

## 5.5. PATENT BIBLIOGRAPHIC DATA SCHEMA

URI: <http://www.epo.org/patent-bibliographic-data/1.1/>

Name Space: [patent](#)

Properties corresponding to bibliographic data available in XML ST36			
name	Type	XML element <sup>3</sup>	property (patent:*)
Publication data	DocId		<b>Publication</b> <sup>4</sup>
country	Text	B190	<b>CountryCode</b>
number	Text	B110	<b>Number</b>
kind code	Text	B130	<b>KindCode</b>
corr. code	Text	B151+B132EP	CorrectionCode
date <sup>5</sup>	date	B140/date ("- " added")	<b>Date</b>
Int. Pub. data	DocId		InternationalPublication
country	Text	B871/ctry	CountryCode
number	Text	B871/dnum/pnum	Number
date	date	B871/date ("- " added")	Date
Application data	DocId		Application
number	Text	B210	Number
date	date	B220/date ("- " added")	Date
Int. App. data	DocId		InternationalApplication
country	Text	B861/ctry	CountryCode
number	Text	B861/dnum/anum	Number
date	date	B861/date ("- " added")	Date
Priority data	bag DocId		Priority
country	Text	B330/ctry	CountryCode
number	Text	B310	Number
date	date	B320/date ("- " added")	Date
IPC <sup>6</sup>	bag Text	classification-ipcr/text or B511, B512	IPC
CPC	bag Text	classification-cpc/text	CPC
Applicant names	bag Text	B711/ttl+fnm+snm+sfx+onm+odv	Applicant
Inventors names	bag Text	B721/ttl+fnm+snm+sfx+onm+odv	Inventor
Proprietors names	bag Text	B731/ttl+fnm+snm+sfx+onm+odv	Proprietor
Represent. names	bag Text	B741/ttl+fnm+snm+sfx+onm+odv	Representative
Title	alt Text	B541, B542	Title (default value depends on @lang in ep-patent-document)
Abstract <sup>7</sup>	seq Text	abstract/@lang, abstract/p	Abstract (first abstract in the ordered list depends on @lang in ep-patent-document)

<sup>3</sup> definition available at <http://www.wipo.int/scit/en/standards/st36-xml-dtd.htm>, EPO's production DTD.

<sup>4</sup> with the exception of CorrectionCode, all the properties contained in **Publication** should be mandatory. All other properties are considered as optional.

<sup>5</sup> all dates must be formatted as follows: yyyy-mm-dd

<sup>6</sup> the content of patent:IPC and patent:CPC is IPC/CPC XML raw data where trailing and duplicated empty spaces are removed. The IPC data (not CPC) contained in dc:subject and pdf:Keywords is not IPC XML raw data, i.e. it is a subset of this XML raw data.

<sup>7</sup> all elements or data that are not compatible with the XMP specification should be removed from the abstract content, e.g. HTML elements such as i, br, etc., or replaced by an empty space.