification level	C,A,S
29	First or later position of symbol	F,L
30	Classification value (inventive or non-inventive)	I,N

31 to 38	Action date	YYYYMMDD date format
39	Original or reclassified data	B,R,V,D
40	Source of classification data	H,M,G
41-42	Generating office	AA,…,ZZ (ST.3)
43-50	For future use	8 blanks

4. Unused positions in the IPC classification fields Group (positions 5-8) and Subgroup (positions 10-15) should be left blank. The only other positions that may be left "blank" are the ones reserved for "future use". All other positions must be assigned one of the acceptable "values" listed in the table of paragraph 3. Any zero appearing in the symbols should be recorded.

5. Considering the numerals appearing after the separating character, the most significant digit (including the case where it is zero, e.g., subgroup 02) should be in position 10. Any unused positions should be left blank.

6. Representation of the indicators

Positions 1 to 19: Recording of the parts of the IPC symbols

IPC symbols are defined in Part 5 of the WIPO *Handbook on Industrial Property Information and Documentation* and in the latest version of the Guide to the IPC.

Positions 20 to 27: Version indicator

Although in the paper publications a version indicator may contain four or six digits, the version indicator in machine-readable records contains eight digits, namely YYYYMMDD with Y for year, M for month and D for day.

Position 28: Classification level

Offices are expected to classify each subject matter only in one level (core or advanced). However, both levels need to be completely represented in the master classification database and thus a level indicator is needed. The level indicator is also useful for indicating situations where an office does not classify in either the core or the advanced level classification, i.e., when an office only assigns classifications to the subclass level. The level indicator enables to make the difference between core, advanced and subclass levels. The letters C (Core), A (Advanced) and S (Subclass) are used for this one-digit field.

Position 29: First or later position of symbols

The position of the first invention information classification can be recognized by this field. The letters F and L are used for first and later position, respectively.

Position 30: Classification value (inventive or non-inventive)

The difference between invention information and other information is important for the retrieval of the information. The letters I and N are used for the invention and non-invention information, respectively.

Positions 31–38: Action date

The date of assigning the classification symbol (action date) is represented by eight digits, namely YYYYMMDD. This date can be used to check if a classification needs to be reviewed after revision of the scheme, e.g., in case of creating new subdivisions.

Position 39: Original and reclassified data

Original data is the first data assigned to the document. In case of a publishing office assigning classification symbols at the core level, another office may also assign symbols at the advanced level as original data.

Reclassified data is data changed due to a change in the classification schemes.

Various data is data changed due to an incidental reclassification of an individual document, such as the correction of a mistake.

Deleted data is data which has to be deleted from the Master Classification Database, due to a change in assigning of classification symbols to a document.

The indication of the different types of data is marked by the letters B for the basic or original data, R for reclassified data, V for various incidental changes, and D for data to be deleted.

Position 40: Source of classification data

The following sources of classification data are foreseen:

- Intellectual classification by persons, value H for human generated data.

- Machine classification by the propagation of earlier intellectual classification through the use of common priorities in the patent application. The value M is used in this case and will facilitate later corrections.
- Classification symbols generated by software using automatic analysis of the content of the patent document. The letter G is used to indicate this source of generated data.

Positions 41–42: Generating Office

Since part of the original data in the advanced level and the reclassified data can be delivered by offices other than the publishing office, the information source of such data is recorded by a field of two characters. The country or office code CC, as defined by WIPO Standard ST.3, must be used.

7. Recording of complete IPC symbols

The full classification symbol must always be used when recording it on machine-readable records. The IPC section, class and subclass should be provided for each group or subgroup classification, even if previously provided with another group or subgroup classification in the same document.

See paragraph 2 of WIPO Standard ST.10/C for the recommended presentation of IPC classifications on machine displays or in printed documents.

8. A schematic representation of the contents of the 50 positions is as follows:

Section	Class		Subclass	Main Group				Separating character	Subgroup						Blanks			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

Version indicator								Classification level	First or later position of a symbol	Classification value	Action date							
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38

Original or reclassified data	Source of classification data	Generating office		Blanks							
39	40	41	42	43	44	45	46	47	48	49	50

Example

One sample representation of IPC classification symbols assigned on June 1st 2007 and their indicators is:
Int. Cl. (2006)

B28B 5/00 (2006.01)	classification in advanced level	invention information
B28B 1/29 (2007.04)	classification in advanced level	invention information
H05B 3/10	classification in core level	non-invention information

According to this Standard, this example would be recorded on machine-readable records as follows:

Record 1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
B	2	8	B				5	/	0	0									2	0	0	6	0	1	0	1

28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
A	F	I	2	0	0	7	0	6	0	1	B	H	E	P								

Record 2

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
B	2	8	B				1	/	2	9									2	0	0	7	0	4	0	1

28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
A	L	I	2	0	0	7	0	6	0	1	B	H	E	P								

Record 3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
H	0	5	B				3	/	1	0									2	0	0	6	0	1	0	1

28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
C	L	N	2	0	0	7	0	6	0	1	B	H	E	P								

ANNEX: Previous version of Standard ST.8

STANDARD RECORDING OF INTERNATIONAL PATENT CLASSIFICATION (IPC) SYMBOLS ON MACHINE-READABLE RECORDS

Revision adopted by the PCIPI Executive Coordination Committee at its fourteenth session on May 20, 1994

Introduction

1. This recording convention provides that symbols of the International Patent Classification (IPC) should be presented on machine-readable records for the exchange of information in machine-readable form in a fixed-length field in 18 positions, each part of the Int. Cl. symbol being recorded in specific positions and in the manner prescribed.

2. The examples given are intended to clarify the text and should not be considered as comprehensive.

Recording

3. For the recording of IPC symbols on machine-readable records a field of 18 positions should be allotted for each symbol, the 18 positions of the field to be used as follows:

POSITION(S)	CONTENT
1	blank (for future use)
2	number of the IPC edition used
3	Section
4	blank (for future use)
5, 6	Class
7	Subclass
8	blank (for future use)
9 to 11	group, right adjusted, leading spaces
12	separating character (as defined in paragraph 4)
13 to 17	subgroup, left adjusted
18	qualifying character (as defined in paragraph 7)

4. The separating character, in position 12, should be an oblique stroke for all symbols except indexing codes where a colon is to be used.

5. Unused positions should be left blank. Any zero appearing in the symbols should be recorded.

6. Considering the numerals appearing after the separating character, the most significant digit (including the case where it is zero, e.g., subgroup 02) should be in position 13. Any unused positions should be left blank.

7. The qualifying character in position 18 is as defined below:

(a) the capital letter “A” defines the first classification symbol representing invention information;

(b) the capital letter “B” defines other classification symbols representing invention information;

(c) the character “-” defines classification symbols representing additional information;

(d) the twenty-three capital letters “C” to “Y” (including I and O) and the eight numerals “2” to “9” identify linked indexing codes as well as the classification symbols to which the codes are linked, the letter “C” identifying symbols and/or codes linked to form a first set of information, the letter “D” identifying symbols and/or codes linked to form the second set, and so on. Following the letter “Y”, the numeral “2” identifies symbols and/or codes linked to form a set in the same way as the letters;

(e) if more than 31 sets of linked IPC classification symbols and indexing codes have to be assigned to a document the thirty second set or more should be recorded by use of the small letter “z”;

(f) the capital letter “Z” defines unlinked indexing codes.

8. *Recording of incomplete symbols*. When symbols are present on patent documents in a truncated form, e.g., as C 23 C 1/00, 7/00, the section, class and subclass must be added when recording these symbols, i.e., as C 23 C 1/00, C 23 C 7/00.

9. *X-notations*. The X should be treated like any other character which is part of the IPC symbol. Thus, if it occurs immediately after the subclass letter it shall be right-adjusted in the positions 9 to 11; if it occurs after the oblique stroke, it shall be left-adjusted in the positions 13 to 17.

10. A schematic representation of the contents of the 18 positions is as follows:

Blank	Number of IPC Edition	Section	Blank	Class		Subclass	Blank	Group (right adjusted)			Separating character	Subgroup (left adjusted)					Qualifying character
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Examples

Three sample representations of IPC classification symbols and indexing codes which could be given on a patent document are:

(a) C 08 F 210/16, 255/04
//A 61 K 47/00, C 09 J 151/06 (C 08 F 210/16, 214:06) (C 08 F 255/04, 214:06)

(b) B 29 C 65/08
//B 29 K 83:00, B 29 L 23:18

(c) C 07 D 401/06, 213/60
// A 01 N 43/40, 43/90 (C 07 D 401/06, 233:32, 213:60)

According to this Standard, these examples would be recorded on machine-readable records as follows:

(a)

	6	C		0	8	F		2	1	0	/	1	6			A
	6	C		0	8	F		2	5	5	/	0	4			B
	6	A		6	1	K			4	7	/	0	0			–
	6	C		0	9	J		1	5	1	/	0	6			–
	6	C		0	8	F		2	1	0	/	1	6			C
	6	C		0	8	F		2	1	4	:	0	6			C
	6	C		0	8	F		2	5	5	/	0	4			D
	6	C		0	8	F		2	1	4	:	0	6			D

(b)

	6	B		2	9	C			6	5	/	0	8			A
	6	B		2	9	K			8	3	:	0	0			Z
	6	B		2	9	L			2	3	:	1	8			Z

(c)

	6	C		0	7	D		4	0	1	/	0	6			A
	6	C		0	7	D		2	1	3	/	6	0			B
	6	A		0	1	N			4	3	/	4	0			–
	6	A		0	1	N			4	3	/	9	0			–
	6	C		0	7	D		4	0	1	/	0	6			C
	6	C		0	7	D		2	3	3	:	3	2			C
	6	C		0	7	D		2	1	3	:	6	0			C

[End of Standard]