

IP Data Quality Management in JPO

Oct. 19. 2015

JAPAN PATENT OFFICE



- Population : 0.127 billion
- Capital city : Tokyo





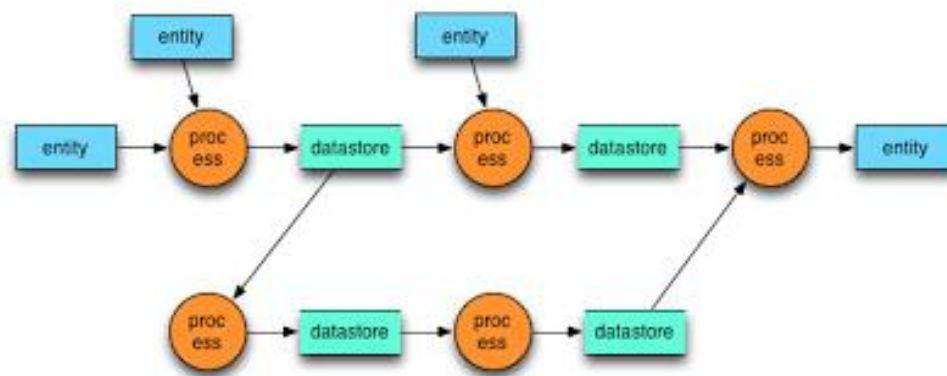
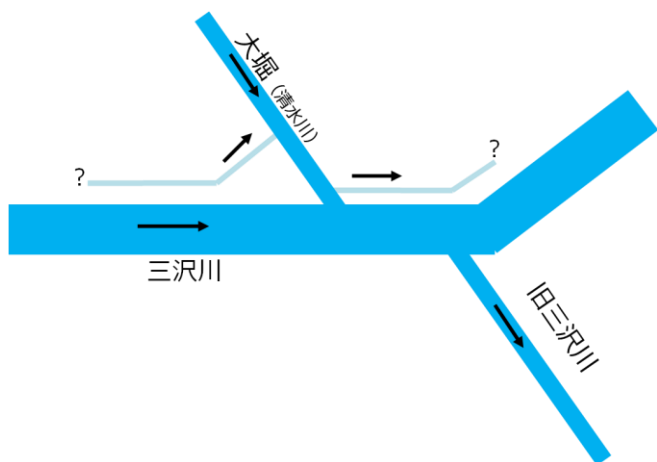
Marugami Falls(Saitama)



Chichibu-Kegon Falls(Saitama)



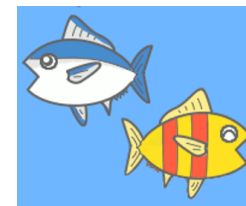
- Upstream to downstream
- Junctions (merge / separate)
- Quality of water (clear / turbid)



- If quality of water is bad(turbid)...
 - Fishes go away, decreasing QOL of human, ...
- How to improve the quality of water?
 - Improvement in the upstream is preferable



- If quality of water is improved...
 - Many fishes, became sightseeing spots, increasing QOL,...



- 2005 : complete graduate school (physics)
- **2005 : join JPO (Assistant Examiner)**
- 2009 : Examiner (man-machine interface)
- 2010 : Assistant Director (search outsourcing)
- 2011 : Examiner (man-machine interface)
- **Apr 2014 – Sep 2015 : Deputy Director (data management)**
- Now : Examiner (digital communications)

1. Introduction

2. Organizations

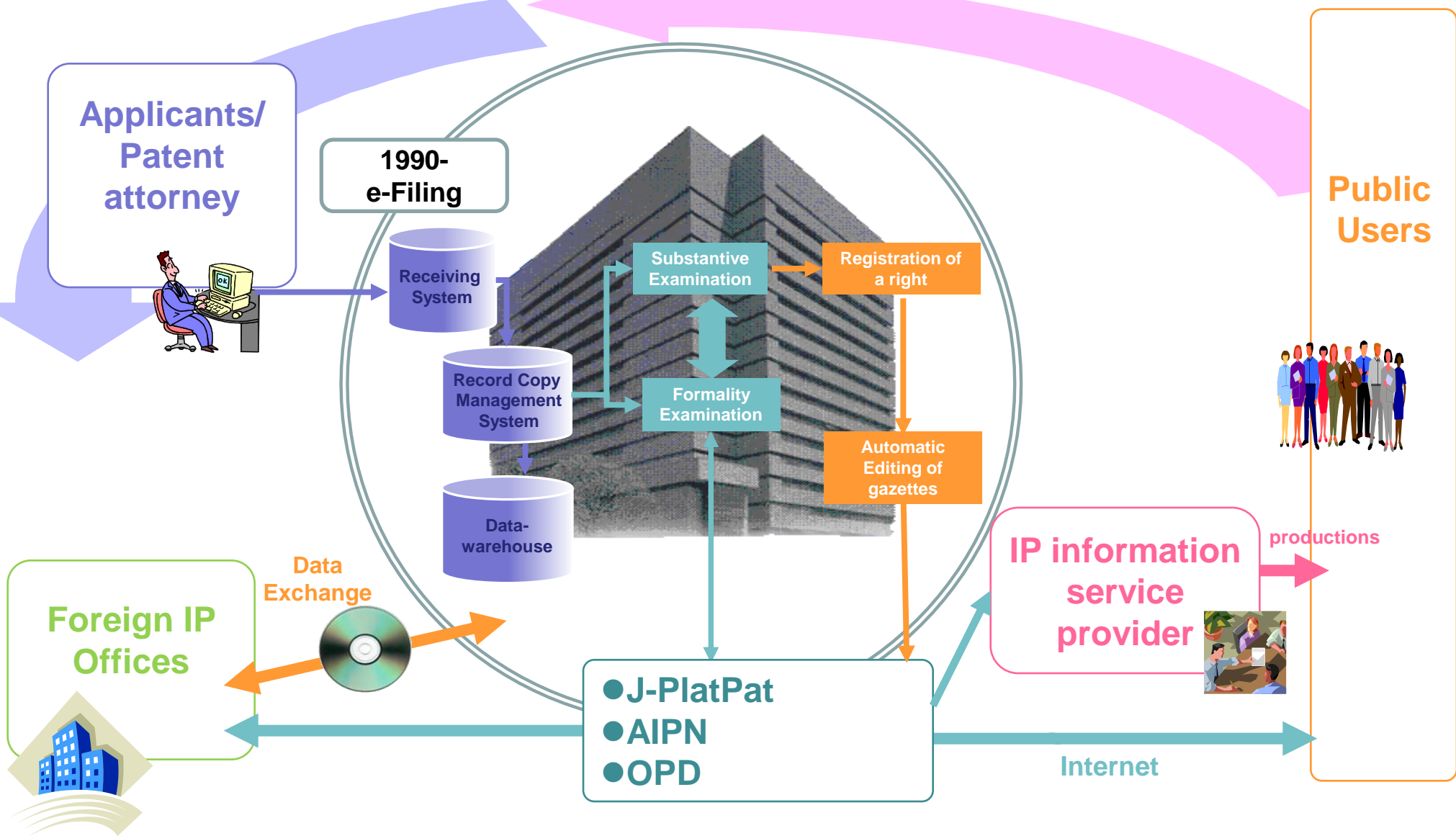
3. Duties of Data Quality Improvement

1. Introduction

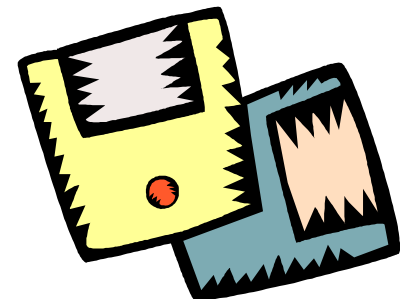
2. Organizations

3. Duties of Data Quality Improvement

Outline of IP information flow

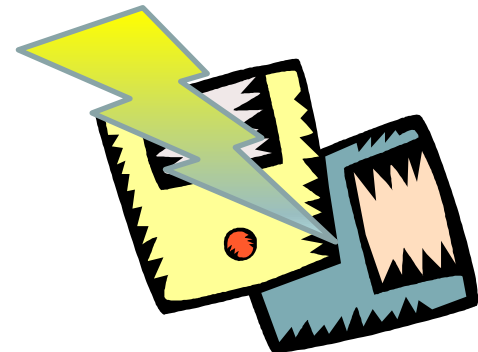
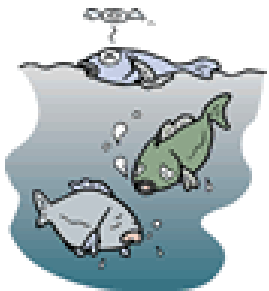


- Information of Application management
 - Application, publication, registration...they are critical information of Patent rights.
- Gazette Information (Internal / External use)
 - We are receiving / accumulating / providing of Foreign Documents of patent, utility model, design, trademark.
 - The large number of documents.



- **Information of Application management**
 - Although the number of errors is small in IP Office, they directly affect the critical information of IP rights.
 - A Data correction requires much influence on many internal systems in JPO.
- **Gazette Information (Internal / External use)**
 - The large number of documents may be affected.
 - Can not be accumulated in retrieval system

⇒ **Serious errors may shake public confidence in the IP system !**



Example of Information of Application management

Identification numbers:
For example,

- (A) application number
- (B) filing date
- (C) priority number
- (D) priority date

(E) final action of
JPO's examiner

(F) internationally
unified classification
based on IPC

[基本項目]		(* 処分済 *)	
出願 (1) (11-123456)(11.04.30)	記号 (J1716)	出願種別(01)	新法
公開 (2000-312770)(12.11.14)	公開基準日 (11.04.30)	国内優先 (0)	
公告 () () ()	優先 () () ()	他 国	
審判 () () ()	担当 () () ()		
登録 () () ()	異議 (0) 請求項数 (6)	出願料金 (21,000)	
公決 (起) (担)	文献 () 新規性 (0)	菌寄託 (0) 公害 ()	
査定 () (起) (担)	前置 () () ()	解除 () 公序・要約 ()	
(完) (官)	審査・評価請求 (0-)	未請求(0) 自動起案 ()	
最終 (A09) (18.07.25)	公開準備 (1)	早期審査 ()	
変更 () () ()	審決 () () ()		
	原出願 () () ()	種別 () () ()	
	期間延長 () () ()	最新起案日 (11.08.02)	
公表 () () ()	翻訳提出 () () ()	国際出願 () () ()	
再公表 () () ()	国際公開 () () ()		
公開IPC = A63F 7/02 330 FN	指定分類IPC		
公告IPC			
名称 検査装置			
出願人 代表 () 種 ()	群馬県桐生		
	代表 () 種 ()		
	岐阜県岐阜市南町		
代理人 種(1) コト (100000000)			
中間 (A05) 特許願 11.04.30 (21,000) 完 (A82-1) 手続補足 11.05.08 () 完			
記録 (A11-2) 中間指令 11.08.10 (6895-) A (A52-3) 補正書 11.08.25 () 完			
(A96-7) 認定情報 11.09.01 () (A96-7) 認定情報 11.09.01 ()			
(A200) 未請求票 18.07.04 ()			
新出願			
国内優先 (先)			

Bibliographic data:
For example,

(G) applicant
(H) inventor

(I) record of
documents
between applicant
and examiner

Example of Patent Information (Publication of Unexamined Patent Application)

(19)日本国特許庁(JP)	(12)公開特許公報(A)	(11)特許出願公開番号 特開2000-244579 (2000-244579A) (43)公開日 平成12年5月20日(2000.5.20)
(31)Int. Cl. ⁷ G01B 3/00 G02C 26/00 23/02	識別記号 101	F I G01B 3/00 101 A G02C 26/00 23/02 A45C 12/00 101 A A47B 23/02
審査請求 未請求 請求項の数 1 O L 外国語出願 公開請求 (全6頁) 最終頁に続く		
(21)出願番号 (22)出願日 (31)優先権主張番号 (32)優先日 (33)優先権主張国	特願平11-123456 平成11年11月10日(1999.11.10) 83304359.9 平成10年11月12日(1998.11.12) フランス(FR)	(71)出願人 390003011 パテント コーポレーション Patent Cooperation アメリカ合衆国ケンタッキーレイビル ビー・オー・ボックス 35090 レイビルガ レリアブラウン タワー 1500 (無番地) 090000423 日本特許発明株式会社 東京都千代田区内幸町4丁目5番6号 (72)発明者 発明 太郎 神奈川県横浜須賀町1丁目2200番地 (74)代理人 123456789 弁理士 代理 太郎 (外2名)
特許法第30条第1項適用申請有り 平成10年9月21日付 画像工学会研究専門委員会主催の1992年度画像符号化シ ンポジウム(RSCJ92)において文書をもって発表 特許法第85条の2第2項第4号の規定により明細書及び 図面の一部は不掲載とする。		
(54)【発明の名称】ファクシミリ走査装置		
(57)【要約】 (修正有) 【目的】ファクシミリ端末パラメータ識別方法に関し、 ファクシミリ装置機能のパラメータ拡張を容易にする。 【構成】通信時の端末パラメータを識別する方法におい て、端末パラメータを含む制御信号の送信端末1a、 bは制御信号のファクシミリ情報フィールドを、複数の サブフィールドに分離し、各サブフィールドの情報を分 離するファクシミリ情報フィールドのデータ中には現れ ない特定の識別コードを挿入してファクシミリ情報フィ ールドを作成する。制御信号の受信端末7はファクシミ リ情報フィールド内の上記特定の識別コードを検出し、 ファクシミリ情報フィールドを複数のサブフィールドに 分離して、各サブフィールドの情報の内容を解析し相手 端末の端末パラメータの内容を検出する。装置機能のパ ラメータを拡張する場合はユニークコードを挿入して可 変長の端末パラメータを分離する。		

(F) internationally unified classification based on IPC

Identification numbers:
For example,
(A) application number
(B) filing date
(C) priority number
(D) priority date

Abstract of the present invention

Publication type:
“A” means publication of unexamined patent application.
“B” means that of examined (granted) patent application.

Bibliographic data:
Necessary information is retrievable.
(G) applicant
(H) inventor

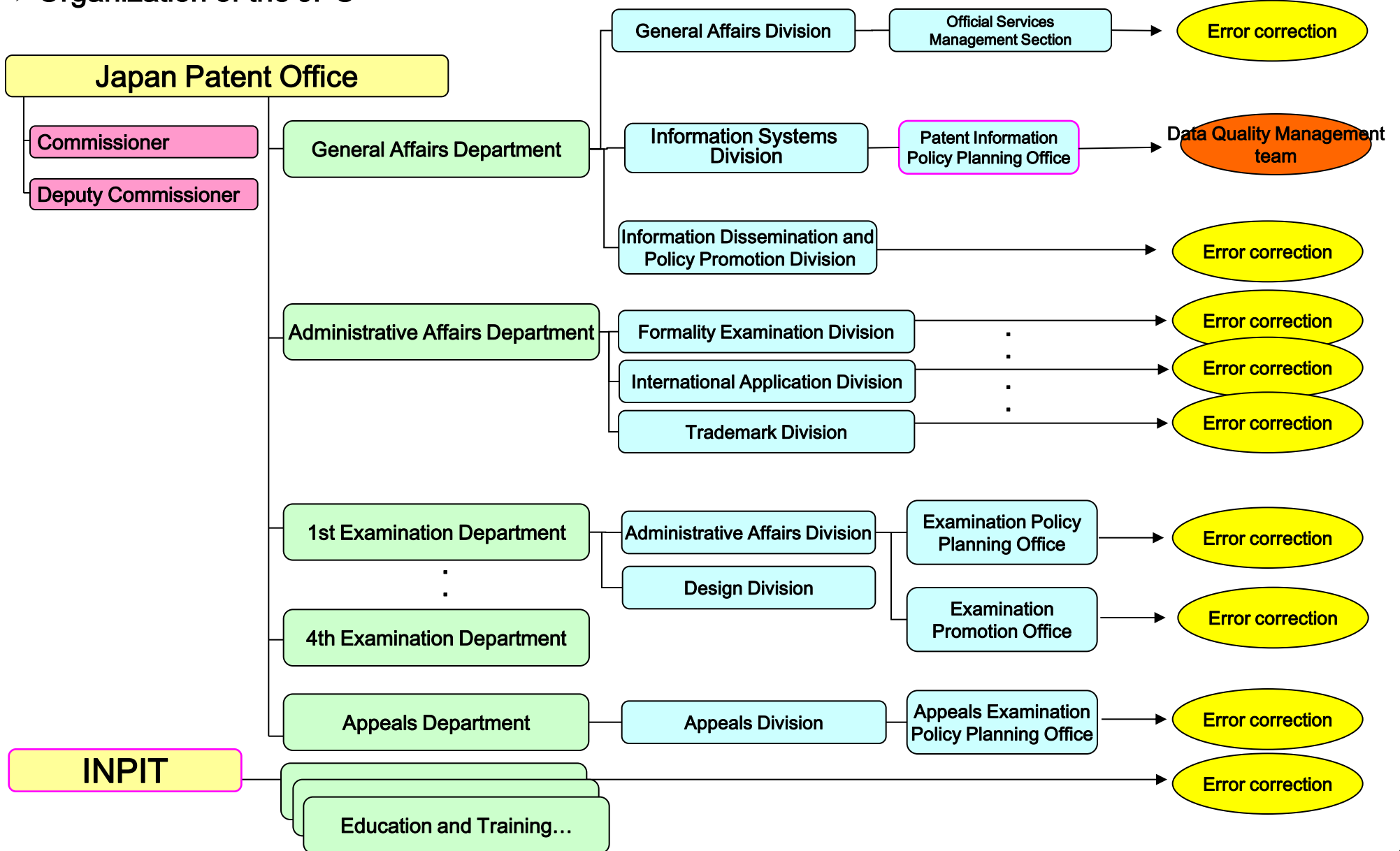
Representative drawing of the present invention

1. Introduction

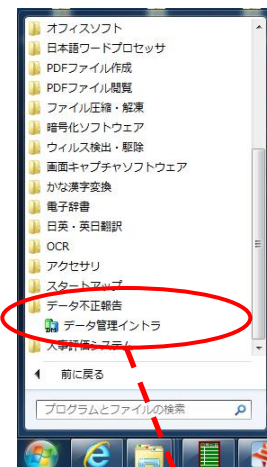
2. Organizations

3. Duties of Data Quality Improvement

Organization of the JPO



Internal user (JPO)



Shortcut in Start-menu of PC of all JPO employees

Mailer launcher



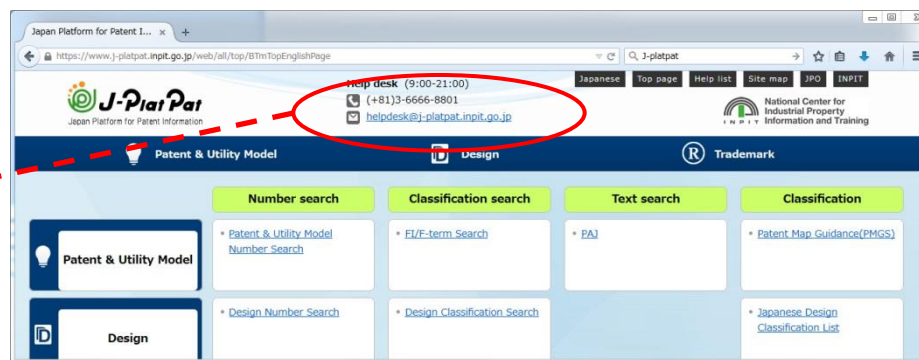
Data management team /
Related divisions



Analysis, Classification,
Research, Correction,
Share (feedback)

External user

HELP DESK (J-PlatPat)



<https://www.j-platpat.inpit.go.jp/web/all/top/BTmTopEnglishPage>

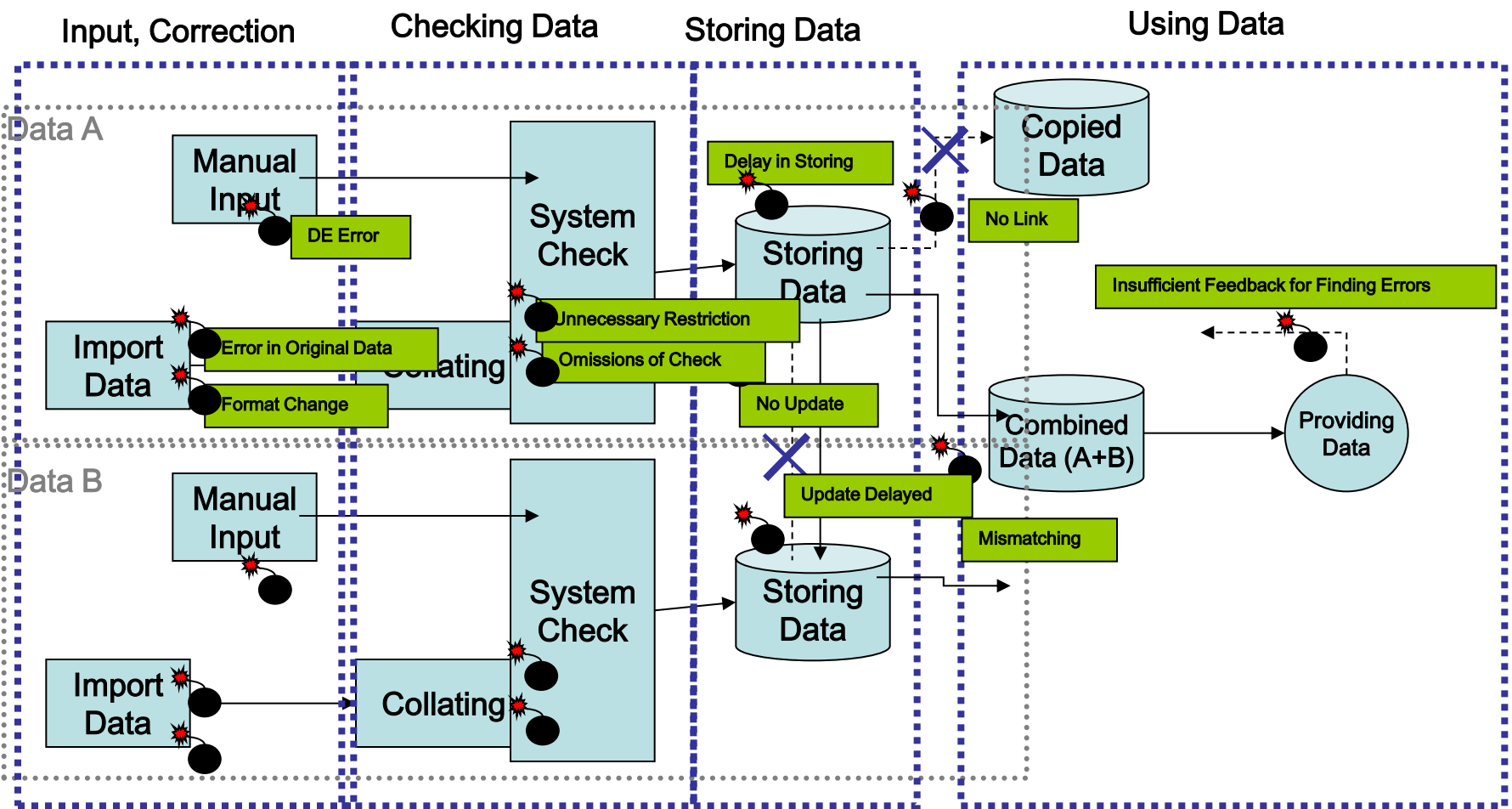
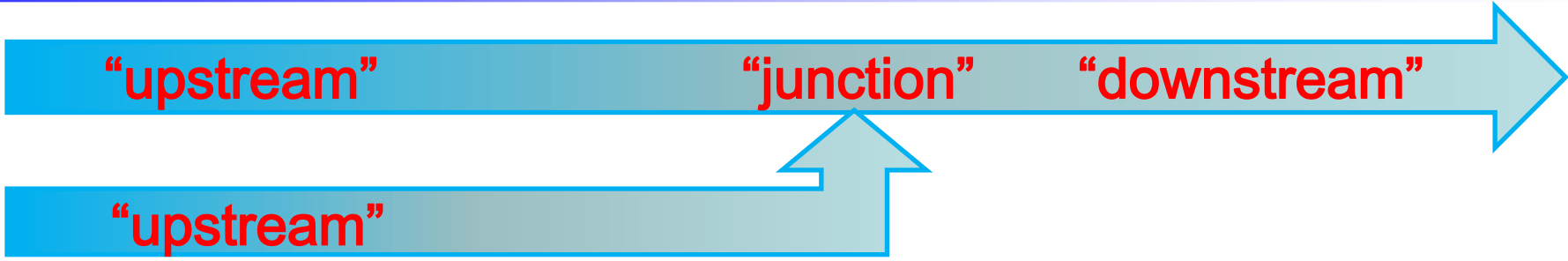
- **Leader (1)**
 - **Sub-leader (1)**
 - **Associate (1)**
 - **IT specialist (3)**
-
- **Review of progress (monthly)**
 - **Adjustment with related divisions**
 - **Communication with foreign offices**
 - **Feedback to reporters of the errors**

1. Introduction

2. Organizations


3. Duties of Data Quality Improvement

Where do errors occur?



- Overview

- Counting the number of errors and corrections(monthly)



water quality
survey

- Actions

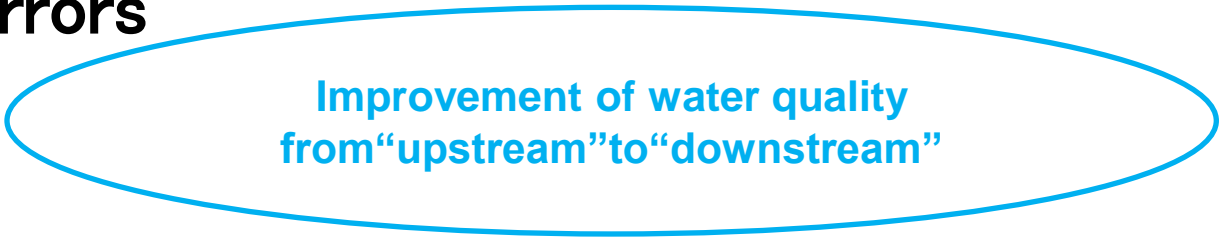
- Prevention of Errors

- Sharing error cases with other Departments / IP types (e.g. Patent → Trademark)

- Monitoring Errors

- Observation of the specific event which may cause the error

- Correction of Errors



Improvement of water quality
from“upstream”to“downstream”

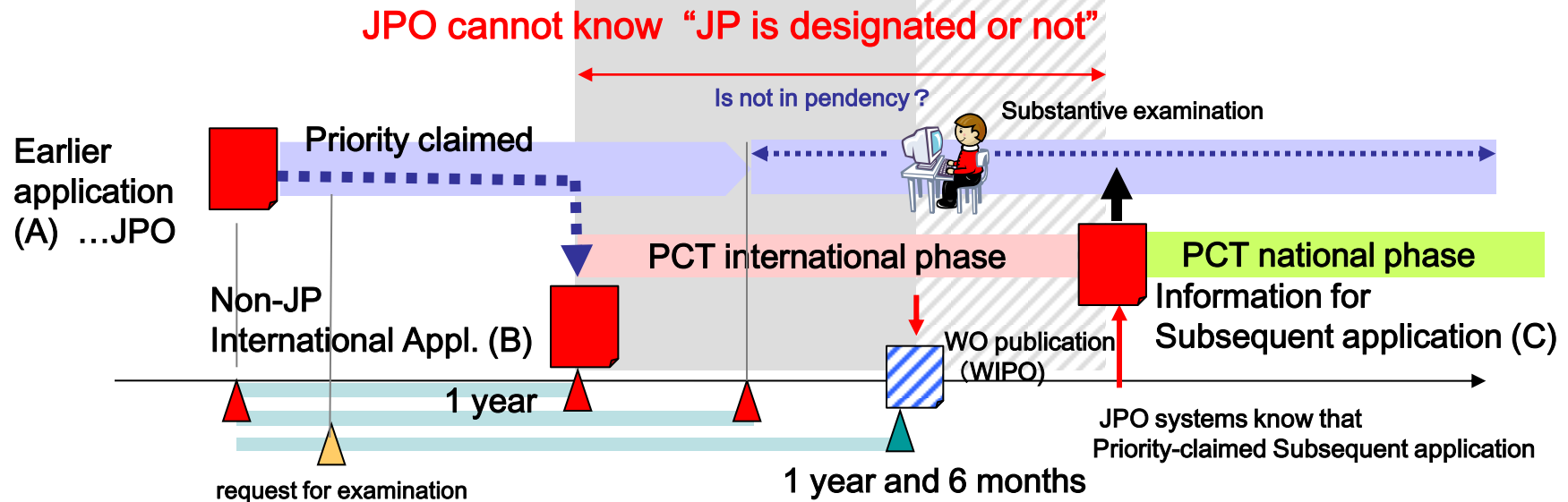
Example1 : Prevention of Errors

[基本項目] (* 処分済 *)

出願 (1) () () (11.04.30) 記号 (J1716) 出願種別(01) 新法
 公開 (20) () (12.11.14) 公開基準日 (11.04.30) 国内優先 (0)
 公告 () () () 優先 () 他 国
 審判 () () () 担当 () ()
 登録 () () () 異議 (0) 請求項数 (6) 出願料金(21,000)
 公決 (起) (担) 文献 () 新規性 (0) 菌寄託 (0) 公害 ()
 査定 () (起) (担) 前置 () 解除 () 公序・要約 ()
 (発) (官) 審査・評価請求 (0-) 未請求 (0) 自動起案 ()
 最終 (A09)(18.07.25) 公開準備 (1) 早期審査 ()
 変更先 () () () 審決 () ()
 原出願 () () () 種別 ()
 期間延長 () 最新起案日 (11.08.02)
 公表 () () () 翻訳提出 () 国際出願 ()
 再公表 () () 国際公開 ()
 公開IPC # A63F 7/02 330 FN 指定分類IPC
 公告IPC
 名称 検査装置
 出願人 代表 () 種 (2) コト (000132747) 国 (10) 株式会社ソフィア *
 群馬県桐生市境野町茅ノ丁目2-0-1番地
 代表 () 種 (2) コト (000132747) 国 (10) 株式会社 マツワラ商会 *
 代理人 種 (1) コト (100096699) 鹿嶋 英貢
 中間 (A63) 特許願 11.04.30 (21,000) 完 (A82-1) 手続補足 11.05.06 () 完
 記録 (A11-2) 中間指令 11.08.10 (6895-) A (A52-3) 補正書 11.08.25 () 完
 (A96-7) 認定情報 11.09.01 () (A96-7) 認定情報 11.09.01 ()
 (A300) 未請求票 18.07.04 ()
 新出願
 国内優先(先)

The last 1 digit of the applicant code is check digit, and we can find some inconsistency before data entry.

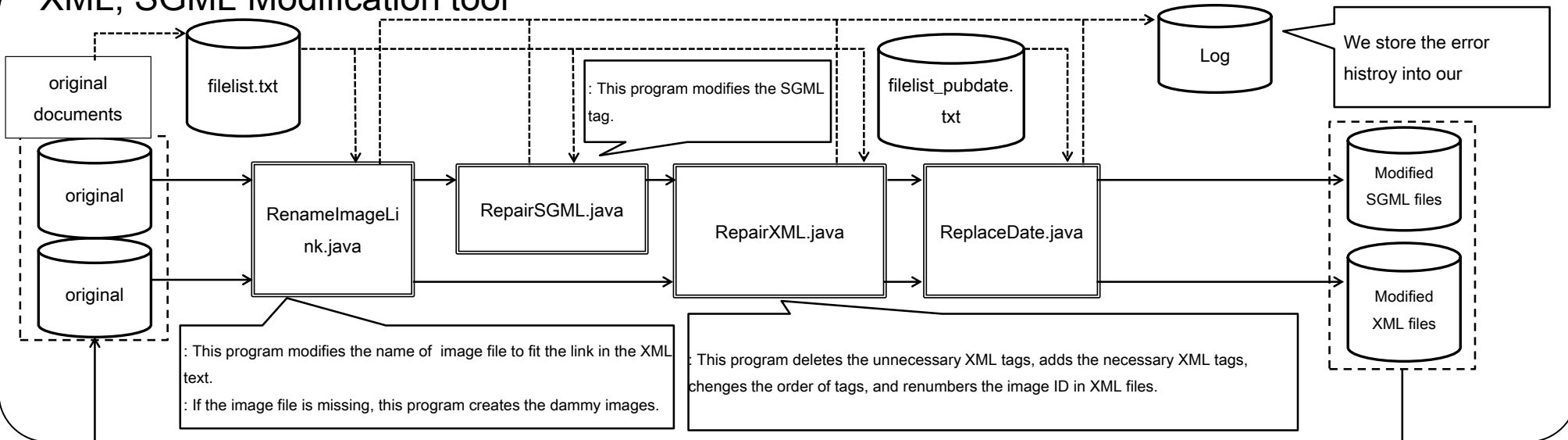
Example2 : Monitoring Errors



- The earlier application (A) should be deemed to have been withdrawn, but JPO sometimes gets behind in recognition under the following conditions.
 - The earlier application (A) is filed with the JPO.
 - The PCT application (B) is based on the earlier application (A), and filed with the other ISA.
 - Japan is designated in the PCT application (B), and the application (C) is filed with the JPO.
- Keep Monitoring the WO gazette, and minimizing the period of JPO not knowing

Example3 : Correcting Errors

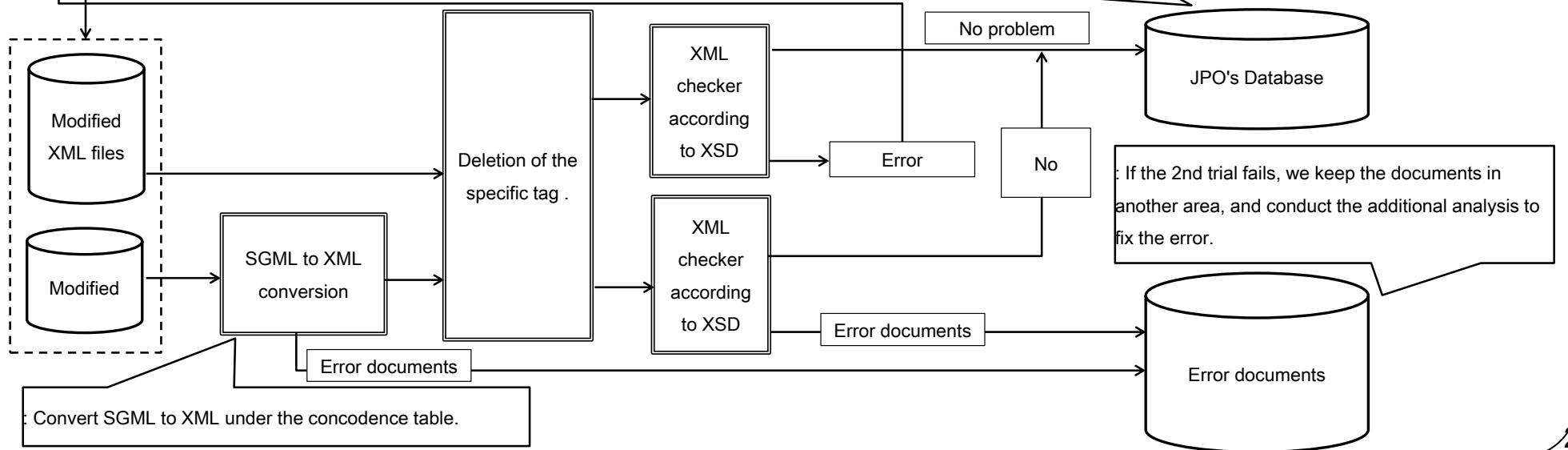
XML, SGML Modification tool



Document translation and storage system

Return the error document to the error modification tool.

The system translates foreign language into Japanese, and stores the database.



Example3 : Correcting Errors

The XML check program...

1. deletes the unnecessary XML tags that XSD does not defines.
2. adds the necessary XML tags that XSD defines.
3. changes the order of tags according to the XSD.
4. rennumbers the image ID in XML files according to the XSD.

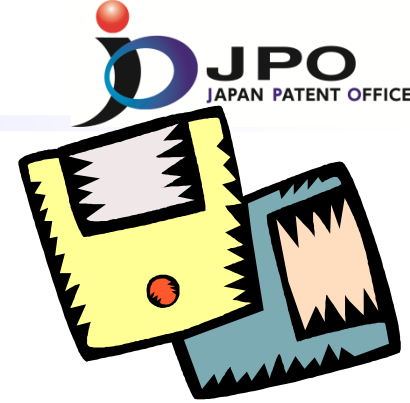
Example

```
<Bibliographic>
  <!-- Temp tag<!DOCTYPE SYSTEM>Temp tag -->
  <!-- Temp tag<APSVER="2.2"><PATDOC>Temp tag -->
  <DP n="1" type="SOFT"/>
  <PatentDOC>
```

The program deletes the tag <DP> that XSD does not define.

```
<Bibliographic>
  <!-- Temp tag<!DOCTYPE SYSTEM>Temp tag -->
  <!-- Temp tag<APSVER="2.2"><PATDOC>Temp tag -->

  <PatentDOC>
```



- We have many error checking tools that confirm the consistency of XML data, corruption of image data that we have received from the foreign offices.
- We are correcting the data error with the error modification tools.
- If you provide us with your gazettes, we check and modify the data consistency.
- We hope we will be of some help.

paper image of old gazettes

後記号なし
特許願(3)
昭和47年4月11日
特許庁長官殿

1. 発明の名称
カソコセイシエンエオ アイリヨクホウキョウキョウノホウ
感光性樹脂版の定置換新方法

2. 発明者
タカツキ シンゴウジ
大阪府高槻市安岡寺3-7-4 付録 改頁

3. 特許出願人
大阪府北区藤田1番地(ほか1名)
(900) 帝人株式会社
代表者 大 塚 善 三

4. 代理人
東京都千代田区内幸町2丁目1番1号
(東京ビル)
帝人株式会社内
(6572) 和歌山 佐藤 弘 徳
和歌山(200-1181) 高 山 秀

5. 発明の要旨
特許請求の範囲
ストレンジタンタニク貯蔵された感光性樹脂版を、
を蒸留より押し出して貯置庫へ搬入させ、版面が
被定置庫に達したことを検出して搬入を停止し、
しかる後、貯置された感光性樹脂版を加圧気流
によつて貯置庫より押し出し、印刷版製作
へ供給することを特徴とする感光性樹脂版
定置換方法。

①特開昭 49-1
③公開日 昭49.(1974)1.5
②特願昭 47-3846Z
②出願日 昭47.(1972)4.17
審査請求 未請求 (全3頁)

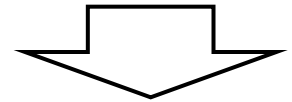
庁内整理番号 ⑤日本分類
6670 46 116 A73
6710 25 116 A7Z
6670 46 116 A73
6670 46 116 A7Z
6906 46 103 H0

47.4.18
47 038462

方法
ト(少な
側版を圧
いる。

OCR text data

このような開一を解決するた○の方法として、
槽一に仕上げられた2枚のプレート(少なく
と・41枚44 \$j iiクラスある)の礪に側8
m液を圧入するというd&巢がなされている。



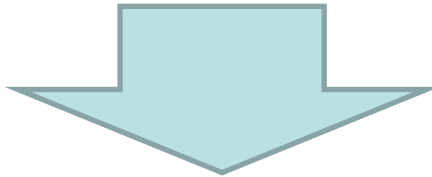
machine translation

As a method of O with resolve, - open
like this Tank - two plates was finished
to-least (The pressure side 8m solution
of boric) some 44 \$ j ii class 41 pieces D
& nest entrance that have been made.

OCR error sometimes causes mistranslation
of machine translation

- The old gazettes are based on the paper document.
- We obtain image data by scanning a paper documents, and retrieve the character data from the images by OCR (optical character recognition)
- The character data contain many recognition errors.

- There are many errors in the OCR recognition results, and they are not suitable for the machine translation.



- We will select the important gazettes and correct the OCR recognition errors manually.
- We will provide the corrected texts via J-PlatPat in the near future, and you will understand the contents using machine translation.
 - URL : <https://www.j-platpat.inpit.go.jp/>



Arakawa river
(Saitama, Nagatoro)



KOBATON
(Mascot character of Saitama)



Based on “shirako-bato”
(National monument in Japan)



Thank you for your attention.