

Chap 1 : Database Design and Schema (Tables)

NUS Patent Database uses PostgreSQL version 7.1.3. PostgreSQL is an Open Source database system.

WARNING: Before upgrading postgres, please read this:
<http://www.postgresql.org/docs/7.3/static/install-upgrading.html>

The various fields of the patent data can be organized into a set of tables as described below:

<u>TABLE NAME</u>	<u>DETAILS</u>
Tables describing Basic Patent Details	
1. TBL_PATN fields created by NUS	: Basic Patent Information, plus custom
2. TBL_INVNT	: Inventor Details
3. TBL_ASSG	: Assignee (Company) Details
Tables describing Patent Classes (These are used to indicate the area of technology being patented)	
4. TBL_CLAS Class	: Classification Details – Main US Patent
5. TBL_CLAS_XCL Patent Classes	: Classification Details-XCL – Other US
6. TBL_CLAS_UCL References	: Classification Details-UCL – Unofficial
7. TBL_CLAS_DCL References	: Classification Details-DCL – Digest
8. TBL_CLAS_ICL Patent Classes	: Classification Details-ICL – International
9. TBL_CLAS_FS Search/Subclasses	: Classification Details-Field of
Tables describing Patent References	
10. TBL_UREF (cited to) by each patent in this database	: U.S. Reference – US Patents referenced
11. TBL_FREF by patentseach patent in this database	: Foreign Reference – Foreign Patents cited
Tables describing Other References	
12. TBL_OTHERS and other	: Other Info (References to Scientific publications sources, Government Interest, Parent Case, Drawing Description, Detail Description)
Tables describing Claims	
13. TBL_CLMS	: Claims Description
14. TBL_DCLM	: Design Claims

Miscellaneous Tables

15. TBL_COUNTRY_CODE USPTO site	: Country Codes/Names : Obtained from
16. TBL_STATE_CODE from USPTO site	: State Codes (of USA)/Names : Obtained
17. TBL_PATENTTYPE_CODE Obtained from USPTO site	: Types of Patent (utility/plant/etc) :
18. TBL_ASSGTYPE_CODE Obtained from USPTO site	: Types of Assignee (US/Foreign/etc) :
19. TBL_PRIR	: Foreign Priority Details
20. TBL_REIS	: Reissue Details
21. TBL_RLAP	: Related U.S. Application Data
22. TBL_PCTA	: PCT Information
23. TBL_LREP	: Legal Information

Note: The tables have 2 kinds of fields :

- **Original Fields** : These are the primary and basic attributes of a patent, as retrieved from the actual Raw Patent Data (text files). These fields are populated by running the Parsers.
- **Custom Fields** : These are additional, custom fields that are Not present in the Raw Patent Data (text files) and are generated by running certain Custom scripts.

DETAILED TABLE DESCRIPTION

1. Table Name : TBL PATN

Logical Group : Patent Bibliographic Information

ID Code : PATN

There is a single unique key, **PATN_WKU**, which is the PATENT NUMBER which uniquely identifies all the data relevant to a particular patent in every table. This unique key is found in the **TBL_PATN** table.

S/N	Field name	Required/ Option	Field Type/Size	Field Desc
ORIGINAL FIELDS (from Raw Patent Data)				
1.	PATN_WKU (PK)	Not Null	CHAR(9)	Patent Number (This nine-character string contains embedded zeroes between the prefix-part and the numerical-part of the patent number) For eg: PATN_WKU= PP004546
2.	PATN_SRC	Not Null	INTEGER	Series Code
3.	PATN_APN	Not Null	CHAR(6)	Application Number
4.	PATN_APT	--	INTEGER	Application Type
5.	PATN_ART	--	CHAR(3)	Art Unit
6.	PATN_APD	--	DATE	Application Filing Date
7.	PATN_TTL	--	VARCHAR(255)	Title of Invention
8.	PATN_NCL	--	INTEGER	Number of Claims
9.	PATN_ECL	--	VARCHAR(14)	Exemplary Claim Number(s)
10.	PATN_EXP	--	VARCHAR(500)	Primary Examiner
11.	PATN_PBL	--	CHAR(1)	Publication Level
12.	PATN_ISD	--	DATE	Issue Date
13.	PATN_EXA	--	VARCHAR(500)	Assistant Examiner
14.	PATN_NDR	--	INTEGER	Number of Drawing Sheets
15.	PATN_NFG	--	INTEGER	Number of Figures
16.	PATN_DCD	--	DATE	Disclaimer Date

17.	PATN_NPS	--	INTEGER	Number of Pages of Specifications
18.	PATN_TRM	--	TEXT	Term of Patent
19.	PATN_ABST	--	TEXT	Abstract - This is taken from the ABST(Abstract) section
20.	PATN_BSUM	--	TEXT	Brief Summary - This is taken from the BSUM(Brief Summary) section
21.	PATN_CLMS_STM	--	TEXT	Claims Statement
CUSTOM FIELDS (Fields that were generated by NUS and Not present in the raw patent data)				
22.	PATN_ISSUE_YEAR	--	INTEGER	Year of Issue (The year field is retrieved from PATN_ISD original field.)
23.	ACTUAL_PATN_WKU	--	CHAR(9)	Patent Number WithOUT embedded zeroes (This field is obtained by removing the embedded zeroes in the PATN_WKU field) For eg: PATN_WKU= PP004546 ; ACTUAL_PATN_WKU= PP4546
24.	NUM_INVTS		INTEGER	Total No. of inventors of this patent
25.	NUM_ASSGS		INTEGER	Total No. of assignees of this patent
26.	NUM_CITETO	--	INTEGER	No. of US Patents cited TO by this patent
27.	NUM_CITEBY_19762000	--	INTEGER	No. of other US patents in 1976-2000 that refer to this patent
28.	NUM_OTHERREF	--	INTEGER	No. of citations to scientific and other publications
29.	NUM_FWD1	--	INTEGER	Forward-1: No. of other US patents issued within 1 year that make citation to this patent
30.	NUM_FWD3	--	INTEGER	Forward-3: No. of other US patents issued within 3 years that make citation to this patent
31.	NUM_FWD5	--	INTEGER	Forward-5: No. of other US patents issued within 5 years that make citation to this patent
32.	NUM_BACK1	--	INTEGER	Backward-1 : No of US patents issued within the past 1 year that this patent cites to
33.	NUM_BACK3	--	INTEGER	Backward-3 : No of US patents issued within the past 3 years that this patent cites to
34.	NUM_BACK5	--	INTEGER	Backward-5 : No of US patents issued within the past 5 years that this patent cites to
35.	FORMAT_OCL	--	VARCHAR(255)	Formatted OCL Class – US Classification
36.	FORMAT_XCL	--	VARCHAR(255)	Formatted XCL Class – Cross Reference
37.	FORMAT_ICL	--	VARCHAR(255)	Formatted ICL Class – International Classification

Notes :

- PATN_WKU : Unique Key Constraint

- PATN_ISSUE_YEAR : This attribute is obtained from 'PATN_ISD' - 'Issue Date' attribute of each patent, as seen in the raw patent text data, and not based on the patent number classification. Hence, there will be a slight discrepancy in the numbers, when compared to those listed by other standard patent websites.
- The NUM_FWD and NUM_BACK fields are based on PATN_ISSUE_YEAR. As such, they do not measure time exactly in calendar years. For that, the user should compute her own citation counts using PATN_ISD

2. Table Name : TBL_INVT

Logical Group : Inventor

(Multiple Inventors can be present for a single patent number)

ID Code : INVT

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	INVT_NAM	Not Null	TEXT	Inventor Name
3.	INVT_CTY	--	TEXT	City
4.	INVT_STR	--	TEXT	Street
5.	INVT_STA	--	CHAR(2)	State
6.	INVT_CNT	--	CHAR(3)	Country
7.	INVT_ZIP	--	CHAR(5)	Zip Code
8.	INVT_R47	--	CHAR(2)	Rule 47 Indicator
9.	INVT_ITX	--	TEXT	Inventor Descriptive Text
CUSTOM FIELDS (Fields that were generated by NUS and are Not present in the raw patent data)				
10.	ISSUEYEAR	--	INTEGER	Year of Issue of the Patent. This is obtained from the TBL_PATN (Field : PATN_ISSUE_YEAR)

Notes:

- The ISSUEYEAR field was duplicated in TBL_INVT in order to speed up searching by inventor. This way, a dynamic join isn't needed to TBL_PATN.

3. Table Name : TBL_ASSG

Logical Group : Assignee

(Multiple Assignees can be present for a single patent number)

ID Code : ASSG

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	ASSG_CUSIP (FK)	Not Null	????????	???????? (For Future Use, to link to Financial Database)
3.	ASSG_NAM	Not Null	TEXT	Assignee Name
4.	ASSG_COD	Not Null	CHAR(2)	Assignee Type Code
5.	ASSG_CTY	--	TEXT	City
6.	ASSG_STA	--	CHAR(2)	State
7.	ASSG_CNT	--	CHAR(3)	Country
8.	ASSG_ZIP	--	CHAR(5)	Zip Code
9.	ASSG_ITX	--	TEXT	Assignee Descriptive Text
CUSTOM FIELDS (Fields that were generated by NUS and are Not present in the raw patent data)				
10.	ISSUEYEAR	--	INTEGER	Year of Issue of the Patent. This is obtained from the TBL_PATN (Field : PATN_ISSUE_YEAR)

Note : ASSG_CUSIP will be obtained from financial database. It is *not* present in the patent-ASCII file.

4. Table Name : TBL CLAS

Logical Group : Classification
ID Code : CLAS

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	CLAS_OCL	--	VARCHAR(255)	Primary US Patent Classification (http://www.uspto.gov/go/classification/)
3.	CLAS_EDF	--	CHAR(1)	Edition Field

5. Table Name : TBL CLAS XCL

Logical Group : Classification-XCL
ID Code : CLAS

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	CLAS_XCL	--	VARCHAR(255)	Cross Reference (Secondary US Patent Class)

6. Table Name : TBL CLAS UCL

Logical Group : Classification-UCL
ID Code : CLAS

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	CLAS_UCL	--	VARCHAR(255)	Unofficial Reference

7. Table Name : TBL CLAS DCL

Logical Group : Classification-DCL
ID Code : CLAS

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	CLAS_DCL	--	VARCHAR(255)	Digest Reference

8. Table Name : TBL CLAS ICL

Logical Group : Classification-ICL
ID Code : CLAS

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	CLAS_ICL	--	VARCHAR(255)	International Classification http://www.uspto.gov/go/classification/international.htm

9. Table Name : TBL CLAS FS

Logical Group : Classification-FS (These are fields searched by the patent examiner)
ID Code : CLAS

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	CLAS_FSC	--	CHAR(3)	Field of Search Class
3.	CLAS_FSS	--	TEXT	Field of Search Subclasses

10. Table Name : TBL_UREF

Logical Group : U.S. Reference (These are references made by each patent in the database to other US Patents)

ID Code : UREF

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	UREF_PNO	--	VARCHAR(255)	Reference Patent Number (?????)
3.	UREF_ISD	--	DATE	Issue Date
4.	UREF_NAM	--	TEXT	Patentee Name
5.	UREF_OCL	--	VARCHAR(255)	US Classification
6.	UREF_XCL	--	VARCHAR(255)	Cross Reference
7.	UREF_UCL	--	VARCHAR(255)	Unofficial Reference
CUSTOM FIELDS (Fields that were generated and are Not present in the raw patent data)				
8.	NEW_UREF_PNO		CHAR(9)	This field is the same as the UREF_PNO field. This was created since certain SQL statements needed a datatype CHAR instead of datatype VARCHAR.
9.	<i>NEW_UREF_ISSUEYEAR</i>		<i>INTEGER</i>	<i>(Please ignore this field. This field is NOT REQUIRED at all. It was created with a certain purpose in mind, which later was not required.)</i>

Note : Only one of the 3 (UREF_OCL or UREF_XCL or UREF_UCL) will exist for a given UREF.

11. Table Name : TBL FREF

Logical Group : Foreign Reference (These are references made by each patent in the database to foreign patents)

ID Code : FREF

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	FREF_PNO	Not Null	VARCHAR(255)	Foreign Patent Number
3.	FREF_ISD	--	DATE	Issue Date
4.	FREF_CNT	--	CHAR(3)	Country Code
5.	FREF_OCL	--	VARCHAR(255)	Classification
6.	FREF_ICL	--	VARCHAR(255)	International Classification

12. Table Name : TBL CLMS

Logical Group : Claims Description

ID Code : CLMS

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	CLMS_NUM	Not Null	INTEGER	Claims Number
3.	CLMS_TEXT	--	TEXT	Claims Text

13. Table Name : TBL DCLM

Logical Group : Design Claims

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	DCLM_TEXT	--	TEXT	Design Claim Text

14. Table Name : TBL COUNTRY CODE

Logical Group : Country Codes

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	CNT_NAME	Not Null	VARCHAR(255)	Country Name
2.	CNT_CODE	Not Null	CHAR(3)	Country Code

15. Table Name : TBL STATE CODE

Logical Group : State Codes (of USA)

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	STA_NAME	Not Null	VARCHAR(255)	State Name
2.	STA_CODE	Not Null	CHAR(2)	State Code

16. Table Name : TBL PATENTTYPE CODE

Logical Group : Types of Patent

[For field TBL_PATN->PATN_APT, Application Type (APT)]

Desc: This field contains a single-digit number which indicates the type of patent, as follows:

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	TYPE_CODE	Not Null	INTEGER	Code for Patent Type
2.	TYPE_DESC	Not Null	VARCHAR(255)	Description of Type

Values:

1 = Utility

2 = Reissue

4 = Design

5 = Defensive Publication

6 = Plant

7 = Statutory Invention Registration

17. Table Name : TBL ASSGTYPE CODE

Logical Group : Types of Assignee Code

[For field TBL_ASSG->ASSG_COD, Assignee Type Code (COD)]

Desc: Defines the type of assignee and is a two digit numeric code (01-09), as follows:

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	TYPE_CODE	Not Null	CHAR(2)	Code for Assignee Type
2.	TYPE_DESC	Not Null	VARCHAR(255)	Description of Type

Values:

01 = Unassigned

02 = United States company or corporation

- 03 = Foreign company or corporation
- 04 = United States individual
- 05 = Foreign individual
- 06 = U.S. Federal government
- 07 = Foreign government
- 08 = U.S. county government
- 09 = U.S. state government

18. Table Name : TBL PRIR

Logical Group : Foreign Priority
(Multiple Foreign Priorities can be present for a single patent number)
ID Code : PRIR

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	PRIR_CNT	Not Null	CHAR(3)	Country Code
3.	PRIR_APD	--	DATE	Priority Application Date
4.	PRIR_APN	--	VARCHAR(255)	Priority Application Number

19. Table Name : TBL REIS

Logical Group : Reissue
(Multiple Reissues can be present for a single patent number)
ID Code : REIS

S/N	Field name	Required/Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	REIS_COD	--	CHAR(2)	Reissue Code

3.	REIS_APN	--	CHAR(6)	Application Number
4.	REIS_APD	--	DATE	Application Filing Date
5.	REIS_PNO	--	CHAR(9)	Patent Number
6.	REIS_ISD	--	DATE	Issue Date

20. Table Name : TBL_RLAP

Logical Group : Related U.S. Application Data
(Multiple RLAPs can be present for a single patent number)
ID Code : RLAP

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	RLAP_COD	Not Null	CHAR(2)	Parent Code
3.	RLAP_APN	--	CHAR(6)	Application Number
4.	RLAP_PSC	--	CHAR(2)	Patent Status Code
5.	RLAP_APD	--	DATE	Application Filing Date
6.	RLAP_PNO	--	CHAR(9)	Patent Number
7.	RLAP_ISD	--	DATE	Issue Date

21. Table Name : TBL_PCTA

Logical Group : PCT Information
ID Code : PCTA

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	PCTA_PCN	Not Null	CHAR(14)	PCT Number
3.	PCTA_PD1	--	DATE	PCT 371 Date

4.	PCTA_PD2	--	DATE	PCT 102(e) Date
5.	PCTA_PD3	--	DATE	PCT Filing Date
6.	PCTA_PCP	--	CHAR(8)	PCT Publication Number
7.	PCTA_PCD	--	DATE	PCT Publication Date

22. Table Name : TBL LREP

Logical Group : Legal Information

ID Code : LREP

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	LREP_FRM	--	VARCHAR(255)	Legal Firm Name
3.	LREP_FR2	--	VARCHAR(255)	Principal Attorney Name
4.	LREP_AAT	--	VARCHAR(255)	Associate Attorney Name
5.	LREP_AGT	--	VARCHAR(255)	Agent's Name
6.	LREP_ATT	--	VARCHAR(255)	Attorney Name
7.	LREP_REG	--	CHAR(6)	Registration Number
8.	LREP_NAM	--	VARCHAR(255)	Representative Name
9.	LREP_STR	--	VARCHAR(255)	Street
10.	LREP_CTY	--	VARCHAR(255)	City
11.	LREP_STA	--	CHAR(2)	State
12.	LREP_CNT	--	CHAR(3)	Country
13.	LREP_ZIP	--	CHAR(5)	Zip Code

The following logical groups are combined to form one table

Logical Group : Other Reference	ID Code : OREF	Optional
Logical Group : Government Interest	ID Code : GOVT	Optional
Logical Group : Parent Case	ID Code : PARN	Optional
Logical Group : Drawing Description	ID Code : DRWD	Optional
Logical Group : Detail Description	ID Code : DETD	Optional

23. Table Name : TBL_OTHERS

S/N	Field name	Required/ Optional	Field Type/Size	Field Desc
1.	PATN_WKU (FK)	Not Null	CHAR(9)	Patent Number
2.	OREF_TEXT	--	TEXT	References to Scientific Papers and Other Publications
3.	GOVT_TEXT	--	TEXT	Government Interest
4.	PARN_TEXT	--	TEXT	Parent Case Text
5.	DRWD_TEXT	--	TEXT	Drawing Description
6.	DETD_TEXT	--	TEXT	Detailed Description