ANNEX 5

STANDARD FOR ELECTRONIC FILING, PROCESSING, STORAGE, AND RECORDS MANAGEMENT OF INTERNATIONAL APPLICATIONS UNDER THE PATENT COOPERATION TREATY (PCT)

Annex F, Version 3.1

Management Summary

This document presents the DRAFT standard for Electronic Filing, Processing, Storage, and Records Management of International Applications under the authority of the relevant Administrative Instructions and pursuant to Rule 89bis of the PCT.

It contains the basic technical and certain legal principles to be adopted for electronic filing and references Technical Appendixes for the details of specific implementations. The current version references Appendix I and II - The Trilateral Technical Standard for the On-line Exchange of Intellectual Property Documents using PKI. Other Appendices may be developed to cover implementations of document exchange on disk and where no PKI is available.

Table of Contents

1 Definitions of relevant terms

For the purposes of Annex F, the expression:

(a) “electronic signature” means data in electronic form in, affixed to, or logically associated with, a data message, and any method in relation to a data message that may be used to identify the signature holder in relation to the data message and indicate the signature holder’s approval of the information contained in the data message;

(b) “enhanced electronic signature” means an electronic signature in respect of which it can be shown, through the use of a security procedure, that the signature:

(i) is unique to the signature holder within the context in which it is used;
(ii) was created and affixed to the data message by the signature holder or using a means under the sole control of the signature holder and not by any other person;
(iii) was created and is linked to the data message to which it relates in a manner which provides reliable assurance as to the integrity of the message”;

One implementation of this is a “digital signature” which is produced using a Public Key Infrastructure generated certificate and corresponding private key.

(c) “data message” means information generated, sent, received or stored by electronic, optical or similar means;

(d) “document integrity check” means a mechanism by which the integrity of a document can be checked both during transmission and storage and at a later date. A digital signature can be used for this purpose;

(e) “signature holder” means a person by whom, or on whose behalf, an enhanced electronic signature can be created and affixed to a data message;

(f) “certification authority” means a person or entity which, in the course of its business, engages in providing identification services which are used to support the use of enhanced electronic signatures.

(g) “certificate” means a data message or other record which is issued by an certification authority and which purports to ascertain the identity of a person or entity who holds a particular signature device;

2 Requirements for applicants concerning the electronic filing of International Applications

This Annex contains a set of recommended practices concerning the submission of electronic documents under the PCT based on the technical implementations set out in the Technical Appendices.

2.1 Submission

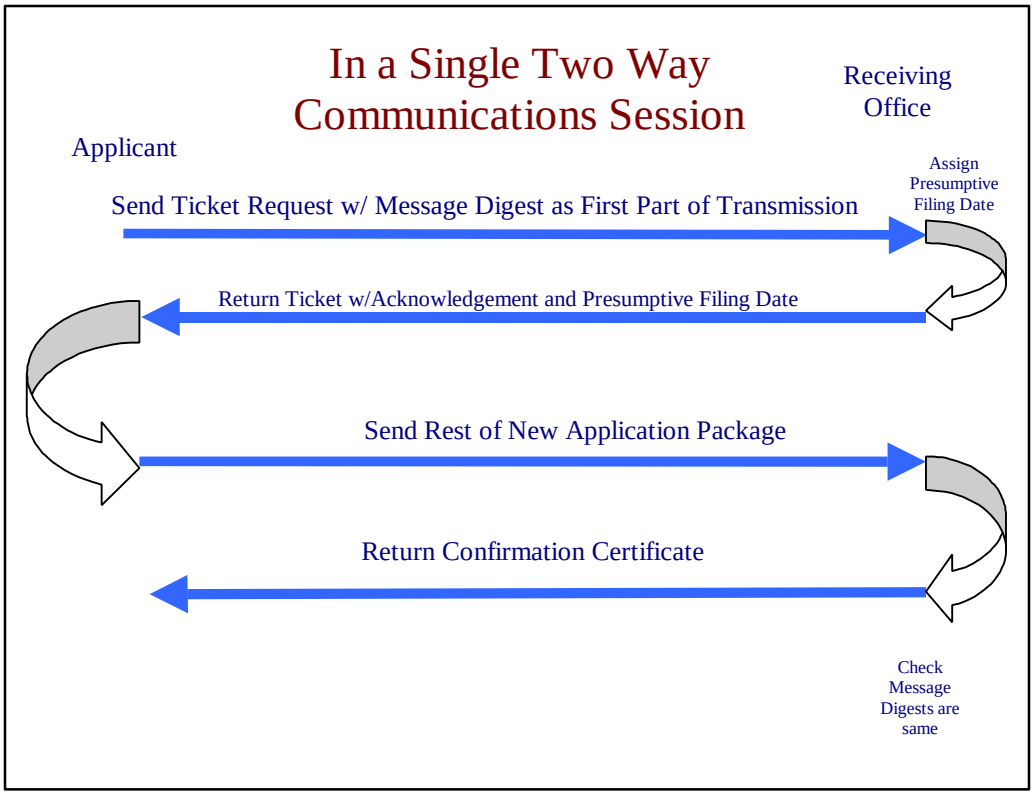
The information exchanged during a transaction is broken into packages. Each phase of the exchange corresponds to the transfer of a package of data sent between the Application and the RO.

For all transactions, there will be the following four packages:

- Ticket Request
- Ticket
- IP Document
- Confirmation Certificate.

For each type of electronic data exchange, the IP Document will contain the data actually prepared by the applicant (e.g. New Application, Fee Payment, Replacement Claims etc.)

In the specific case of PCT On-line Filing, this "**Ticket Mechanism"** operates as follows. :



- An electronic session is established between the applicant and the receiving Office for the purpose of filing an international patent application.
- Near the beginning of that session, a message digest code is transmitted to the Office, which is uniquely derived from the combined files constituting the necessary parts of the international application. The code is such that even the tiniest change to any of those files will be indicated by a change in that message digest code.
- On receipt of that message digest code, and as part of the session, the Office sends an acknowledgement to the applicant indicating the date of receipt of the message digest.
- The applicant then continues the session and transmits the complete set of files constituting the international application.

On receipt of the full set of files, the application files are processed to develop their unique message digest. This is compared to the original message digest that was sent at the beginning of the session. If they match, an acknowledgement of receipt is sent to the applicant. If they do not match, the applicant is informed accordingly. The session can then be ended.

This submission mechanism is particularly appropriate for the on-line filing of normal sized files. For this type of submissions, the transmission of the full set of application files is expected to be completed in the single session. However, this submission mechanism should not be used for on-line filing of exceptionally large files. Appendix I contains details concerning acceptable file sizes for on-line submissions.

Technical details of the implementation of this mechanism are given in the Annex Appendix I.

2.2 Security

Security is required in the receipt, storage and processing of electronic patent records to maintain the following qualities of those records:

- Authenticity
- Integrity
- Confidentiality
- Non-repudiation

The authenticity of an electronic record refers to the quality of being what on its face it is presented to be, expressing the intent and purpose of the party who signed or submitted it. The test of authenticity is one of the more critical and potentially difficult tests of admissibility since it is somewhat broadly defined and typically relies more on the testimony of a knowledgeable witness (e.g. The Records Manager) to establish.

The requirement of authentication or identification as a condition precedent to admissibility is satisfied by evidence sufficient to support a finding that the matter in question is what its proponent claims.

This requirement suggests that, in addition to the record having been accurately and reliably created, received, inputted and stored, it was:

- created, transacted and/or communicated by an identifiable and verifiable party for a specific purpose,
- received or created and stored at a specified point in time.

Confidentiality means ensuring that information concerning the content or existence of an international application is not disclosed or revealed to unauthorized persons. Maintaining the *confidentiality* of international patent records is mandated by Article 30 and Art 38 of the PCT. Confidentiality must be maintained until the international publication of the application, if there is one, unless requested or authorized by the applicant. [Insert text from WIPO here]

Patents can only be enforced in courts of law if the patent records are admissible as evidence of the patentee's rights. It is thus essential that they be properly created and protected during and after their pendency, so that they have the necessary credibility to establish the patentee's rights. The key objective of an electronic commerce environment is to assure those qualities of the records. In addition, the PCT and national laws require that the confidentiality of the patent application information be maintained by all Offices and systems processing those files.

2.2.1 Authenticity - Electronic Signatures

In patent law, we typical rely on the **signature** on a document to establish its authenticity. Authenticity is the confirmation that a document is what it claims to be, authored by the person who purported to author it, and expressing the intent of the author. In electronic documents, **electronic signatures** serve the same purpose.

An electronic signature expresses the intent of the signing party to represent his identity and his endorsement of the contents of the document for the purposes indicated in the document. The signature may be:

- a Basic Electronic Signature such as a series of characters chosen by the author to express his or her identity and the intent to sign that document or
- a Enhanced Electronic Signature associated with the document. In this case, the intention to use the Enhanced Electronic Signature is added to the text of the document.

The acceptable formats for electronic signatures are set out in Appendix I.

2.2.2 Integrity – Document Integrity Check

Integrity means ensuring consistency of data, in particular, preventing (including detecting) unauthorized alteration or destruction of data. Maintaining the integrity of patent records is essential for ensuring the protection of the intellectual property rights and the value of the inventor's business assets. For these reasons, checks and balances must be in place to assure the preservation of the integrity, authenticity and trustworthiness of the Intellectual Property Office's records over time. This may be accomplished by managing and protecting the electronic records from any loss, removal, or unauthorized alteration or destruction. Since evidence of record tampering may not be as readily identifiable with electronic records, as it might be with paper records, it is even more important that the controls are in place to adequately protect and preserve the electronic record.

2.2.3 Confidentiality - Encryption

All designs for the receipt, generation, storage, processing and transmission of the electronic records will take all reasonable steps to enable maintaining the confidentiality of the electronic records. All implementations will protect the metadata concerning the electronic records of the international application. The confidentiality of the connection between the applicant and a Receiving Office or between an RO, Authority and the IB will be protected using encryption technology during transmission.

2.2.4 Non-repudiation

This ensures that strong and substantial evidence is available to the sender of data that the data has been delivered (with the cooperation of the recipient), and, to the recipient, of the sender's identity, sufficient to prevent either from successfully denying having sent or received the data. This includes the ability of a third party to verify the integrity and origin of the data.

2.3 Formal Document Requirements

Documents exchanged electronically should, in addition to satisfying all the content requirements of the PCT, meet the technical specifications set out in Appendix I.
Each Receiving Office can choose whether to accept other document formats not specified in Appendix I, but if they do so, then they must invite the applicant to resubmit the application in a supported format before exchange with an Authority or the IB. If such a resubmission took place, the original filed documents must also be transmitted along with the resubmitted ones.

3 Requirements for Offices and Authorities

This section contains a uniform set of procedural requirements (more than simply recommendations) which receiving Offices, International Searching Authorities, International Preliminary Examining Authorities and the International Bureau will be required to use in relation to electronic filings. The purpose of this section is to provide for specific record keeping

procedures that will assure the authenticity, confidentiality, integrity and non-repudiation of the electronic records by all parties in the chain of custody, and thereby assure that these records can be presented as true wherever they may be required.

Whatever automated record management systems may be used by the PCT Offices, they must be designed to protect the confidentiality, preserve the integrity and maintain the authenticity of all submitted and generated electronic documents and records. The automated record management systems and workflow systems must restrict access to only authorized parties and maintain an audit trail of when and by whom the records were accessed.

3.1 Record Copy

The RO will transmit to the IB, within the timeframe set prescribed under PCT regulations, the electronic data of a New Application that it receives from the Applicant. If the document is received in a format consistent with Annex F-Appendix I then the unencrypted data will be transmitted. If the RO has accepted a filing in a format different from Annex F-Appendix I and II, then the RO will invite the applicant to resubmit the data received in an accepted format as described in Annex F – Appendix I and then the RO will also send this resubmitted data to the IB.

Both the original document as submitted and the format, which complies with the standard, will become the Record Copy. Any enhanced electronic signatures required for the resubmitted document will be provided by applicant. Any electronic signatures will be copied unchanged into the new document.

3.2 Exchange of Records

The chain of custody and the records of all changes must be maintained by each custodian organization of the PCT that receives, stores, processes or transmits the electronic records associated with the electronic international application. All changes to the electronic records must be traceable to either a requirement of the Office or an instruction of the applicant, expressed in a suitable document of instruction to make that change.

At the various points in the PCT procedure where one RO, Authority or the IB needs to pass on the records of the procedure to another RO, Authority or IB, then, in addition to transferring the electronic record to that RO or Authority, an electronic copy of the dossier should be made and transmitted to the IB to become part of the permanent Record copy. (This will capture in a safe repository a back-up copy of the dossier as it existed at larger steps in the prosecution).

ROs, Authorities and the IB are not required to pass on copies of the dossier after all individual changes are performed within that Office.

3.3 Integrity of Transmission

Electronic transmission of the dossier includes both the transmission of documents associated with international applications, and the metadata that describes information concerning those documents. The multiple files associated with a dossier and metadata shall be combined in an archive file format for transmission as a single object. A document integrity check shall be applied to the archive file so that integrity of transmission may be ensured by validating the received file with the document integrity check.

3.4 Acknowledgement of Communication

The submission mechanism shall provide indication of successful transmission using its built-in acknowledgement function. Beyond this, to assure the integrity, confidentiality, non-repudiation and authenticity of the electronic dossier, the result of initial validation processing of electronic

dossiers will be expressly acknowledged by the receiving Office, and such acknowledgement will be both transmitted to the sending Office, and placed as a record into the electronic dossier. In instances where the initial validation is unsuccessful, this acknowledgement shall be the means of notifying the sending office of a problem.

3.5 Integrity of Storage

Both the storage media and the recording method for the electronic records will be selected to ensure the integrity of the electronic dossier during its storage at all Bureaus, Offices or Authorities for the full life of the patent information. The integrity of the storage specifically includes the continued readability of the information over the years, as technologies change and as media degrade. The storage mechanism shall be selected with the objective of preventing the need to re-record the information for the full life of the patent information.

3.6 Electronic Records Management

Technical standards are specific and precise, as the signals that the computers generate must be precisely those that the receiving apparatus is expecting to see. However, *legal standards* are more of an expression of a desired quality, which can be satisfied by a number of possible techniques. For example, any number of physical, electrical or encoding mechanisms can assure security of the record.

The various PCT Offices may satisfy the legal standards of this Annex in different ways. Variations in cost of these methods in different regions of the world may lead one Office to choose to protect, for example, the authenticity of the records under its control using PKI or Personal Identification Numbers, while another office may choose to protect those same qualities using far different techniques.

In order to allow the holders of patents derived from international applications to establish the validity of their intellectual property rights, the admissibility of the electronic records may have to be established in any court of any PCT member state.

All Offices that participate on electronic document exchange should comply with the requirements set out in Attachment 5 of Appendix I. In order to ensure compliance with these Electronic Records Management requirements, it could be foreseen that regular external audits of an Office's ERM implementation will be undertaken with the results being published by the IB.

3.6.1 Statement of the ERM Standards

Standards are presented below, each followed by Commentary. The comments are intended to be helpful, suggestive and explanatory, but only the standard is mandatory.

S1. All documents filed electronically must be capable of being printed as paper, and transferred to archival media, without loss of content or material alteration.

S2. Information that is routinely collected by the automated systems concerning the record, often called metadata, is to be considered part of the electronic records and maintained by the automated systems.

*Commentary:*A complete record as defined in archival science, and being more broadly accepted as a best practice consideration in electronic records management, has three primary elements:
(1) content, (2) structure, and (3) context.

- **Content** *Content* is the actual data resulting from a transaction conducted in the normal course of business, such as from a receipt or creation process. For example, the filing of a patent application includes the application form with various data fields and a signature.

- **Structure** *Structure* is generally defined in two parts: logical structure and physical structure. The logical structure of a record are the identifiable parts of the record, such as the title, applicant and inventor(s) name, date, and signature on a patent application. These parts may be both computer identifiable, as in metadata, and/or human identifiable, as when rendered on a viewing screen or printer.

  The physical structure relates to the format of the record, such as the type font, spacing, page margins, logo, and the encoding of the file, which provide information for processing (rendering) or transferring of the record over the full retention period.

- **Context** *Context* is the meaning of the record, or the "what" and "why" of the business transaction from which the record was created or received. The context may be implicit in the content and structure of the record, such as a patent or trademark application which contains a form number or description and a signature block which states the specific *intent* of the signer.

One of the key requirements for admissibility in evidence in a legal court is that the record and the system receiving or creating the record store an "accurate" representation of the record. A record is more likely to be perceived as accurate and complete when as many elements of the record as possible are recorded, either within the content of the record or as metadata. The more complete a record can be shown to be, the more weight will be attributed to it for admissibility and for any subsequent cross-examination. It is recommended that the full content, structure and context of electronic records be acquired and stored. The complete electronic record document should be accessible and be able to be rendered on display screen or printer without any loss or alteration of content or structure for the full retention period. If the electronic document was originally generated on or from paper, then the appearance on the paper should be maintained, at least in the archived copy of the received document. If the submitted document only existed as a string of text, then that is all that must be archived. If the received information is simply bibliographic data, then the context of that data must be preserved. For example, if the applicant has answered questions on an input form, either on paper or on a web page, then at least the wording of the questions must be captured and associated with the data elements entered by the applicant. Thus if we capture "1997" entered by the applicant, then at least the question to which she was responding, "What was the date of your invention?", must be captured. Even better would be the ability to recreate the screen into which the applicant placed her answer.

S3. Electronic documents must be submitted in an Office-designated electronic file format, and the archive copies must be retained in the electronic format in which they are submitted. Conversions to other formats may be made in working copies of the submittals.

S4. All electronic submissions must generate a positive acknowledgment to the submitter indicating that the Office has received the document. The positive acknowledgment must include the identity of the Office, date and time of the document's receipt (which is the Office's official receipt date/time), and an Office-assigned reference number or application number, if assigned.

*Commentary:*In addition to providing a document receipt to the submitter (which merely acknowledges the receipt of the submitted document), the Office may also wish to provide a document integrity check (e.g., document checksum) by which the filer may be assured that the submitted document was received correctly by the Office. Provision of a document validation is optional, but is recommended if enhanced electronic signature methods are being used, since document validation is a common feature of digital signature technologies. The Office needs to identify the document message digesting algorithms it will support.

S5. Every Office that accepts electronic filing must also accommodate the submission of paper documents. These paper documents may be imaged to facilitate the creation of a single electronic case file.

*Commentary:*While direct electronic submission is the preferred way to capture documents in electronic form, Offices will still need to accommodate paper submissions as a component of a comprehensive

electronic case files system. To facilitate the creation of a single electronic case file, it will be necessary to convert paper submissions to electronic form. A paper document can generally be imaged in a way that avoids loss of content or appearance. The approved technical standards are recited in Appendix I.

To make imaged documents text-searchable, the images may be converted to text via optical character recognition, or OCR. It should be noted, however, that OCR conversion of an imaged document to text would often introduce significant errors in the converted text. The OCR process may be acceptable as a means of creating searchable text for Office use, but not for creating official records for public access or for retaining archival records. The image (or the paper original) corresponding to OCR-created text must be retained for archival purposes.

While creation of a single electronic case file is encouraged, imaging of non-electronic submissions is not required. Offices may choose to maintain (some or all) paper submissions separately from electronic submissions.

S6. A mechanism must be provided to ensure the authenticity and integrity of the electronically filed document. This requires the ability to verify the identity of the submitter-- (the applicant or authorized representative)--, as well as the ability to verify that a document has not been altered without authorization since it was filed.

S7. Electronic filing systems must provide backup and recovery mechanisms to protect electronic filings against system failures.

S8. The electronic records must be maintained for long-term access and retention.

*Commentary:*The fundamental principle of electronic records management is to provide for long-term access and retention over the complete life cycle. Providing long-term access requires that every electronic case file or individual electronic record is searchable and retrievable and that it can be rendered (displayed) on a display screen or printer without loss of content or structure.

Retention requirements stipulate that the integrity of the electronic case file and associated records be preserved over the full life cycle, independent of changes in media technology or system obsolescence. As such, electronic records management solutions should not be predicated on a particular media or a single systems or application environment. The records should migrate to the newer technologies as they are installed, with no loss of content or structure.

S9. Electronic files must be scanned for viruses prior to processing.

S10. Access to computers used for electronic filing must not jeopardize the security of other Office networks and applications.

Commentary: The public must not be permitted access to internal Office networks or computers upon which Office operations are performed. One way to isolate Internet web sites that may be used for electronic filing is to use an Internet firewall. Additional network security methods can be combined with a firewall to further enhance network security. Similar security precautions should be taken for other electronic filing implementations.

S11. Electronic Records Management Systems must provide mechanisms for quality assurance and quality control of the submitted documents.

*Commentary:*The Office is responsible for ensuring the accuracy of its application file management data. How an Office chooses to ensure accuracy is a local Office management decision. Electronic filing systems should enable the Office to review submissions and validate the accuracy of the application file management data before accepting and docketing an electronic filing. These functions need to be supported by both the electronic filing system and the receiving Office's other automation systems.

S12. The electronic record management systems will maintain an audit trail of all additions to or alterations of the electronic records, recording the receipt information or other information about the generation of each record and of all changes to the records.

*Commentary:*The audit trail essentially shows the use history of the electronic case file or record. From a records management and legal perspective, an audit trail details the "chain of custody" whereby the who, what, and when of each action, or event related to an electronic case file or record is evident. Events could include the creation or receipt, processing, access, routing, dissemination, copy, reformat, transfer, or disposal of an electronic patent or trademark case file.
From a legal perspective, an audit trail can also be used to provide proof that, for admissibility as evidence in a court, the authenticity of a record has been maintained - - that the integrity of the content, structure and context has not been altered, misused or inappropriately destroyed.

S13. If access to confidential data by electronic means is allowed, this access must be secure and available to authorized viewers only. Measures to assure the protection of these files from alteration must be taken. Such access by applicants, representatives or authorized members of the public by electronic means must be documented concerning the identity of the party, the date (and optionally time) of the transaction, and the details of any submissions. Such documentation should be maintained as confidential data.

S14. To the extent provided in the PCT, adequate public access to the published international applications held by the Offices must be provided.

*Commentary:*The published case files and dockets of the Offices are public records where so provided by the PCT. Regardless of the electronic filing process that is adopted, adequate public access outside the Office is recommended. If an Office chooses to image its paper submissions and combine them with electronically filed documents to form a single electronic case file, then the public should have electronic access to all documents in the electronic case file, whether or not they were originally submitted in electronic form. This standard is not intended, however, to require extensive conversion of non-electronic filings if the Office chooses not to maintain a single electronic case file for all its filings.

4 Options for Receiving Offices

4.1 Background

As proposed to be revised, draft Section 701(c) of the Administrative Instructions under the PCT states: Each receiving Office shall notify the International Bureau of:

i) the electronic form or forms and the electronic means acceptable for filing for international applications which that receiving Office accepts;
ii) the requirements of the receiving Office for electronic filing and the electronic methods of communication accepted by the receiving Office;
iii) the conditions, rules and procedures relating to electronic receipt, including hours of operation, choices for acknowledgment of receipt and standard electronic acknowledgment processes, details concerning help desks, electronic and software requirements and other administrative matters related to the electronic filing of international applications and related documents;
iv) any change to the accepted electronic form or forms indicated under subparagraph (i)--the effective date of a change of electronic form accepted by a receiving Office shall be two months after the date of publication of the notification of the change in the *Gazette;*
v) recommended procedures which applicants should follow for alternatives to electronic filing when electronic systems of the receiving Office are not available by reason of malfunction or scheduled maintenance — in such cases of system unavailability, the receiving Office shall take all reasonable steps to notify applicants of such conditions.

4.2 Categories of Options (under draft Section 701(c)):

There are two categories of information concerning which the receiving Offices should notify the International Bureau:

(i) the choice of the receiving Office among its (limited) options concerning some matters; and

(ii) simple notification by the receiving Office concerning other informational matters.

Under Section 701(c), receiving Offices would be limited to the following options, from which they must choose:

1--"the electronic form or forms and the electronic means acceptable by the receiving Office for the filing of international applications:"

"Electronic forms" (i.e., the physical form of the medium containing the electronic data)

(a) electronic file (file formats apply as indicated in Annex F, Appendix I)
(b) on physical media [The necessary file formats are defined in Appendix III)
(i) 3.5 inch diskette, preferably produced using PCT-EASY (full application in electronic form)
(ii) CD-ROM or CD-R/DVD-ROM or DVD-R
(iii) hi-density diskettes (e.g., IOMEGA Zip disk, Super disk, etc.)]

"Electronic means" (i.e., manner of delivery or transmission)
(a) Electronic File Transfer (Mechanisms apply as indicated in Annex F, Appendix I)
[(b) Internet via browser (with or without encryption)
(c) email
(i) with body of application in email message
(ii) with body of application contained in email attachments (in accepted formats)]

2--requirements of the receiving Office for electronic filing and the electronic methods of communication accepted by the receiving Office;

(a) communications model used by the Receiving Office--
(i) notices/documents sent directly from the RO/Authority to the applicant at provided email address
[(ii) "mailbox" model, where applicant must log into system administered under the direction of the Authority to obtain messages, notifications]

3--the conditions, rules and procedures relating to electronic receipt, including hours of operation, choices for acknowledgment of receipt and standard electronic acknowledgment processes, details concerning help desks, electronic and software requirements and other administrative matters related to the electronic filing of international applications and related documents;

(a) In place of the default acknowledgment mechanism of the Confirmation Certificate, type of acknowledgment of receipt of electronically submitted applications that an RO may offer on request of an Applicant:--
(i) electronic (e.g., notice sent by receiving Office to supplied email address [or to internal electronic mailbox])
(ii) paper (mail, fax, etc.)
(b) hours of operation:
(i) open all the time for the purposes of electronic filings?
(ii) restricted hours?

4--[recommended] procedures [,if any,] which applicants should follow for alternatives to electronic filing when electronic systems of the Office are not available by reason of malfunction or scheduled maintenance—in such cases of system unavailability, the receiving Office [shall] [should] take all reasonable steps to notify applicants of such conditions.

(a) For reasonably predictable events, the description of the mechanisms in place to meet ERM Standard S7
  (i) RO Home Page on the Internet
  (ii) E-mail
  (iii) Fax

[Appendix I follows]

APPENDIX I

TRILATERAL TECHNICAL STANDARD FOR THE ON-LINE EXCHANGE OF IP DOCUMENTS IN A PKI ENVIRONMENT

1 Background

This document presents the technical requirements for on-line filing in a PKI environment. This standard is expected to evolve to cover all types of exchange of Intellectual Property Documents including the PCT procedure as well as non-patent procedures such as trademarks. Text in square brackets "[ ]" indicates requirements that are recommended to become part of the standard in future.

This standard represents a maximum set of measures that an IP Office can require of an applicant.

It has been adopted by the Trilateral Offices as the standard for the implementation of on-line filing pilots and for interoperability testing.

2 Scope

This specification allows all types of IP documents (Patents, Trademarks, Utility Models etc) to use the same mechanisms for electronic communication between applicants, Receiving Offices, Authorities and the IB for PCT as well as National Procedures.

The technical implementation for different types of exchange rely on the following aspects:

- Security and PKI
- Electronic Signatures
- Document Packaging
- Submission

3 Security

This section addresses requirements for:

1. the security of electronic documents that are transmitted over communication networks in the course of conducting electronic commerce with Intellectual Property (IP) Offices, and
2. the security of electronic documents held by IP offices subject to electronic record management practices. In particular, this covers the mechanisms needed to protect computer systems at IP Offices against unauthorised penetration.

After the requirements are identified, the section specifies technical standards to be used when implementing systems to fulfil the requirements.

Security requirements for the exchange and storage of IP electronic documents are derived from treaty and law protecting the content of IP electronic documents from inappropriate disclosure and the need to support the validity, admissibility and weight of IP electronic documents in legal proceedings. Consequently, IP office automation and electronic filing systems must reliably preserve the confidentiality and integrity of IP electronic documents while implementing features to ensure originator authentication and non-repudiation. Concise definitions of these qualities follow.

3.1 Public Key Infrastructure

Public Key cryptography techniques supply most of the accepted methods and defacto standards for providing these qualities in Internet electronic commerce. In public key cryptography key pairs are created. The key pairs have the property that each key of a pair can be used to decrypt data that was encrypted using the other key. Users of Public Key

cryptographic systems publish one key from the pair (the public key) and protect the other key from disclosure (the private key). Public Key cryptography is supplemented by the use of strong one-way message digest functions to establish the integrity of electronic documents. The two technologies are used to create digital signatures. To make a digital signature a message is input to a strong one-way message digest function and the result of the function, called a message digest, is encrypted using the private key.

Digital certificates have been developed as a means to bind the identity of parties with their public key. A digital certificate is a compact data message bearing the identity of its owner, the owner's pubic key and a means of independently verifying that the certificate can be trusted. Digital certificates may be issued either by one of the parties to the electronic business transaction after establishing the identity of the other party, or by a trusted third party. In the third party case, both parties to the electronic business transaction trust the policies and practices used by the third party to establish identities. For digital certificates to be useful it must be possible to revoke them when they are no longer valid. Situations leading to certificate cancellation include:

- a compromise of the private key paired with the public key on the certificate,
- the address or other identity information changes,
- a mistake is discovered in the certificate,
- an affiliation (such as employer) of the holder changes or
- the expiration date of a certificate is past.

To provide comprehensive security on a large scale, large numbers of digital certificates are required. An organization such as an IP office that desires to conduct bi-directional electronic commerce on the Internet with its customers protected by confidentiality, integrity, authentication and non-repudiation features needs to recognize a digital certificate for each customer. The policies and procedures plus the suite of systems and software required to issue, revoke, retrieve and manage large numbers of digital certificates is called a Public Key Infrastructure (PKI).

A PKI includes the following components:

Registration Authority (RA)	interacts with parties requesting a certificate to establish identities
Certification Authority (CA)	issues certificates based on applications received from the RA
Certificate Management	handles certificate renewal, revocation for expiration, revocation for cause, distributes Certificate Revocation List (CRL), validates certificates, validates digital signatures, provides API for use by software applications
Directory Service	maintains database of certificates and retrieves certificates for use
Certification Policies and Practices Statement	documents the operational practices and rules under which the PKI operates

3.2 Certificates

All digital certificates used in IP Document Exchange shall comply with the International Telecommunication Union (ITU) X.509 Version 3 Recommendation for certificate format. See: Recommendation X.509 (08/97) – Information technology – Open Systems Interconnection – The Directory: Authentication framework.

Requests for digital certificates shall be prepared in compliance with the PKCS#10 Standard.

See: RSA Laboratories, PKCS #10 – Certification Request Syntax Standard Ver. 1

Certificates may be issued as follows:

- On Smart Cards
- On Diskette
- On-line

Public Key Infrastructure (PKI) systems shall interoperate based on use of the X.509 Version 3 Certificates and X.509 Version 2 Certificate Revocation Lists . Implementations of PKI systems shall comply with the recommendations established by the Internet Engineering Task Force (IETF) Working Group on PKI Interoperability (PKIX) and documented in IETF RFC 2459. Draft standards documents are available from the University of Southern California Information Sciences Institute at ftp://ftp.isi.edu/internet-drafts/. See the following draft standards documents:

Internet Draft Filename	*Title of Specification*
draft-ietf-pkix-ipki3cmp	Internet X.509 Public Key Infrastructure Certificate Management Protocols
draft-ietf-pkix-crmf	Certificate Request Message Format
draft-ietf-pkix-ipki-part4	Internet X.509 Public Key Infrastructure Certificate Policy and Certificate Practices framework
draft-ietf-pkix-ipki-part1	Internet Public Key Infrastructure X.509 Certificate and CRL Profile
draft-ietf-pkix-ldapv2-schema	Internet X.509 Public Key Infrastructure LDAPv2 Schema

Implementations of PKI systems shall use separate key pairs and digital certificates for the purpose of authentication and confidentiality. The authentication keys shall be the property of the IP office customer, and the private key of the authentication keypair shall never leave the IP office customer's custody.

An IP Office may decide to offer Key Recovery for the security keypair when allowed under national laws.

3.3 Certification Authorities

Certification Authorities are responsible for maintaining the accuracy of the electronic certificates that "prove" a party is who he says he is. There can be many such authorities. Certificates will be issued as determined by each CA according to national law.

[The IB maintains a list of root Certification Authorities for the International IP community. The Offices will select from this list those root authorities that they will accept for certificate validation. The IB may also act as a CA.]

Each IP Office will subscribe to Certificate Revocation Lists for all CAs that it accepts. Whenever a certificate is used to authenticate an individual, these Certificate Revocation lists will be consulted by the IP OFFICE to ensure that the certificate has not been revoked.

3.4 Certification Management

The process of issuing, managing, and revoking certificates is divided into eight parts, referred to as the Certificate Life Cycle. The major life cycle processes are:

- Certificate Application
- Validation of Certificate Application

- Certificate Generation, Issuance, and Distribution
- Acceptance of Certificate by Subscriber
- Certificate Use
- Certificate Expiration and Renewal
- Certificate Revocation
- Key Recovery

Each of these processes has its own set of policies and procedures that will be followed, assuring that the PKI will provide a trusted environment. The first three phases are directed towards assuring that certificates are issued to appropriate individuals. The fourth and fifth phases refer to usage of the certificate by a certificate holder (called a subscriber). The last three phases address the end of the life cycle, where a certificate expires naturally, or a certificate may be revoked and replaced with a new one.

3.4.1 Certificate Application

This phase is the beginning of communication between subscriber and Certification Authority and thus initiates the certificate life cycle. IP Office personnel will not have to complete a certificate application. Employee status will provide the required evidence of identity and need for these certificates. Others will normally apply for a certificate by completing and submitting a certificate application that provides specific subscriber information, including name, organization, and certificate type. Written applications may be required initially; however, future enhancements to the PKI will implement on-line certificate applications. An external requester will be required to complete a subscriber agreement that sets out his obligations regarding the use of the certificate issued to him.

3.4.2 Validation of Certificate Application

The Registration Authority has responsibility for authenticating the identity of the certificate subscriber and affirming the accuracy of information submitted, including the need for the certificate. After validation of information in the certificate application, the Registration Authority authorizes the creation of certificates by the Certification Authority for the subscriber. The Certification Authority validates the authorization from the Registration Authority, to make sure that the authorization was issued by a valid Registration Authority and that it contains all of the required information. The Certification Authority then provides the subscriber with the information necessary to complete the certificate issuance process. The identity proofing function may be delegated to a Local Registration Authority (LRA) with organizational or customer focus. Certificate validation is closely tied to certificate application.

3.4.3 Certificate Generation, Issuance and Distribution

The subscriber uses PKI client software to complete a series of steps that results in the creation of key pairs, and the generation, issuance and distribution of public key certificates. Two key pairs (four keys) are created in the process: one encryption key pair and one signing key pair. The encryption key pair consists of the encryption public key and decryption private key. When created, the encryption public key is automatically sent to the Certification Authority platform where it is entered in a public key certificate and signed by the Certification Authority. The signing key pair consists of a signing private key and a verification public key. The verification public key is automatically sent to the Certification Authority where it is entered in a public key certificate and signed by the Certification Authority. If key recovery is implemented, a copy of the private decryption key is stored in a key recovery system for use if the decryption key becomes unavailable. The Certification Authority posts the encryption public key certificate to the appropriate Directory (certificate repository) and returns the verification public key certificate to the subscriber's PKI client software.

3.4.4 Acceptance of Certificate by Subscriber

The external subscriber can indicate acceptance of the certificate by various means such as by written agreement, or by use of the certificate to send a signed message to the Registration Authority acknowledging receipt of the certificate, or by use of the certificate to establish an encrypted session.

3.4.5 Use of Certificate

Subscribers encrypt objects (e.g., files, forms, documents, email) for intended recipients by using the recipient's public encryption key obtained from their encryption public key certificate; only the intended recipient is able to decrypt the object using his/her private decryption key.

Subscribers digitally sign objects using their private keys. Relying parties can verify the signatures of subscribers and the integrity of the signed object by obtaining the signer's public verification key from their verification public key certificate, which is provided with the signed object, and using it to verify the digital signature.

In both cases, the public key certificate and current certification revocation list are obtained by the relying party's PKI client software. The PKI client software then verifies the Certification Authority's signature on the certificate and from the Certificate Revocation List, verifies that the certificate has not been revoked, and in the case of digital signature verification, that the signature verification certificate was valid when the digital signature was executed. These activities are accomplished via simple icon or pop down menu choices executed by the user. This process will be automated.

3.4.6 Certificate Expiration and Renewal

Each certificate has a set life span after which it expires and needs to be renewed. The life span is set to avoid vulnerabilities that may occur if an attacker has a large collection of messages signed or encrypted with the same key and sets about breaking the key, a time intensive process. Normally, internal subscribers' certificates will be automatically renewed before they expire, with new key pairs generated and certificates issued. External subscribers may be required to request renewal. External subscriber software notifies the end user of pending expiration.

When key pairs are updated, they are replaced with new key pairs and new public key certificates are created. If a subscriber's certificate requires update for other than normal time expiration reasons, the subscriber and Registration Authority will need to be involved. Such reasons include the need to modify subscriber identification information, policy that requires periodic confirmation of subscriber information, or to resolve suspected misuse or key compromise.

3.4.7 Certificate Revocation

A subscriber's certificate may be revoked for any of several reasons. Certificate revocation may be initiated by the subscriber, the Registration Authority or Local Registration Authority, and/or authorized IP Office management. Subscribers should advise the cognizant Registration Authority or Local Registration Authority if they 1) no longer require use of the certificate (e.g., termination of employment, change of job responsibilities), 2) know of or suspect a compromise of their private key, or 3) have changed their name. In the absence of a request by the subscriber, the cognizant Registration Authority or Local Registration Authority should request revocation of a subscriber's certificate for any of the above reasons. The cognizant Registration Authority or Local Registration Authority should also initiate revocation of a subscriber's certificate if there is a material breach of the subscriber agreement.

3.4.8 Key Recovery

A subscriber should be able to recover data, which they have encrypted or that was encrypted for them, even though their decryption private key becomes unavailable. The key may become unavailable for a variety of reasons including, inability to access the stored key (e.g., forgets password), corruption of the stored key, failure of the storage medium, and theft of the key or storage medium. An organization should be able to recover its data, which has been encrypted by subscribers, when the subscriber is unable or unwilling (e.g., disgruntled, incapacitated, unavailable) to decrypt the data.

The IP Office PKI may provide the capability for key recovery of internal and external subscriber decryption keys. In order to meet these requirements, a copy of each user's private decryption keys must be obtained and securely stored to enable the authorized recovery of encrypted data.

Key recovery does not apply to the subscriber's signing keys. The subscriber's private signing keys are not recoverable due to the requirement for effective non-repudiation. Non-repudiation is supported by having the subscriber generate his signing key pair on his own system and only transferring his public verification key to the Certification Authority during the registration process. The private signing key must remain under the sole control of the subscriber so that there is no opportunity to mascarade.

The following discussion applies to decryption key recovery only. It is a highly sensitive PKI function since it deals with the confidentiality of communications and files which may, as with patent application prosecution, be held in confidence by law.

Key recovery for external subscribers may only be initiated by the subscriber, a Registration Authority, or a Local Registration Authority by following established key recovery procedures and interacting with the Registration Authority.

For internal subscribers, a Registration Authority or Local Registration Authority should initiate key recovery only after authorization by appropriate IP office management. Such authorization may result from a request from the internal subscriber or from a requirement by management to access data encrypted by the subscriber.

3.5 Cross Certification

The Trilateral Offices have indicated that they will maintain their own Certificate Authorities. As a result, there are likely to be several root certificates in IP Document Exchange. Cross certification provides a method to greatly reduce the need to distribution root certificates throughout the system and a resulting simplification of root certificate management.

When two communities of users each have a Certification Authority, a process called cross certification can be used so that the users of one community can trust the public key certificates of the users in the other community, and visa versa. This is accomplished by having the two Certificate Authorities issue certificates for each other's public key. User PKI client software can process these certificates to determine whether the public key certificate of a user in another community should be trusted. It is important that a determination is made that the certificate policies of each community of users are at equivalent levels of assurance. The questions and issues arising related to accepting policies between offices will be more challenging than the technical issues of providing cross certificates.

For the present, until a model Certificate Policy can be accepted by the various IP Offices, cross certification will be accomplished on a case-by-case basis, based on a successful review of the parties respective Certificate Policies. In addition, some third party certification authority, as yet

unidentified, can be requested by the Office to certify compliance with the published described procedures, which certification would also be published.

3.6 Digital Signatures

Digital signatures used to sign electronic documents for IP Document Exchange shall conform to the format and practice specified in RSA Laboratories, PKCS #7 – Cryptographic Message Syntax Standard Version 1.5 definition of Signed-data content type. To build these signatures, a certificate is needed. These are X.509 version 3 certificates signed by an approved Certification Authority (CA).

3.7 Directory Services

All Offices participating in electronic document exchange must subscribe to the CRLs described above and must use these when decoding a PKCS#7 package.

PKI Systems shall store certificate information in a directory structure complying with ITU Recommendation X.500. Such systems shall provide an external interface for publishing and retrieving user digital certificates that complies with the Lightweight Directory Access Protocol (LDAP). See ITEF Network Working Group RFC 1777 dated March 1995.

3.8 Encryption Algorithms

Both symmetric (secret key) and asymmetric (public key) algorithms may be used as necessary. Algorithms that are prohibited under national law of a country shall not be used for IP Document Exchange from that country. Algorithms implemented in hardware or software shall not be used in any manner that is contrary to export restrictions of the country of origin for the hardware or software. Any algorithm used between IP Offices must be disclosed to both parties.

3.9 Data Encryption

Electronic document data that is encrypted to ensure confidentiality for IP Document Exchange shall conform to the format and practice specified in RSA Laboratories, PKCS #7 – Cryptographic Message Syntax Standard Version 1.5 definition of Signed and Enveloped-data content type.

3.10 Strong One-Way Message digest Algorithms

The message stream shall be input to a strong one-way message digest algorithm to create a message digest. The one-way message digest algorithm shall be SHA-1.

3.11 Security and Payment Mechanisms

The security offered under the PKI system for data confidentiality shall also be deemed sufficient to protect the confidentiality of Credit Card information transmitted online for fees or other payments.

3.12 Summary of Security mechanisms

The following table shows how, in a PKI environment, various technical components meet the Confidentiality, Integrity, Authenticity and Non-repudiation:

	Confidentiality	Integrity	Authenticity	Non-repudiation
Certification Authority			X	X
Digital Certificates		X	X	X
Certification Management		X	X	X
Cross Certification		X	X	X
PKCS#7 Signed Data Type		X	X	X
Directory Services			X	X
Encryption Algorithms	X			
PKCS#7 Signed and Enveloped Data Type	X			
Message digest Algorithm		X		
PKI Policy Statements		X	X	X

4 Signatures Mechanisms

A signature functions in the electronic world to identify a particular person as a source of the electronic message. It also indicates such a person's approval of the information contained in the electronic message.

For the purposes of this document, there are two signature mechanisms:

- Basic Electronic Signature
- Enhanced Electronic Signature

This signature, which includes the full name of the signature holder as well as place and date of signature, is embedded in the document as XML tagged data (See Appendix II for the XML DTD). The XML tagged data either indicates that the user wishes to apply their Enhanced Electronic Signature to the electronic message or their Basic Electronic Signature.

4.1 Basic Electronic Signature

To indicate the human intention to perform a certain action, the standard includes the specification of a basic electronic signature. This is can be one of the following types of signature:

- A particular string of text entered by a user
- A facsimile image of the handwritten signature

The Basic Electronic Signature is encoded within the "party" structure of the XML document shown below:

```
...
<!ELEMENT electronic-signature
             (date-signed,
              place-signed,
              ((signature-mark,signature-file?)
             | (signature-file,signature-mark?)
             | use-digital-signature)) >
...
```

A Basic Electronic Signature within an XML document may be supplemented by the addition of a Digital Signature of the Signer's Representative to the Wrapped Documents.

4.2 Enhanced Electronic Signature

This is a PKCS#7 Signed Data Type generated from the electronic message by the act of the signer invoking the use of their private authentication key to encrypt the message digest (Digital Signature). The PKCS#7 Signed Data Type includes a copy of the Digital Certificate of the signer issued by a recognised Certification Authority.

5 Document Packaging

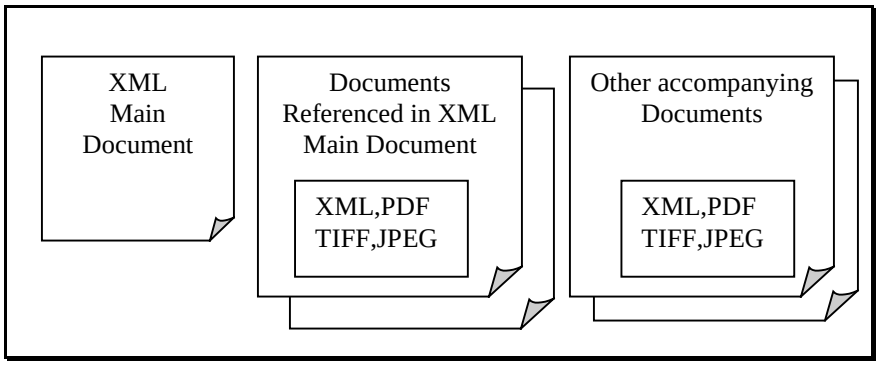
The document packaging mechanism is used to combine into a single binary object both the data about what is being transmitted with the contents of the transmission and to then apply the appropriate digital signatures and encryption.

5.1 Document Preparation

For each IP Document Exchange there is an XML Main Document that may explicitly reference and be hyper-linked to other documents. These referenced documents are logically part of the Main Document (e.g. a New Patent Application). In addition, a document exchange may include other accompanying documents (e.g. Designation of Inventor or a Fee Payment).

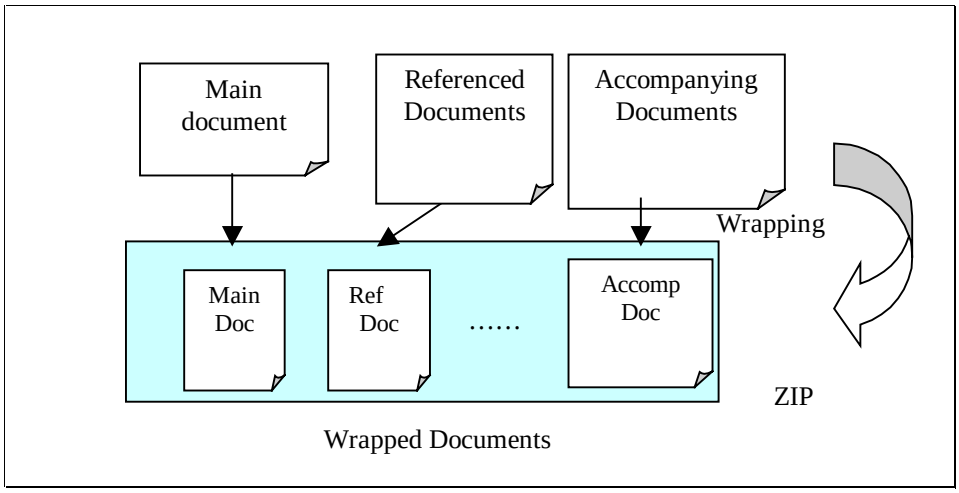
The XML Main Document must conform to one of the DTDs specified in Appendix II. The Referenced Documents (external entities) are typically embedded images, tables, drawings or other compound documents and may be encoded as either XML, PDF, TIFF and JPEG. See Attachment 1 for details.

The accompanying documents are separate, but related documents that may be encoded as either XML, PDF or Image. See Attachment 1 for details.



5.2 Wrapping the Documents

The Main Document with any Externally Referenced Documents and Accompanying Documents are wrapped and treated as one data block. This data block is called the Wrapped Documents and is created using the wrapping standard (ZIP). Applicants shall use ZIP format archiving and compression software to package the document files constituting an electronic application. The software used to create the ZIP file shall conform to the ZIP format standard as published in the PKWARE® PKZIP® Application Note. National Offices, WIPO and Third party vendors and implementers of filing software shall verify that any ZIP software used complies with the Application Note standard. Attachment 2 describes the ZIP standard in detail.

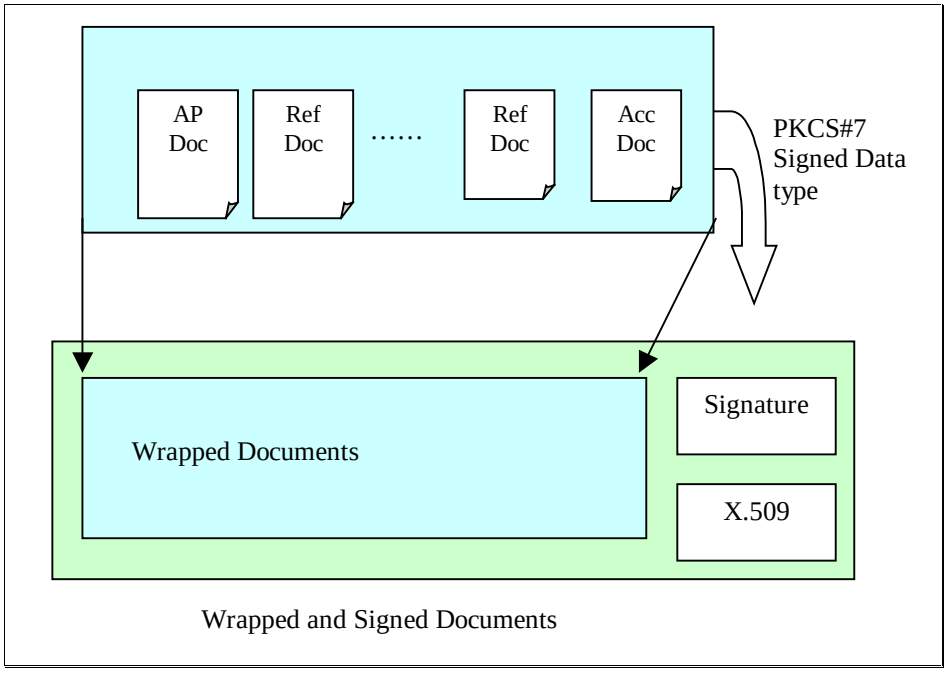


See Attachment 1 for details of the allowed document format types.

5.3 Signing the Wrapped Documents

To bind the person submitting the package to the electronic Wrapped Documents , a Digital Signature is added to create the Wrapped and Signed Document Data item. The purpose of adding the signature is to identify the applicant and to ensure that the recipient is able to detect any unauthorized alternation during the transmission.

PKCS#7 is used to produce a Signed Data Type for the signature. Detailed information on PKCS#7 is provided in Attachment 3.



5.4 Packaging the Wrapped and Signed Documents

A package is the actual transmission data that is exchanged between the applicant and RO.

The package contains various data items according to the individual package type. Data items include:

- Header Object Data item
- Document Data item that is made by wrapping and signing Documents
- Transmission Data such as the Ticket.

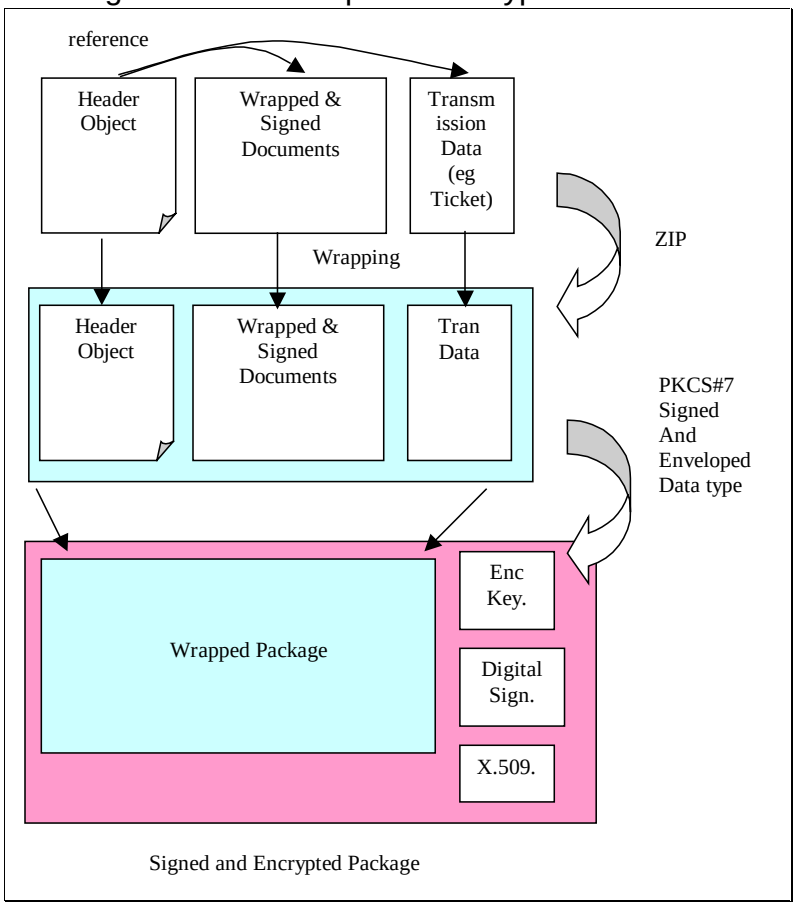
The Header Object Data item indicates the package type, file name of data item, etc. The Header Object Data item is always found in the package. The Header Object Data item is written in XML in accordance with DTD defined in Appendix II.

A package for network transfer is created by making one data block from the multiple data items. The procedure for creating the package is as follows:

- Create a wrapped package by wrapping multiple data items using ZIP
- Create a signed and encrypted package for network transmission by encrypting using the Signed And Enveloped Data Type in PKCS#7

The purpose of the signature is to assure the combination and contents of individual data items, and to ensure that the recipient is able to detect any unauthorized alterations during the transmission. Encryption is to prevent unauthorized interception during data communications.

The Digital Signature for the Wrapper Application Documents may be produced either by the applicant or their representatives. The person that starts the transmission produces the Digital Signature for the final Signed and Enveloped Data Type.



The IP Office then receives the package, opens the data items in the package and decides the role of individual data items in accordance with the documentation in the Header Object Data items.

6 Submission

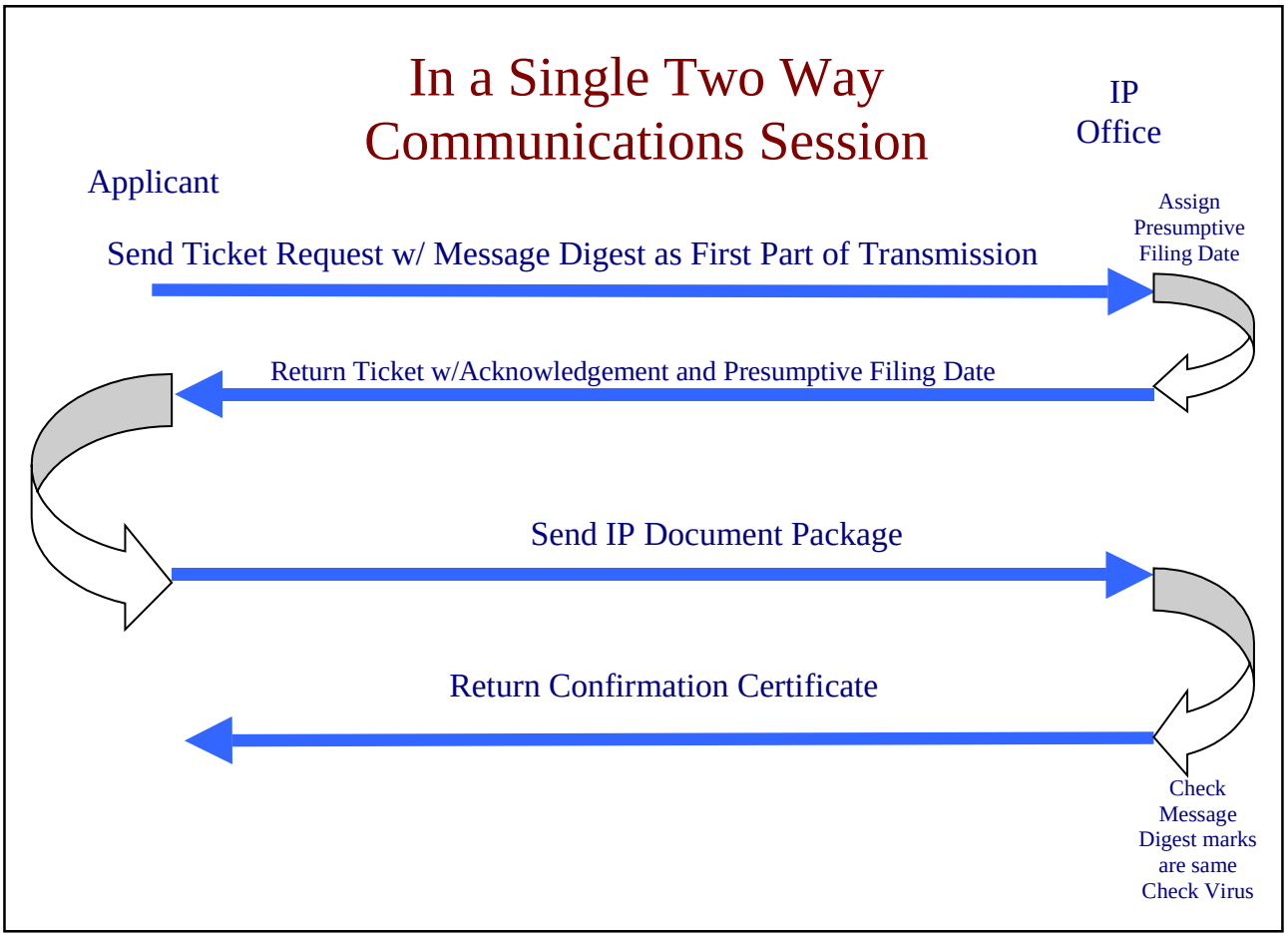
As shown above, the information exchanged during a transaction is broken into packages. Each phase of the exchange corresponds to the transfer of a package of data sent between the Application and the IP Office.

For all transactions, there will be the following four packages :

- Ticket Request
- Ticket
- IP Document Package
- Confirmation Certificate.

For each type of electronic data exchange, the IP Document package will contain the data actually prepared by the applicant (e.g. New Application, Fee Payment, Replacement Claims etc.)

The "**Ticket Mechanism"** operates as follows. :



- An electronic session is established between the applicant and the IP Office.
- Near the beginning of that session, a message digest is transmitted to the Office which is uniquely derived from the wrapped files. The code is such that any change to any of those files will be indicated by a change in that message digest.
- On receipt of that message digest, and as part of the session, the Office sends an acknowledgement to the applicant indicating the date of receipt of the message digest, which will apply to the full documents to come.
- The applicant then continues the session and transmits the complete set of files constituting the IP Document package.
- On receipt of the full set of files, the documents are then checked for the presence of viruses and processed to develop their unique message digest. This is compared to the original message digest that was sent at the beginning of the session. If they match, an acknowledgement of receipt is sent to the applicant. If they do not match, the applicant is informed accordingly. The session can then be ended.

Attachment 4 gives complete details of the communication flow for ticketed communications.

In the event of a communications or message digest comparison problems, the Confirmation Certificate contains information about the problem detected.

6.1 Transfer Protocol

To increase the probability of successful validation, a reliable transport layer protocol shall be used for all transfers between offices. The reliable transport layer protocol serves to assure that all data in transmission has been transferred and re-assembled correctly between the sending and receiving software applications. For this standard, the FTP or the HTTP protocol is used to transfer the package.

The security of communication provided by the encryption within a PKCS#7 Signed And Enveloped Data type will normally be sufficient but, if an IP Office considers it necessary, it may opt for channel level encryption such as SSL or IP Sec to enhance this security.

7 Types of Document Exchange

This Document Exchange Standard includes the following procedures:

National Patent Procedures

- On-line Filing of new patent applications
- [Procedural communications between the Applicant and IP Offices]

PCT Procedures

- On-line Filing of new PCT applications
- [Procedural communications between the Applicant and RO/IB
- Receiving Office to IB (Record Copy)
- RO to ISA (Search Copy)
- IB to ISA
- IB to IPEA]

Non-Patent Procedures

- [Trademark Applications
- Procedural communications between the Applicant and IP Offices]

8 Reference Implementations

As part of the preparation of this standard, the Trilateral offices have prepared two reference implementations (both in JAVA and C++ running on Win NT) that allows other developers to re-use and extend the basic source code provided to build client and procedure specific implementations.

The reference implementations covers the following areas:

- ZIP
- PKCS#7
- Packaging
- Ticket based transfer including return of a Confirmation Certificate.

These are available in source as well as object code.

In addition, standard test data sets are available to verify third-party implementations.

Attachments

Attachment 1. Document Format Requirements

The Trilateral Offices and WIPO are committed to the principle of establishing an open standards environment for electronic exchange of Intellectual Property documents. A notable result of this is: the standard for submitting electronic documents emphasize the use of open standards and will not promote proprietary vendor formats for electronic documents. The reasons for this policy include avoiding the need to maintain the record copies of electronic filings in specific versions of proprietary formats over which the offices have no control.

One desirable feature of commonly used proprietary word processor systems is that they package their electronic documents as a single file such as a .doc or .wps. The proprietary .doc, .wps and other word processor formats combine text, processing instructions, page layout information, raster graphics, vector drawings, tables and other types of data in a single proprietary word processor file.

The Trilateral Offices have selected an open systems alternative to word processor files where electronic documents will be based on using the eXtensible Markup Language (XML) which is being developed to deliver structured data on the World Wide Web. XML documents, like HTML web pages, consist of a character coded text file and zero or more additional files that may be more text or contain binary data such as images and drawings. A typical XML patent application electronic document will consist of a collection of files. An example would be a text file for each procedural document submitted as part of the application plus a text file for the specification of the invention that is accompanied by multiple graphics files (one graphics file for each drawing in the specification). While the XML approach has freed the offices from investing in proprietary word processor formats, the simplicity of the single word processor document file has been lost.

1.1 Images

The facsimile images for use in IP document exchange must meet the following requirements:

- Format
  - TIFF V6.0 with Group 4 compression, Single Strip, Intel Encoded or
  - JPEG
- 200, 300 or 400 dpi
- Max size A4 or Letter size

1.2 PDF

The PDF documents for use in IP document exchange must meet the following requirements:

- Acrobat V3 compatible
- Non-compressed text to facilitate searching
- Un-encrypted text
- No Digital Signatures
- No embedded OLE objects
- All Fonts must either be embedded, Standard PS17 or built from Adobe MM fonts

1.3 XML

All XML documents must conform to one of the DTDs specified in Appendix II.

The character set for all XML documents must be either UTF-8 encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP encoded JIS-X0208. [For PCT Applications, Chinese GB2312 and Korean KSC 5601 are also acceptable)

Attachment 2. Wrapping Specification (SDIF V2)

The ZIP format published in the PKWARE® PKZIP® Application Note and by Info-ZIP is suitable for use in IP document exchange. Commercial off-the-shelf software libraries and applications for creating ZIP format files are available from several vendors. Use of the ZIP format will provide the benefit of archiving and compression.

2.1 Wrapping of Application Documents

An easy to use, open standards approach is needed to wrap or pack a multi-file electronic document into a single file object for delivery from the applicant to the patent office. Having the applicant create the single file greatly simplifies the handling of the document by the IP Office as it need not track the successful transmission/reception of the individual files. A single file also means that a digital signature can be computed for the application file which can then be used to ensure the data integrity of the entire application.

2.2 Archiving and Compression

The creation of archive files is an approach that has been adopted by the PC, UNIX and MacIntosh environments. An archive is a collection of computer files that have been packaged together for backup, to transport to some other location, for saving away from the computer so that more hard disk storage can be made available, or for long term storage. An archive can include a simple list of files or files organized under a directory or catalog structure (depending on how a particular program supports archiving).

Compression is the reduction in size of data in order to save space or transmission time. For data transmission, compression can be performed on just the data content or on the entire transmission unit (including header data) depending on a number of factors.

Content compression can be as simple as removing all extra space characters, inserting a single repeat character to indicate a string of repeated characters, and substituting smaller bit strings for frequently occurring characters. This kind of compression can reduce a text file to 50% of its original size. Compression is performed by a program that uses a formula or algorithm to determine how to compress or decompress data.

The above definitions of archiving and compression include features that are desired for the electronic filing of multi-file patent documents. A suitable archive technique will produce a single file that includes all the component files of an electronic application plus a master directory of the files with information on their type, size, date and time they were last changed and a CRC code for error detection. Data compression of the application content will reduce the amount of time required for online submission and reduce the likelihood of experiencing a transmission error.

2.3 Use of ZIP Files

The ZIP format is a widely used open standard that provides both archiving and compression of data files. The archiving features of ZIP allow the user to collect all the files in a single ZIP directory. All the files in the zip directory along with the directory information are compressed into a single ZIP file object which is suitable for input to a digital signature process. The compression algorithms in the ZIP standard are lossless so the user can be assured that the decompressed result (unZipped) file is identical to the original. The compression techniques used by the ZIP standard achieve the greatest reduction of size (greater than 50%) for text files, but reductions on the order of 10 to 20% for compressed image files are achievable. The error detection features of the ZIP format (which are based on using a 32 bit CRC code) add additional assurance of data integrity.

2.4 ZIP Usage

The files to be zipped shall include all parts of the document identified elsewhere in this specification. All external files referenced by the Specification of the Invention must be included in the ZIP file submission. Filenames included in the central directory of the ZIP file shall comply with the specification for the applicable operating system given elsewhere in this specification.

2.4.1 Directory Structure

All ZIP files must have a flat directory structure. If a collection of files need to be embedded in the ZIP file, then these should be included an single flat embedded ZIP file.

2.4.2 Compression Algorithms

The ZIP standard allows the compression software to select from among a number of compression algorithms. The default compression method shall be “Deflation” with the normal compression option. This format can be most readily dealt with by UNZIP packages. The “Shrinking” compression method shall not be used because it makes use of a patented Lempel-Ziv-Welch (LZW) compression algorithm protected by a patent held by the UNISYS Corporation.

2.5 PKZIP® Application Note

The ZIP format was originally developed by Phil Katz and incorporated in PKZIP® software for DOS which is available as shareware. PKWARE®, Phil Katz’s company sells commercial versions of ZIP software for many platforms. Phil Katz published the standard for the ZIP format, making it an industry open standard. C programming language source code for the ZIP and UNZIP functions was originally published by a group of independent software developers and appears on the Info-ZIP website http://www.cdrom.com/pub/infozip/. Info-ZIP also publishes the ZIP standard on its website. While the ZIP standard is not a formal international standard, the information available on the ZIP format has allowed several third party vendors to develop products that implement ZIP functions on a wide variety of computer platforms. There is good interoperability achieved among the products of these vendors. Most of the third product vendors offer packages that are designed to interoperate with PKWARE® software. A sound strategy is to require applicants to prepare files for submission with any software that claims to be compatible with PKZIP and PKUNZIP. There are currently products available that meet this requirement from WINZip®, DynaZip, NetZip, Info-Zip and others.

The PKZIP® standard can be found as an application note (Revised: 08/01/1998) on the PKWARE® web page http://www.pkware.com/appnote.html.

Attachment 3. PKCS#7 Envelope Formats

This document describes the basic specifications of the digital envelope for IP Document Exchange. Following are the preconditions of this specification.

3.1 Scope

In this document, the structure of the digital envelope at the data transfer layer and business data processing layer is defined. The required functions of the digital envelope are given below:

- Giving a digital signature to user data (for authentication or detection of illegal data alteration)
- Encryption of user data

The following matters are NOT described in this Attachment:

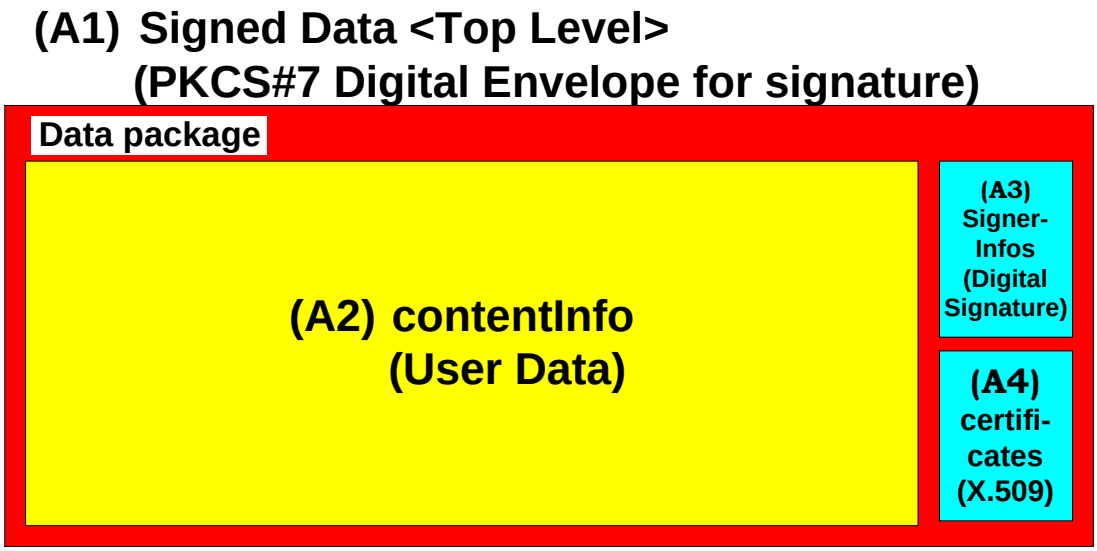
- The structure of user data that will be encrypted or to which a digital signature will be given. For example, method of packaging or document format structure etc.
- The method of describing additional information used for data processing at each national patent office. These data elements are not essential information for the applicant and RO.

3.2 Definitions

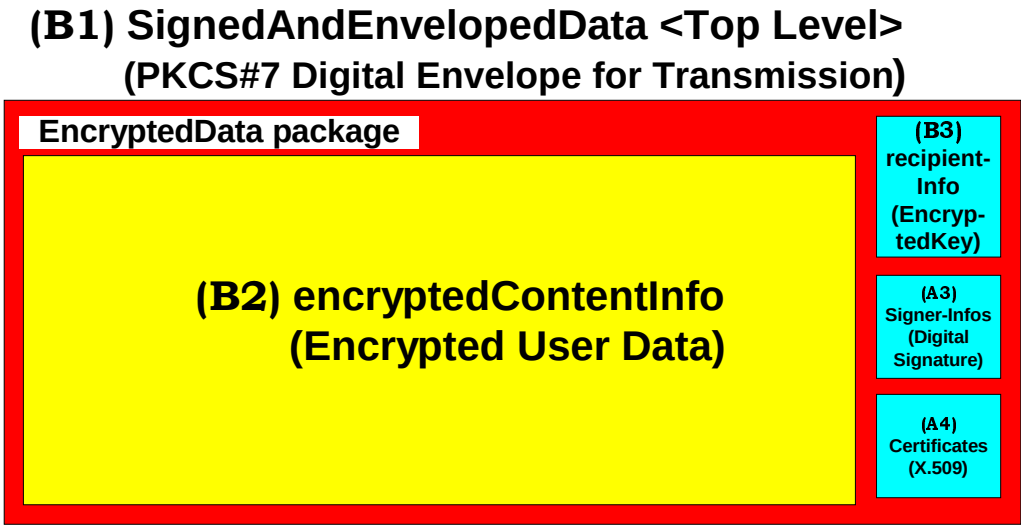
PKCS#7	The PKCS#7 Cryptographic Message Syntax specification, as defined in ***Internet Draft <draft-hoffman-pkcs-crypt-msg-03.txt> Version 1.5***
X.509	X.509 digital certificate standard, as defined in ITU-T Recommendation X.509 (06/97).
Object identifier for sha-1	The object identifier for sha-1 that we adopt is defined in OIW interconnection protocols: Part 12. The definition is below: **Sha-1 OBJECT IDENTIFIER ::= {iso (1) identified-organization(3) oiw(14) secsig(3) algorithm(2) 26}**
Object identifier for RSA encryption	The object identifier for RSA encryption is defined in *RSA Encryption Standard PKCS#1.* The definition is below: **Pkcs-1 OBJECT IDENTIFIER ::= iso(1) member-body(2) US(840) rsadsi(113549) pkcs(1) 1}** **RsaEncryption OBJECT IDENTIFIER ::={pkcs-1 1}**

Digital envelope for certification

This envelope is used for detecting alterations to user data. The user data, digital signature, and digital certificate of the signer are stored in this envelope.



Digital envelope for transmission
This envelope is used for data transmission, to achieve both encryption and alteration detection at the same time. In the case of this envelope, the digital signature is used, not for certification, but for detection of data alteration on the network, so this digital signature does not form part of the business data.



3.3 Common rules for digital envelopes for IP Document Exchange
All digital envelope data should be encoded under DER rules. The DER encoding rule helps the application program to analyze the digital envelope data following only one unique rule.

3.3.1 Digital envelope for the signature
This digital envelope is SignedData type PKCS#7.
Rules for producing the PKCS#7 digital envelope for certification

Table A1 SignedData top level

No.	Item name	PKCS#7 item	Content
1	Version	Version	Set integer value '1'
2	Set of algorithm identifiers	DigestAlgorithms	
2.1	Algorithm identifier	AlgorithmIdentifier	Set ONLY ONE **set of** algorithm identifiers {sha-1}[1]
3	Content information	ContentInfo	Set one content info (see table A2)
4	Certificates	Certificates	Set one Certificates (see table A4)
5	Certificate revocation lists	Crls	Not used (Set no data)
6	Signer information	SignerInfos	Set one signerInfos (see table A3)

Table A2 contentInfo top level

No.	Item name	PKCS#7 item	Content
1	Content type	ContentType	Set object identifier {pkcs-7 1}
2	Content	Content	Set user data (binary)

Table A3 signerInfos top level

No.	Item name	PKCS#7 item	Content
1	Version	Version	Set integer value '1'
2	Issuer and serial number	IssuerAndSerialNumber	Issuer of certificate and its serial number defined in X.509 spec. (for signer's certificate)
3	Set of digest algorithms	DigestAlgorithm	
3.1	Algorithm identifier	AlgorithmIdentifier	Set ONLY ONE **set of** algorithm identifiers {sha-1} for making digest of digital signature.
4	Authenticated attributes	AuthenticatedAttributes	Not used (Set no data)
5	Digest encryption algorithm	DigestEncryptionAlgoritm	Set object identifier {pkcs-1 1} (rsaEncryption[2])
6	Encrypted digest	EncryptedDigest	Message digested data; content is encrypted with signer's private key.
7	Unauthenticated attributes	UnauthenticatedAttributes	Not used (Set no data)

[1] sha-1 OBJECT IDENTIFIER ::= {iso(1) identified-organization(3) oiw(14) secsig(3) algorithm(2) 26}

[2] rsaEncryption OBJECT IDENTIFIER ::={pkcs-1 1}

Table A4 certificates top level

No.	Item name	PKCS#7 item	Content
1	Set of certificates	ExtendedCertificatesAndCertificates	
1.1	The X.509 certificate	Certificate (defined in X.509 spec.)	Set ONLY ONE **set of** X.509 certificate data

4.3.2 Details of digital envelope for transmission

This digital envelope is SignedAndEnvelopedData type PKCS#7.

Rules for producing the PKCS#7 digital envelope for transmission

Table B1 SignedAndEnvelopedData top level

No.	Item name	PKCS#7 item	Content
1	Version	Version	Set integer value '1'
2	Recipient information	RecipientInfos	Set ONLY ONE **set of** recipientInfo (see table B3)
2	Set of algorithm identifiers	DigestAlgorithms	
2.1	Algorithm identifier	AlgorithmIdentifier	Set ONLY ONE **set of** algorithm identifiers {sha-1}
3	Encrypted Content information	EncryptedContentInfo	Set one encrypted content info (see table B2)
4	Certificates	Certificates	Set one Certificates (see table A4)
5	Certificate revocation lists	Crls	Not used (Set no data)
6	Signer information	SignerInfos	Set one signerInfos (see table A3)

Table B2 EncryptedContentInfo top level

No.	Item name	PKCS#7 item	Content
1	Content type	ContentType	Set object identifier {pkcs-7 1}
2	Content encryption algorithm	ContentEncryptionAlgorithm	Algorithm OBJECT identifier of content encryption. (JPO's tested system: DES in CBC)
3	Encrypted content	EncryptedContent	Encrypted user data

Table B3 recipientInfo top level

No.	Item name	PKCS#7 item	Content
1	Version	Version	Set integer value ‘1’
2	Issuer and serial number	IssuerAndSerialNumber	Issuer and serial number of certificates that includes the public key for encrypting user data encryption key.
3	Key encryption algorithm	KeyEncryptionAlgorithm	Algorithm OBJECT identifier for encrypting user data encryption key. (JPO’s tested system: RSA1024)
4	Encrypted key	EncryptedKey	Encrypted decryption key for user data.

Attachment 4. Ticket Mechanism

This Attachment describes the Data format for each application phase of the Ticket Mechanism.

Ticket is exchanged between the applicant and IP Office with the following protocol.

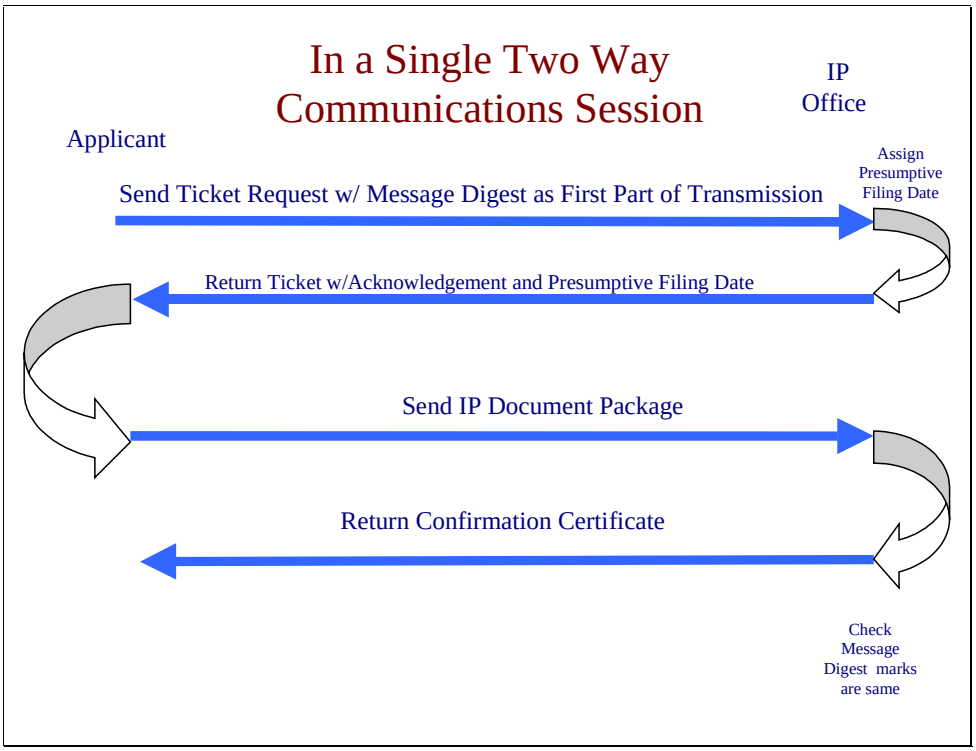


Fig1.Ticket Protocol

Individual protocols in the Ticket system are described below.

4.1 Ticket Request

The Ticket Request is the first packet that is sent to RO. The purpose of sending this packet is to prevent disadvantageous events in recovery measures and transmission speed if the line fails during the submission of an application made with a large Application Documents Data item.

A Ticket Request Data item includes:

- A Message digest created by a Message digest algorithm after wrapping the Documents using ZIP,
- Bibliographic Data and
- A Header Object specifying that the corresponding packet is a Ticket Request.

The Ticket Request is created by wrapping these three Data items with ZIP and packing the resulting ZIP file into a Signed And Enveloped Data type in PKCS#7. This package shall be verified by the client prior to submission. If any errors are detected, this shall be reported to the user and the submission cancelled.

Detailed information on the Header Object Data item is provided in Appendix II. Information on PKCS#7 is provided in Attachment 3.

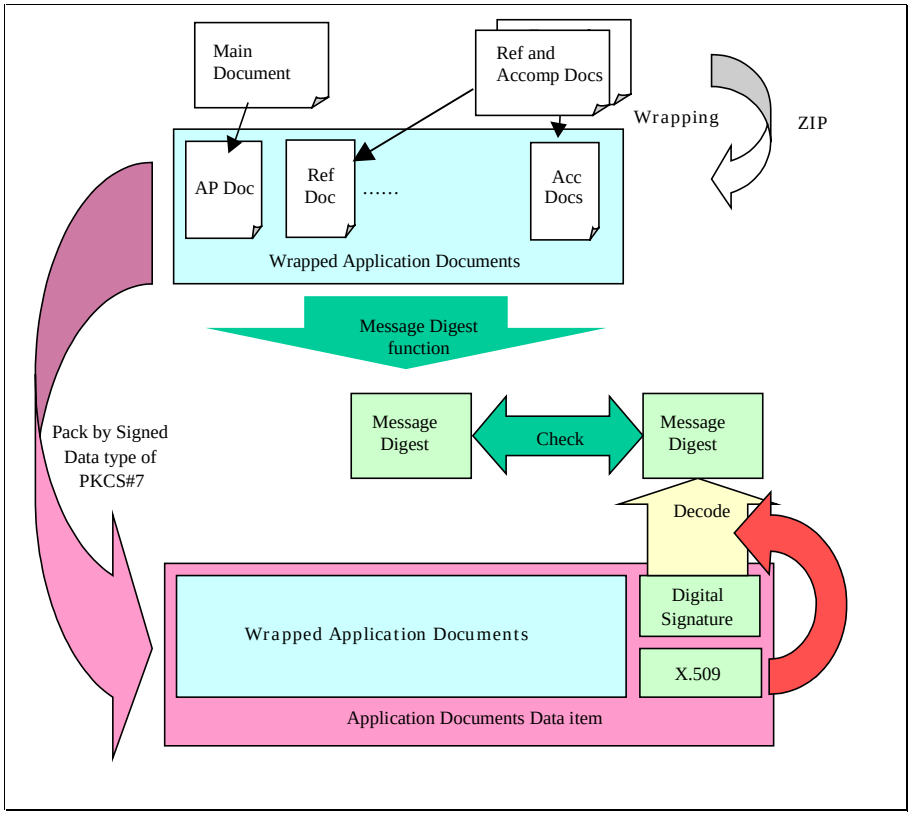


Fig2. Outline of Message digest

Ticket Request

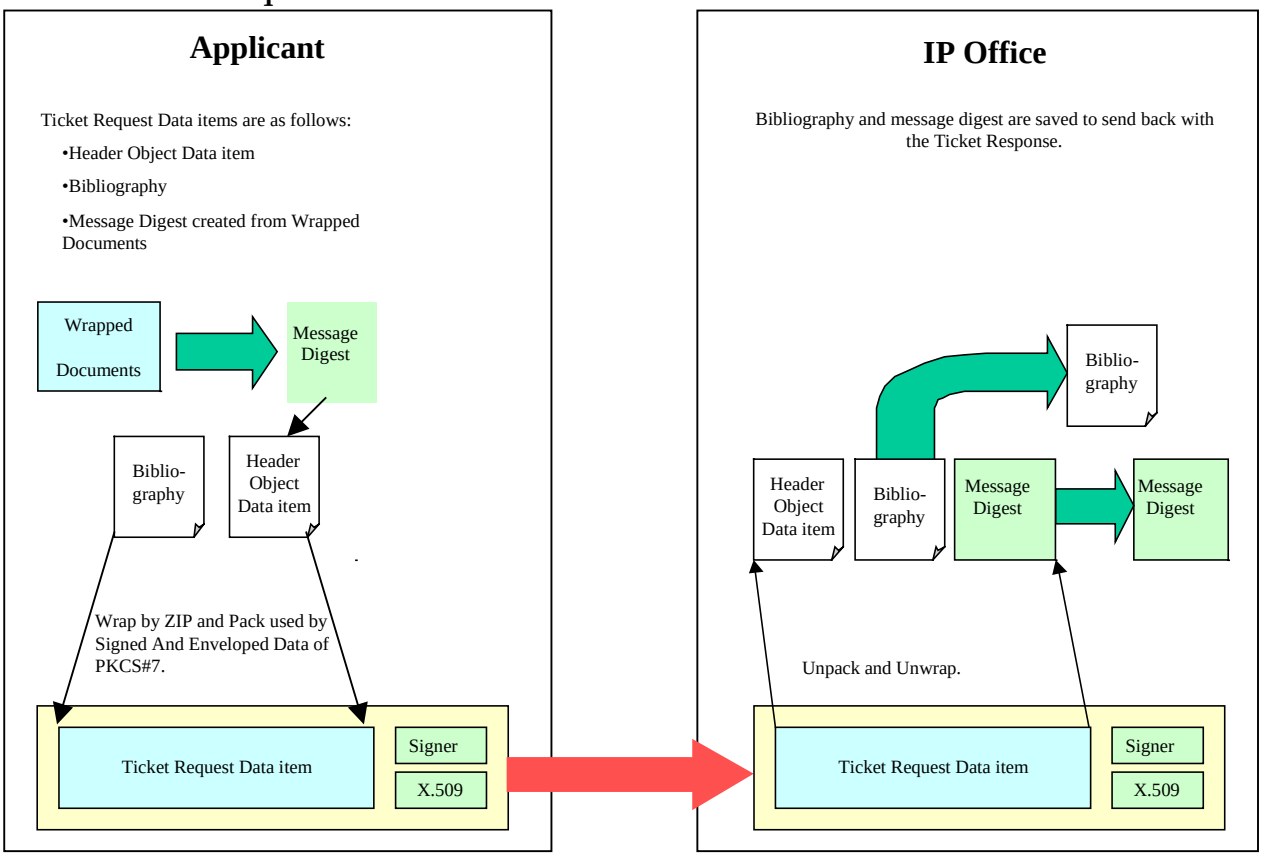


Fig3.Ticket Request

4.2 Ticket Response

The Ticket Response Data item includes a Header Object Data item, a Ticket Data item, and Bibliographic Data.

- The Header Object Data item specifies that the corresponding packet is a Ticket Response.
- The Ticket Data item is created as Signed Data type in PKCS#7 by wrapping the Ticket that includes the Ticket Request Receiving Number, Date Stamp and Expiration Date, and Message digest received from the Ticket Request by ZIP.
- The Bibliographic Data is attached to specify the Ticket Response to the Ticket Request.

The Ticket Response is created by packing with Signed And Enveloped Data type in PKCS#7 after wrapping the Data items by ZIP. New Application can include this Ticket Data item in the Ticket Response.

Detailed information on Header Object Data item is provided in Appendix II. Information on PKCS#7 is available in Attachment 3.

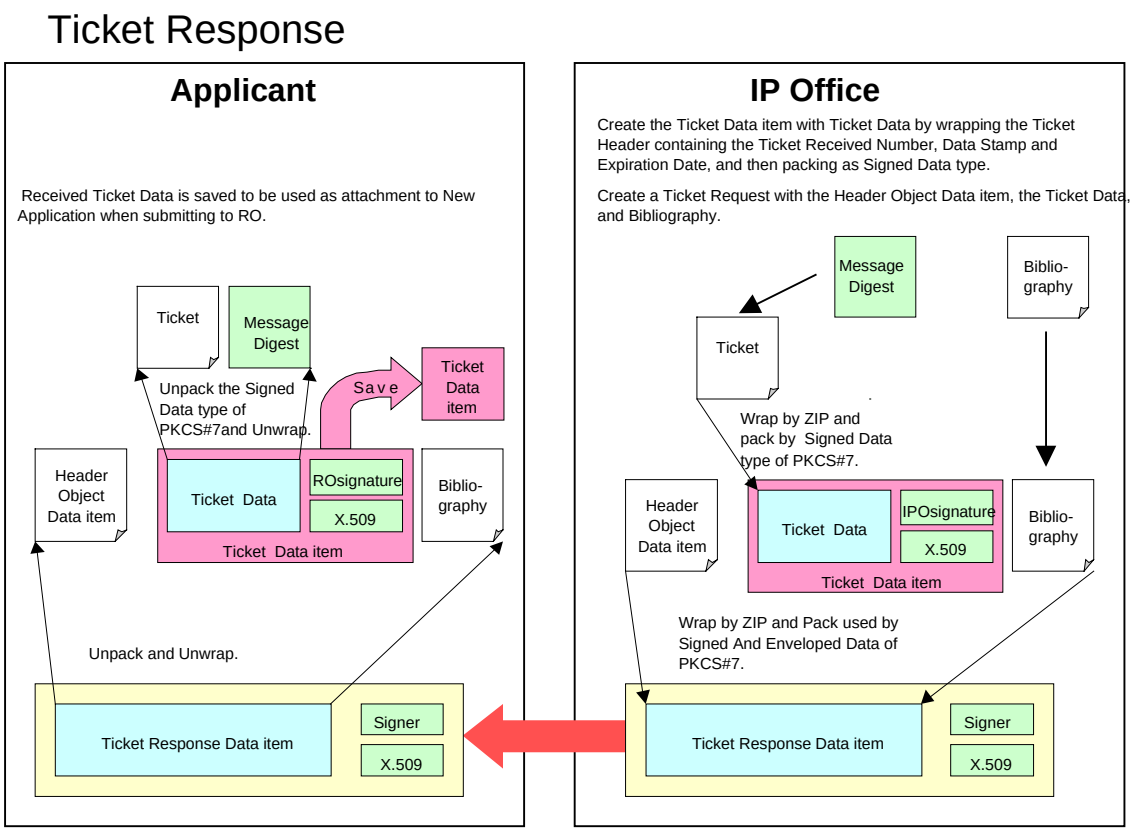


Fig4.Ticket Response

4.3 IP Document

IP Document Data item is created by packing Wrapped Documents using Signed Data Type in PKCS#7. The Wrapped Documents are wrapped in advance using ZIP according to the Wrapping Standard.

IP Document data items include a Header Object Data item specifying what the corresponding packet contains as well as the Ticket Data item received with the Ticket Response.

An IP Document is created as an Envelope by Signed And Enveloped Data type in PKCS#7 after wrapping these data items using ZIP. The Message digest from the unwrapped Ticket Data item is compared with the Message digest of Wrapped Documents included in the IP Document Data item.

Detailed information on the Header Object Data item is provided in Appendix II. Information on PKCS#7 is available in Attachment 3.

IP Document

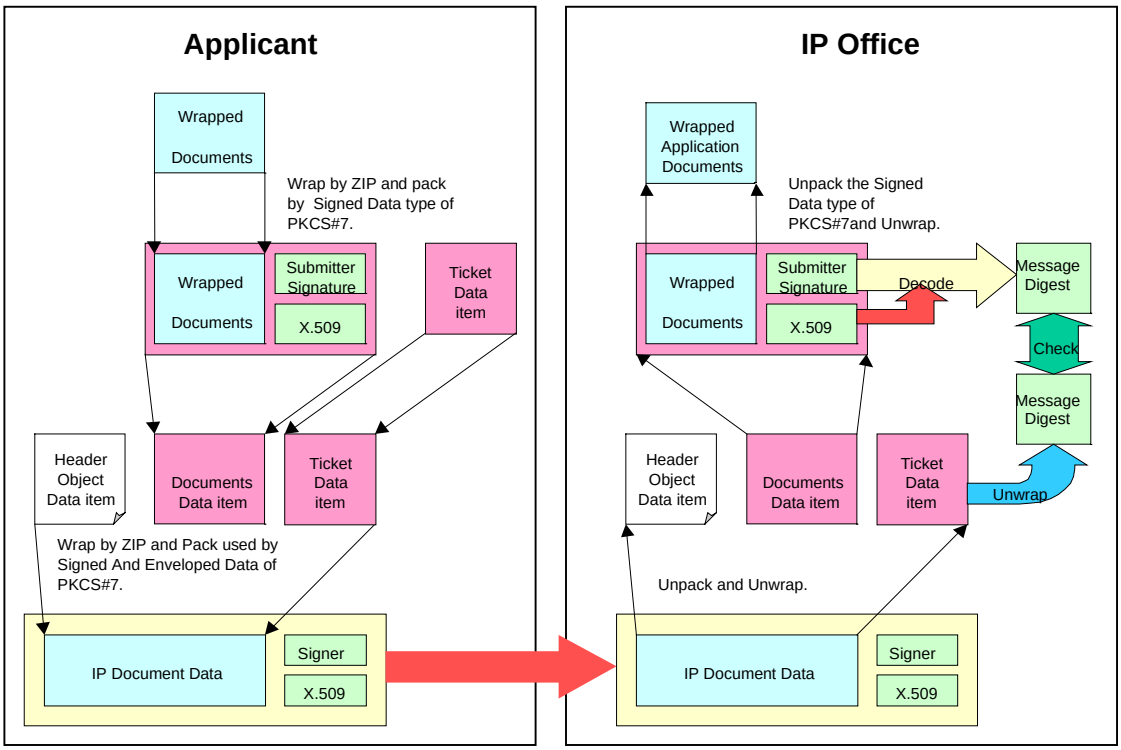


Fig5.New Application

4.4 Confirmation Certificate

The Confirmation Certificate Data item includes a Certificate Data item, a Header Object Data item specifying that the corresponding packet is a Confirmation Certificate, and an Application Documents Data item received with a New Application as an option.

The Confirmation Certificate is created by wrapping and packing the Data item using Signed And Enveloped Data type in PKCS#7.

All application procedures are concluded with a Confirmation Certificate.

Confirmation Certificate

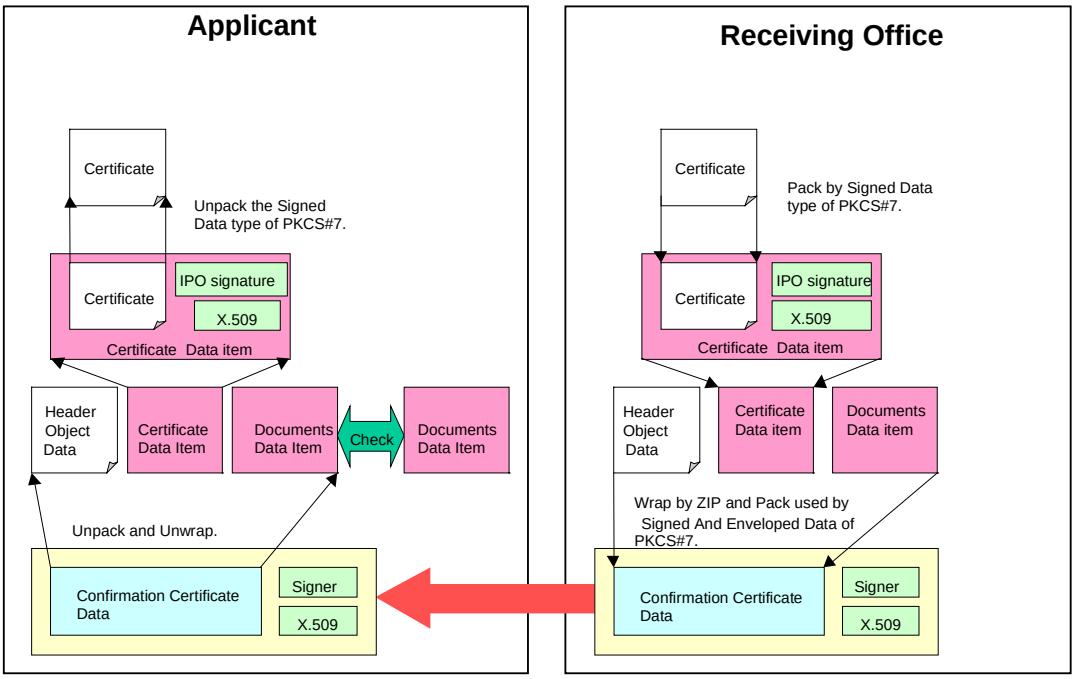


Fig.6 Confirmation Certificate

The Confirmation Certificate is used to inform the applicant of the receipt of the application must contain an XML version of this information. It may contain a formatted version of the data in PDF, TIFF and JPEG. These to files are combined in a single ZIP file and signed using the Digital Certificate of the IP Office.

Attachment 5. Checklist of Requirements for Electronic Records Management

5.1 PURPOSE

The purpose of this "Checklist of Requirements for Electronic Records Management" is to identify and define operational requirements for electronic records management that Intellectual Property Offices (IPOs) must consider when developing and operating information systems that support the management of electronic intellectual property dossiers and records.

5.2 OVERVIEW

The Standing Committee on Information Technologies' (SCIT) Strategic Information Technology Plan describes a vision for the twenty-first century. The SCIT vision is to focus on a global information technology architecture that links intellectual property offices in WIPO Member States, regional intellectual property offices and the International Bureau, for the purpose of generating, communicating and distributing information about intellectual property rights, protecting intellectual property knowledge and rights for the global economy of the twenty-first century, and global worksharing. The Plan outlines governing strategies to enhance information technology management, to achieve intellectual property office links through a global information technology architecture.

There are legal and technological considerations that require that the policies, practices, procedures, processes, requirements and solutions for managing electronic intellectual property dossiers by IPOs be addressed prior to full deployment of a global information technology architecture.

- An IPO must provide for adequate management of all records, including electronic records.

- An IPO must ensure that the integrity of all records be preserved over the full retention period, and that the confidentiality of records be protected, as required.

- Retention periods, where ready accessibility to and maintenance of electronic dossiers and records must be provided, can be permanent.

- Information technology is certain to change and advance with a frequency that will require multiple migrations of electronic intellectual property records to new storage media and new hardware and software during the prescribed records retention period.

- IPOs must provide for the adequate management of all records, including electronic records, for the required retention period. Intellectual property dossiers and records are subject to legal discovery and must, therefore, also meet the tests of admissibility in courts and during appeals and interferences.

Records management is the control of informational material that should be preserved because of its important content. Record material needs to be safeguarded because it is evidence of official policies and transactions or it contains valuable information needed long-term. Records management is required throughout the life cycle of record material—creation, maintenance, use, and disposition.

As such, records management entails much more than just the “retention” of IPO electronic records. Accurate and reliable capture and storage of electronic (and paper) records received or created by an IPO must be assured. The appropriate metadata and file formats must be applied to ensure accessibility to and migration of dossiers for the full retention period. Integrity must be preserved and confidentiality protected, as required, from the moment of receipt or creation for the full retention period. Maintaining the integrity and confidentiality, as required, of electronic intellectual property records is also essential for protecting the intellectual property rights and the value of the inventor’s business assets. This can be accomplished by ensuring that accurate, reliable electronic copies of the records are captured and then assuring that they are protected from any loss, alteration, removal, or premature destruction over the complete life cycle.

Management of electronic records information has many similarities with the traditional management of digital data in information systems. However, there are certain unique requirements. The unique areas relate to: long-term accessibility (decades); long-term retention

of records (in many cases permanently); renewal of electronic records to new storage media; transfer to new hardware, software and application systems; and unique metadata requirements that enable the management of electronic records for long-term accessibility and retention.

Given the Intellectual Property Community's plans and programs to move to electronic filing, processing and management of applications, there is a clear need for developing requirements that define policies, practices, procedures and automated information system features and controls that will provide for the effective management of the electronic records that will be received, created, stored, accessed, retained, reproduced and disseminated by these new business processes.

The overwhelming prevalence of electronically created, stored, and reproduced records (copies or duplicates) is now recognized by judicial systems. Being able to demonstrate that records offered into evidence are authentic and that they were accurately and reliably produced in the ordinary course of business is critical to admissibility. Valid tests for the trustworthiness of electronic records are more easily satisfied when proper records management procedures are part of the normal course of business.

The test of authenticity is one of the more critical and potentially difficult tests of admissibility since it is somewhat broadly defined and typically relies more on the testimony of a knowledgeable witness to establish.

The electronic record must be accurately and reliably created, received, inputted and stored. In addition it must be:

- created, transacted and/or communicated by an identifiable and verifiable party,
- received or created and stored at a specified point in time, and
- maintained in its originally inputted or transacted form; protected from any alteration or unauthorized destruction (i.e., the integrity of the record is preserved.)

5.3 DEFINITIONS

It is important to understand what is considered to be a "record" and to define what is the "official" IPO record that will be committed to the electronic dossier and managed for the full retention life. In this overall context it is also relevant to understand the concepts of "working files," "non-records" and "vital records."

5.3.1 Records

Documentary materials is a collective term for records and non-record materials that refers to all media on which information is recorded, regardless of the nature of the medium or the method or circumstances of recording. Documentary materials can become records when they are made or received by the IPO in connection with business transactions and if they are worthy of preservation as evidence of official policies or activities or because of the long-term value of the information they contain.

Working files are also documentary materials that can be deemed appropriate for preservation. Working files, such as preliminary drafts and rough notes, and other similar materials should be maintained for purposes of adequate and proper documentation if:

- They were circulated or made available to IPO employees, other than the creator, for official purposes such as approval, comment, action recommendation, follow-up, or to communicate with IPO staff about IPO business; and
- They contain unique information, such as substantive annotations or comments, that adds to a proper understanding of the IPO's formulation and execution of basic policies, decisions, actions or responsibilities.

It is important to distinguish between records and non-record materials. Applying the definition of records to most documentary materials created or received by IPOs presents few problems when the IPOs have established and periodically updated recordkeeping requirements covering all media and all activities at all levels and locations.

Certain types of documents and materials are not considered to meet the definition of a record copy and, therefore, are considered a non-record. The most notable example of a non-record is

the extra copy of a document kept for reference only. Another example is stocks of publications and processed documents. Non-record material has no real long-term informational value.

5.3.1.1 IPO Record Copy

In this Checklist, any record that is stored and managed as evidence of activities or events related to an intellectual property dossier, whether paper or electronic, is referred to as the "record copy."

A record must meet the following tests. It is:

- Made or received in connection with the transaction of IPO business, and
- Preserved or appropriate for preservation by that IPO as
- Evidence of the organization, functions, policies, decisions, procedures, operations or other activities, or because of its long-term informational value.

An extra reference copy of a record is a non-record copy.

Essentially, any incoming documents from an applicant to an IPO, whether paper or electronic, that are materially related to an intellectual property dossier is a "record copy" because the record is evidence of transacted business by the IPO, and also because the record represented what an applicant would have considered to be their record copy at the time of submission to the IPO.

Any document created by an IPO when finalized and communicated to an applicant and/or otherwise stored to an IPO dossier is a record.

Any working file that was circulated or made available to IPO employees other than the creator for official purposes, and that contains unique information that adds to the understanding of basic policies, decisions, actions or responsibilities, is considered record copy and must be stored in the electronic intellectual property dossier. This definition also implies that all working files created and kept solely by an IPO employee, for instance, and not circulated to other employees for any official purpose, would not be considered a record copy.

5.3.2 Electronic Record

An electronic record is defined within the context of a general record, but with the added characteristic of information being represented in a digital form that only a computer can process. Electronic records include numeric, graphic, and textual information, which may be recorded on any medium capable of being read by a computer and which satisfy the definition of a record. This includes, but is not limited to, magnetic media, such as tapes and disks, and optical disks. Unless otherwise noted, these requirements apply to all electronic records systems, whether on microcomputers, minicomputers, or mainframe computers, regardless of storage media, in network or stand-alone configurations.

5.3.3 Vital Record

Essentially, all intellectual property dossiers, including any related annotations, links and metadata are considered vital records. As such, a backup copy of all electronic intellectual property dossiers, including associated metadata, must be made and managed for their full retention period, as a means of providing for business recovery from a disaster.

Vital records are essential to the preservation of the legal rights and interests of IPOs and individuals conducting business with IPOs. While these records require protection, storage points need not be at or in the vicinity of emergency operating centers.

5.3.4 A Complete Record

A complete record has three primary elements: content, structure, and context.

Content is the actual data resulting from a transaction conducted in the normal course of business, such as from a receipt of a patent application. For example, the content of a patent application transmittal record includes various data fields related to specific components of the form plus a signature.

Structure is generally defined in two parts: logical structure and physical structure. The logical structure of a record includes the identifiable parts of the record, such as the title, applicant

address, date and “signature” on a patent application form. These parts may be both computer identifiable, as in metadata, and/or human identifiable (graphical) when rendered on a viewing screen or printer. The physical structure relates to the format of the record, such as the type font, spacing, page margins, logo, and the “encoding” or format of the file, which provide information for processing (rendering) or transferring the record over the retention period.

Context is the meaning of the record, or the “what” and “why” of the business transaction from which the record was created or received. The context may be implicit in the content and structure of the record, such as a patent application form, which contains a form number, or phrase and a signature block, which states the *intent* of the signer. The context may also include the general environment within which the records are stored and managed, e.g., records managed within an intellectual property dossier, or dossiers that are managed as part of a larger IPO dossier repository.

One of the key requirements for admissibility as evidence in a court proceeding is that the system receiving or creating the record store an “accurate” representation of the record. A record is more likely to be perceived as accurate and reliable and, therefore, trustworthy when as many elements of the record as possible are documented. The more “complete” a record can be shown to be, the more likely it would be considered “authentic” (that it is what it purports to be) for purposes of admissibility as evidence, and the more weight it would likely carry as evidence.

A record may consist of one or more files (such as in the case of a compound record consisting of a text file and a graphics file), with the content, structure and context of each being separately identified, either as part of the record or as metadata about the record and files. The electronic dossier will consist of one or more records with the content, structure and context of each record being separately identifiable.

5.3.5 Record Retention

Retention periods vary based on the status of the dossier, i.e., whether the patent is issued or abandoned, and also based on the period of time during which the application was filed.

The determination of retention period for any given record is a matter of business need, legal requirement, or historical value. In other words, is the record needed to conduct business? If yes, it should be retained so long as that need exists. It is more obvious and less of an intuitive decision if a legal requirement or historical value determine retention period.

An example of a business need determining retention is documentation of an automated system, which should be retained for the life of the system. An example of a legal requirement determining retention is a treaty prescribing record retention for a specific period. An example of historical value determining retention is the files of the head of an IPO, which document the history and intellectual property process of the IPO.

5.3.5.1 Dossier Retention

For the scheduling of records retention, the complete electronic intellectual property dossier or dossier is assigned a single retention period. Accordingly, all of the electronic records in the dossier can be managed and retained as a single entity.

5.3.5.2 Retain Only One Record Copy

For operational and for legal purposes only one “record copy,” plus a vital records disaster backup copy, of an electronic record and electronic dossier file should be retained.

From an operational perspective, having a single record copy of the electronic dossier avoids any confusion as to what the “official file copy” is and also may avoid someone retaining and accessing a copy that is not up-to-date. This means that copies retained purely for reference should be deleted at the earliest practical point in time. Working files should either be committed to the electronic dossier if they meet the retention criteria or they should be deleted at the earliest practical point in time.

From a legal discovery perspective, *all* records pertaining to a subpoenaed subject matter, such as one or more specific patent applications, are discoverable and must be produced if requested. As such, any reference copies or working files that have not been deleted are subject to discovery and must be searched, retrieved and produced if requested.

5.3.6 Usage Periods

When viewed from the perspective of records management, there are two logical, progressive periods of time in the life of an electronic intellectual property dossier.

1) **In-Process and Use**: This period defines the "pending" period for a patent application. During this period, the electronic dossier and all electronic records are being managed by the information systems and the associated work flows that control the filing, reviewing, examination, and the issuance, registration or abandonment processes.

2) **Maintenance and Use**: This period defines the management of electronic intellectual property dossiers subsequent to the issuance or abandonment of an intellectual property application. In the current paper process, this is the point when the dossier is transferred to an IPO repository and enters a "maintenance" period for the remainder of the authorized retention period prior to being transferred for archival storage. Also at this point, the retention period for the dossier is established.

   During the maintenance and use period, electronic records and associated metadata may be added to the dossier, such as assignments, or certain metadata may be added or updated, such as the address of the applicant or inventor. It is also during this period when migration and retention management events are most likely to occur. Examples include media renewal; transfer to another hardware, software or application system; transfer to archival storage; and the disposal of records at the end of the retention period. On an exception basis, the electronic dossier may be subjected to a more active process, such as a reexamination, or an appeal or interference proceeding.

5.4 REQUIREMENTS

This section defines and describes the requirements for electronic records management over the complete life cycle of the intellectual property dossier. The following requirement areas are addressed:

- Records Acquisition
- Metadata
- File Management
- Preserve Integrity
- Protect Confidentiality
- Access Controls and Authentication
- Search, Retrieval and Reproduction
- Audit Trail
- Vital Record Backup and Recovery
- Records Retention
- Migration
- Transfer to Archival Storage
- Records Hold

5.4.1 Records Acquisition

The capture of electronic records by IPO information systems applies to all documentary materials that meet the tests for a record copy, including records received or created electronically and those received and converted from paper or microfilm to electronic form. The preciseness and reliability of the acquisition process is critical to obtaining an accurate and complete electronic record—a record that preserves the content, structure and context of the information.

5.4.1.1 Capture of Complete Electronic Records

It is important to acquire a complete electronic record: a) to ensure that an accurate copy of all elements of the record are captured and b) to ensure long-term processability and transferability.

The requirements for capturing a complete record are:

- ❏ Provide for the capture of all received and created electronic records that meet the tests of a record copy and commit the records to an electronic intellectual property dossier. Such records might include:
  - ❏ Records received from an applicant.
  - ❏ Office actions created by an examiner.
  - ❏ Communications between the examiner and the applicant, including e-mail.
  - ❏ Working files that meet the retention criteria.
  - ❏ Other documentary materials that are *appropriate for preservation*.

- ❏ The act of "committing" documentary intellectual property material as a record copy to a dossier should be a distinct, conscious and auditable event.
- ❏ Capture the content, structure and context of each record.
- ❏ Capture attachments and addenda as separate records such that they do not alter the record to which it is linked.
  - ❏ Logically link the attachment or addenda to the associated record copy.
- ❏ Capture metadata associated with the electronic dossier and records that allows for search, retrieval, routing, confidentiality status, migration and retention management. (see Section 4.2, Metadata)
- ❏ Establish file format standards for receiving, creating and storing records that are processible and transferable for the full retention life of the record.
- ❏ If multiple renditions (same content but different file format, such as a MS Word and an XML rendition of the same document) of the same record are captured, each should have the same content, structure and context.

5.4.1.2 Capture Links to Notes and Annotations

- ❏ Capture notes and annotations as "logical" additions to the record and ensure that they do not alter the content or structure of the record.
- ❏ Allow users to position notes and annotations on the document in a meaningful location. Retain the location of the annotation, like a bookmark.
- ❏ The electronic record should be viewable and reproducible at any time without the notes or annotations.

5.4.1.3 Capture Hyperlinks

- ❏ Capture hyperlinks within an electronic record that refer to another part of the same record.
- ❏ Capture hyperlinks within an electronic record that refer to other electronic records within the same dossier.
- ❏ Hyperlinks to electronic records outside of the electronic dossier should not be allowed, unless a method is provided to update the hyperlinks whenever the record location or the record or hyperlink changes.

5.4.1.4 Working Files

Capture working files when they meet the requirements for preservation as a record copy. Working files, such as preliminary drafts and rough notes, and other similar materials should be maintained for purposes of adequate and proper documentation if:

- They were circulated or made available to IPO employees, other than the creator, for official purposes such as approval, comment, action recommendation, follow-up, or to communicate with IPO staff about IPO business; and
- They contain unique information, such as substantive annotations or comments, that adds to a proper understanding of the IPO's formulation and execution of basic policies, decisions, actions or responsibilities.

Electronic working files are managed outside of the dossier until it is determined that they are to become a record copy, then they are committed to and managed as part of the dossier.

- ❏ Establish a process for committing a working file as a record copy.
- ❏ Use version control to capture and track the creation history of working files and link successor records to the predecessor records.

5.4.1.5 Conversion of Paper Records to Electronic Form

When it is required that documentary materials in paper form be converted to electronic form, specific guidelines should be followed in order to ensure that accurate and complete records are captured and that the process is consistent and reliable.

- ❏ Ensure that the scanning resolution for image capture is sufficiently high to capture a readable, usable and reproducible copy of the original.
- ❏ Perform a high level of quality control to ensure that accurate and complete records have been captured.
- ❏ Periodically test the scanners to ensure that they are operating according to manufacturers' specifications and are producing the desired quality level for the documents being scanned.
- ❏ Provide a means to rescan documents that quality control has shown to be of insufficient quality, or mark the documents prior to scanning as "best copy."
- ❏ Track the batch number for both the imaged and paper documents as a means of accessing any documents determined to be of insufficient quality.

5.4.1.6 Quality Control

A quality control step should be an integral part of the record capture process, independent of whether the documentary materials are being scanned from paper or are being automatically acquired from electronically received or created sources.

- ❏ As an integral part of the capture process, conduct quality control on a sample of the captured electronic records, whether scanned from paper or acquired from electronically received or created sources.

- ❏ The sampling level required will depend on the quality level of: a) the source materials, higher for image-scanned paper and lower for documents acquired from electronic sources, and b) the error levels encountered during the quality process. Sampling rates should be adjusted to reflect the level of errors per sample determined during the quality control process.
- ❏ For records captured from paper, a thorough quality control should be conducted to ensure accurate, complete and readable information is acquired from the original paper document.
- ❏ If Optical Character Recognition is used to convert scanned image documents to computer readable format, the converted text should undergo meticulous quality control.
- ❏ For records captured from an electronic source, the sampling level would normally be lower, depending on the accuracy and completeness of prior samples measured.

5.4.1.7 Quality Assurance

In addition to the quality control process, a quality assurance sampling process should be conducted by the IPO to ensure that all records are accurately and reliably captured. The sampling level can generally be lower than that used for quality control, unless the observed quality level of the captured records warrants an increased sample size.

- ❏ Regularly conduct a final quality assurance process as the means of being certain that all records and associated metadata are being accurately and reliably captured.
- ❏ It is preferable that the IPO conduct the final quality assurance, particularly if a third party contractor is performing the record capture and/or quality control.

5.4.1.8 Audits

IPO personnel should periodically conduct audits of the document capture process to ensure that all procedures and guidelines are being followed and that accurate and reliable records and associated metadata are being acquired.

- ❏ Conduct period audits of all elements of the document capture process and adjust scanning, quality control and assurance procedures and levels accordingly.

5.4.2 Record metadata

Record metadata can be defined as “data describing stored data”, that is, data describing the schema or structure of data elements, their interrelationships, and other characteristics of electronic record data. The term "metadata" has been widely used to characterize the descriptive information that supports search and retrieval of both hardcopy (paper and microfilm) and

electronic material. Metadata also includes additional information, such as file formats and creation sources that must be acquired and retained in order to effectively manage electronic records over long periods of time, including those designated for permanent retention.

5.4.2.1 Metadata Management

The records for intellectual property applications are managed in dossiers, whether paper or electronic. The metadata required to provide accessibility, auditability and transferability of the electronic records must also be managed in conjunction with the dossier. It is recommended that the metadata be integrated or encapsulated in the electronic dossier.

5.4.2.1.1 Metadata Retention

The retention management requirements for the electronic dossier and records must also be applied to the related metadata.

- One of the fundamental requirements for long-term access and retention of metadata is that it must be retained for the same period of time as the electronic dossier and records. This includes maintaining the accuracy and completeness of all dossier and record-related metadata through media renewals and through transfers to new hardware or software or new application information systems, and transfers to archival storage.

5.4.2.1.2 Metadata Profiles

It is recommended that the following metadata record profiles be implemented and that metadata elements be defined, acquired and retained for the following areas:

Dossier and Dossier Records profiles: These profiles contain metadata for storage, search and retrieval, confidentiality and tracking accesses.

Use History (see Section 4.8 Audit Trail): This profile documents an audit trail of events and actions (such as accesses to and migrations of electronic intellectual property dossiers) in order to provide a basis for: a) establishing the reliability and general trustworthiness of electronic intellectual property dossiers and records, b) use in meeting the test of authenticity for

admissibility, and c) for tracking or researching events related to proving that integrity has been preserved or that confidentiality has been protected.

Copy, Reformat and Transfer Profiles: These profiles document the copy or reformat of electronic dossiers in the event of media renewal; the transfer to a new hardware and software system or information system; or the transfer to archival storage.

The Metadata Profiles do not identify *how* each metadata element should be captured. Nonetheless, it assumes that many metadata elements, such as those for the Dossier Profile and Dossier Record Profile could be automatically acquired from the IPO information systems. Many of the other metadata elements, such as access dates, may be supplied automatically by the information system or, as a last resort, via manual data entry.

5.4.2.1.3 Dossier Encapsulation

It is recommended, primarily for ease of migration and for auditability purposes, that all metadata information be encapsulated within the electronic intellectual property dossier. Encapsulation of all metadata as part of the dossier provides a "single digital object" that contains the elements necessary for accessing and for tracking activity related to the dossier. Encapsulation also makes it easier to identify all information that needs to be migrated in the event of media renewal or the transfer of dossiers. Encapsulation can be done using either a logical or physical approach.

Logical encapsulation: "Links" all of the records, annotations and notes, and metadata profiles that comprise a dossier as a "logical" single digital object. However, the records and metadata related to a dossier may reside on different storage servers and different media volumes. While logical encapsulation may represent the most flexible method for storage, it creates a potentially complex management environment and, as such, a higher risk of loss or corruption when performing media renewal or transfer of dossiers.

Physical encapsulation: This means that all of the information associated with a specific electronic intellectual property dossier, such as the records, related annotations and notes, metadata, and intra-or inter-dossier or record linkages, exists as a single physical object or entity residing on the same volume of media. Physical encapsulation may provide the most

straightforward environment for media renewal and transfers of dossiers, however, it could pose issues related to the resources and time required to keep the dossier physically encapsulated.

5.4.2.2 Dossier and Record Profile Metadata

One of the primary purposes of metadata is to provide an easy and efficient means for retrieval of the electronic records by authorized personnel. One type of metadata is defined as "profile data." The Dossier and Dossier Record profiles are designed to include information such as: a unique identifier for the dossier and each record in the dossier, the subject or title of the dossier or record, the creator or origin of the record; why, when, and how the record came into existence; accuracy of the source data; source granularity; processing status; use history; and the quality and extent or scope of the resource.

The following guidelines are provided for capturing and managing metadata for purposes of retrieving electronic records.

- ❑ Each electronic dossier and each record in an electronic dossier should have at least one unique metadata identifier that differentiates them from other electronic dossiers and records.
- ❑ Encapsulate Dossier and Dossier Record profile metadata in the related electronic dossier for effective retention and migration management. (See electronic Dossier and Dossier Record metadata requirements below).
- ❑ Use IPO information systems to provide the primary source of metadata for the Dossier and Dossier Record profiles.
- ❑ Use IPO information systems for search and retrieval of electronic dossiers and records.
- ❑ Provide a link from IPO information system data bases to the electronic dossiers.
- ❑ Define specific document-type designations that allow for retrieval of only specific records within an electronic dossier.
- ❑ Provide for the identification of hard copy records and other non-electronic materials, where relevant, as Dossier Record Profiles so that retrievals are kept simple, e.g., one search to identify all records available, regardless of media type or physical location.
- ❑ Use the Dossier Metadata profile to indicate the status of confidential records, e.g., identify confidential, privileged and private records so they can be secured from unauthorized users.
- ❑ Where possible, automatically index all records received or created in accordance with an IPO-wide filing scheme for electronic dossiers that is to be developed using the metadata profiles defined for Dossiers and Dossier Records.
- ❑ For ease of mass retrievals, allow for designation of common predetermined metadata fields, e.g., application number, subject, date, originator, record type or form number, and disposition code.

- ❑ No alteration of metadata associated with (or encapsulated with) an electronic dossier should be allowed, unless it is determined that the metadata contains errors that must be corrected.
- ❑ Should corrections to metadata be required (such as due to errors induced during automatic or manual data capture), access to and use of modification programs or tools for correcting metadata should be limited and controlled for use only by authorized personnel. Audit trail information about any changes should be added to the Use History metadata profile.

5.4.2.2.1 Dossier Profile

The recommendations for metadata in the Dossiers and Dossier Record profiles are based on the needs for filing, retrieval and disposition of electronic intellectual property records.

The data elements comprising the Dossier Metadata Profile capture core information about the dossier as a logical entity that supports long-term access and retention management. All of these data elements are considered non-revisable from the moment that they are registered as part of the dossier. The only instances where an authorized modification is allowed in the metadata of the dossier profile is when there are updates to the dossier for error correction, or after issuance, registration or abandonment, such as assignment of rights or reexamination, or if the storage location status changes, or when access dates are transferred from the individual record profile.

SAMPLE DOSSIER METADATA PROFILE

- ❑ Unique Identifier
- ❑ File Type
  - Patent
- ❑ File Subject
- ❑ Applicant* (may be changed if attorney or agent changes)
- ❑ Inventor(s)* (may be changed via an assignment)
- ❑ Filing Date
- ❑ Closure Date (defines the beginning of the authorized retention period)
- ❑ Process Status Code (pending, issued, registered, abandoned)
- ❑ Representation
  - Binary
  - ASCII

- UNICODE

❑ Encryption of Dossier (optional, if required)

- Name of Algorithm Used
- Name of Software and Version Used

❑ Formats Used

- Text
- Image
- Compound (text and image)
- Data Base

❑ File Formats Used

- TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)
- JPEG
- PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[1]
- XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)
- SGML
- Other (e.g., complex work units)

❑ Dossier Size

- Logical Record Length
- Logical Record Count*
- Physical Record Count*
- Byte Count*
- Total Number of Records*

❑ Dossier Authentication (If Used)

- Cyclical Redundancy Check (CRC)
- Hash Digest

❑ Updates After the Pending Period*

---

[1] PDF may be used if permissible for long-term archival storage by individual IPOs.

- Assignment
- Reexamination
- Change of Address
- Etc.

❑ Access/Event Dates List*
- Date of Access/Event
- Type of Access/Event
- I.D. of Access/Event User

❑ Location of Dossier*
- In Process and Use
- Maintenance and Use
- Archival Storage

❑ Location of Paper or Microfilm Records Related to Dossier

❑ Disposition*
- Instruction Code
- Action Date

*Indicates update to the profile is permitted

5.4.2.2.2 Dossier Record Profile

The data elements comprising the dossier record profile capture core information about each record as a logical entity that can be used individually or collectively to provide control of the records. For example, the Unique Record Identifier provides for direct access to the record, and the Access/Event Date element contains all instances of access to or other actions related to the record. A specific date can be linked to the Use History Profile, for example, that would disclose if the access was linked to an update, reformat, copy, or transfer activity. Each of these data elements is considered non-revisable from the moment the record becomes part of a dossier. The only instance of a change that could occur in the dossier record profile is when access dates are added. An asterisk ("*") identifies these instances.

SAMPLE DOSSIER RECORD METADATA PROFILE

- ❑ Dossier Unique Identifier
- ❑ Unique Record Identifier (e.g., serial number)
- ❑ Record Descriptor (form number, alpha code)
- ❑ Date Received/Created
- ❑ Name of Record Recipient/Addressee
- ❑ Subject
- ❑ Originating Organization
- ❑ Author
- ❑ Security Level (Pending, Issued)
- ❑ Representation
  - Binary
  - ASCII
  - UNICODE
- ❑ Encryption of Record (optional, if required)
  - Name of Algorithm Used
  - Name of Software and Version Used
- ❑ Formats Used
  - Text
  - Image
  - Compound (text and image)
  - Data Base
- ❑ File Formats Used
  - TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)
  - JPEG

- PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[2]

- XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)

- SGML

- Other (e.g., complex work units)

❑ Record Size in Bytes

❑ Record Authentication (If Used)

- CRC

- Hash Digest

❑ Access/Event Dates*

- Date of Access/Event

- Type of Access/Event

- I.D. of Access/Event User

*Denotes that updates are permitted

5.4.3 File Management

The file management system must protect the physical and referential integrity of electronic dossiers and records and associated metadata over the full life cycle. The file management system must protect an electronic record and associated metadata from being overwritten or inadvertently deleted. It should also provide the means to enable access to electronic dossiers, records and metadata for the full retention period, independent of the storage media—digital, paper or microfilm.

❑ The file management system must not allow electronic dossiers and records or related metadata to be overwritten or inadvertently deleted.

---

[2] PDF may be used if permissible for long-term archival storage by individual IPOs.

- ❑ Using access controls, the disposal of electronic records and associated metadata should be restricted to specific personnel who are authorized by the IPO Records Officer or equivalent/designate to destroy or transfer dossiers and associated metadata.
- ❑ For accessibility purposes, a unique link or pointer should be created and maintained between the IPO's retrieval and tracking information systems, the file management directory and/or the metadata, and the physical location of the electronic dossier and records.
- ❑ For vital records backup of electronic dossiers and records and related metadata, provide a link or pointer between the retrieval and tracking information systems, the file management directory and/or the metadata profiles, and the physical location of the electronic dossiers, records and related metadata.
- ❑ Capture metadata to locate any records related to a dossier, which are maintained only in physical form or on physical media (such as CD-ROM).

For management of physical storage media, see Section 4.11.2 Media Management.

5.4.4 Preserve Integrity

Integrity means ensuring consistency of data, in particular, preventing (including detecting) unauthorized alteration or destruction of data.

Maintaining the integrity of intellectual property records is essential for ensuring the protection of intellectual property rights and the value of the inventor's business assets. For these reasons, checks and balances must be in place to assure the preservation of the integrity, authenticity and trustworthiness of the IPO's records over time. This may be accomplished by managing and protecting the electronic records from any loss, alteration, removal, or premature destruction. Since evidence of record tampering may not be as readily identifiable with electronic records, as it might be with paper records, it is even more important that the controls are in place to adequately protect and preserve the electronic record.

5.4.4.1 Protect Against Alteration

- ❑ Access controls to electronic dossiers and records should be applied on a "least privilege" basis, with access limited based on the role or function of an individual.
- ❑ No alteration of the content, structure and context of a record committed to an electronic dossier will be allowed. The integrity of each record must be preserved for the full retention period of the dossier.

- ❑ Any annotations and addenda related to an electronic dossier or record should also be protected from alteration or loss and should be maintained as data that is separate from, but logically linked to the record, so that an electronic record can be viewed or reproduced in the original manner it was stored.
- ❑ No alteration of metadata associated with (or encapsulated with) an electronic dossier or record should be allowed, unless it is determined that the metadata contains errors that must be corrected.
- ❑ Should corrections to metadata be required (such as due to errors induced during automatic or manual data acquisition), access to and use of modification programs or tools for correcting metadata should be limited and controlled for use only by specifically authorized personnel. An audit trail of any modification activity should be tracked in the Use History metadata profile.
- ❑ If new electronic records are to be created using an existing electronic record as the basis, the new electronic version must be committed to the electronic dossier as a new record with new dossier record profile metadata. Version controls (such as check-out/check-in procedures) must be used to ensure that the original electronic record remains unaltered.
- ❑ Access to programs or functions used for destroying electronic dossiers and records at the end of the authorized retention period (including associated attachments and annotations, and the metadata) should be limited to persons or system functions specifically authorized by the Information Systems Security Officer or equivalent/designate.

5.4.4.2 Validating Integrity

- ❑ An electronic record should contain some feature, such as a cyclical redundancy check (CRC) or a digital signature hash, that allows the integrity of the record to be validated at the time of each receipt/creation and for each access, and allows for detection of any attempt at or alteration of the record to be detected.
- ❑ Each validation of integrity should be recorded in the Use History metadata profile.

5.4.5 Protect Confidentiality

Confidentiality is defined as ensuring that information is not disclosed or revealed to unauthorized persons. Confidentiality applies to pending and abandoned electronic patent records and associated metadata, except when the abandoned patent is referenced by an issued patent. Confidentiality is not a requirement for trademark records. Maintaining the confidentiality of pending or abandoned patent dossier and record content is crucial because the information is confidential and highly valuable intellectual property of the owner.

The following requirements apply only to protecting the confidentiality of pending and abandoned patent dossiers.

- ❑ Protect confidential electronic records and associated metadata from unauthorized viewing during transmission, using encryption.
- ❑ Protect confidential electronic records during the review, pre-examination, examination, and publication steps, and during maintenance of abandoned patents, and ensure that they are not disclosed or revealed to unauthorized persons.
- ❑ Provide a metadata field that indicates when the status of an electronic dossier is confidential.
- ❑ Access controls to confidential records should include the ability to restrict access on a "least privilege" basis within one or more specified functions or roles for an individual.
- ❑ Physically secure off-line media that contain confidential records from unauthorized access.

5.4.6 Access Controls and Authentication

Providing access controls (identification) and authentication (validation of identity) is critical for protecting the confidentiality and preserving the integrity of electronic dossiers and records. It is required to prevent unauthorized viewing, modification, destruction and, generally, the unauthorized issuing of commands. Access controls should be based on "least privilege," granting users access only to those electronic dossiers and records, and to associated annotations and metadata, that are minimally required to perform their roles or functions. Controls may selectively limit access to electronic dossiers and records, to any action or event related to electronic dossiers, and to specific computing resources related to migration or purging.

- ❑ Provide access controls and authentication, such as public/private key pair, to authenticate the identity of the sender, creator, and user of the record.
- ❑ Identify and authenticate each user at the time of log-on to all IPO information systems that receive, create, process, maintain or otherwise manage electronic dossiers and records.
- ❑ Access controls and profiles should include the ability to restrict access by an individual on a "least privilege" basis to one or more specific functions or roles as well as to selected computing resources based on the process status of the dossier.
- ❑ Access should be strictly limited to any information system or other computing functions or resources that provide for the modification of metadata, creating new versions of documents, and the scheduled purging (deleting) of electronic dossiers and associated metadata.
- ❑ Use an audit trail to track all events related to access of the record, including the detection of unauthorized access attempts.

5.4.7 Search, retrieval and reproduction

5.4.7.1 Search and retrieval

- ❏ Search for all dossiers or records relevant to an authorized request, whether they are stored on on-line, near-line, or off-line media.
- ❏ Search options should include:
  - ❏ Searching the metadata for a specific "dossier."
  - ❏ Searching a dossier for a specific record.
- ❏ Allow for a variety of types of searches:
  - ❏ Linguistic-based word matching, e.g., knife or knives; record keeping, recordkeeping, or record-keeping.
  - ❏ Boolean searches, e.g., water AND H_2O.
  - ❏ SQL expressions, e.g., like, contains.

5.4.7.2 Store Search Results

- ❏ Allow users to store the search results as a record, e.g., a list of all intellectual property documents that were reviewed during a search process of an intellectual property application examination.
- ❏ Allow users to temporarily store and then retrieve search results until a business task or event is complete.

5.4.7.3 Display/Print the Record, the Index and Annotations

- ❏ Ensure that a human–readable copy (screen display and hard-copy) of the complete record can be generated over the life of the record.
- ❏ Display/print options should include any combination of the following:
  - ❏ Display/print of the record.
  - ❏ Display/print of the annotations.
  - ❏ Display/print of Dossier and Dossier Record metadata.
- ❏ Distinguish notes and annotations from the record.
- ❏ Permit printing and display with annotations on the record, annotations following the print of the record, or no annotations.
- ❏ Provide for zooming in and out to easily distinguish hard-to-read text and graphics.

5.4.7.4 Applicant Access

Where access to electronic dossier, records and/or metadata information is provided, measures must be taken to restrict access only to the records and metadata that an applicant is authorized to

view. The integrity of the electronic dossier records and metadata must be preserved, and the confidentiality of pending patent application information must be assured.

- ❑ Provide a means of identifying and authenticating the applicant, preferably through digital certificates and Public Key Infrastructure technology.
- ❑ Provide a means of restricting access to only those dossiers, records and/or metadata that the applicant is authorized to access and view.
- ❑ Log all accesses by applicants in the Use History Metadata Profile as an audit trail of these events.

5.4.8 Audit Trail

An audit trail, also know as an event log, is the "chain of custody" that records the "who, what, and when" for each action or event, including creation or receipt, processing, access, distribution, dissemination, migration, transfer and disposal of an electronic intellectual property dossier. As such, an audit trail essentially captures specific metadata that documents the intent and result of activities related to electronic dossiers and records. The audit trail can provide an independent (computer-controlled, not human-controlled) element of proof that policies and procedures were followed and that the integrity and confidentiality of the information was not compromised. From a legal perspective, an audit trail can be used as "management data" to help prove the authenticity of a record for admissibility in evidence in an interference or appeal proceeding. It is recommended that audit trail information for accesses and events related to intellectual property dossiers be documented using the metadata in the Use History Profile.

5.4.8.1 Use History Profile

As set forth in Section 4.2.1.2 Metadata Profiles, it is recommended that a Use History metadata profile be created for each intellectual property dossier, and encapsulated in the electronic dossier, as the audit trail of accesses and events related to the dossier. The Use History metadata profile logs accesses and events (such as migrations of electronic dossiers). These provide a basis for: a) establishing the reliability and general trustworthiness of electronic intellectual property dossiers and records, b) meeting the test of authenticity in admissibility, including appeals, and c) tracking or researching events related to proving that integrity has been preserved, that confidentiality has been protected, that system procedures have been followed and that the

management of electronic records has been conducted in a trustworthy manner. The following metadata are recommended for inclusion in the Use History profile:

SAMPLE USE HISTORY METADATA PROFILE

- ❑ Dossier Unique Identifier
- ❑ Event/Activity Date (Repeatable)
- ❑ Access Dates (Repeatable)
- ❑ Reformat (Repeatable)
  - Date
  - Reformat Iteration Number
  - Logical Record Size
  - Logical Record Count
  - Byte Count
  - CRC (If Used)
  - Hash Digest (If Used)
- ❑ Copy (Repeatable)
  - Date
  - Copy Iteration Number
  - Logical Record Count
  - Byte Count
  - CRC (If Used)
  - Hash Digest (If Used)
- ❑ Transfer (Repeatable)
  - Date
  - Transfer Iteration Number
  - Logical Record Count

- Byte Count
- CRC (If Used)
- Hash Digest (If Used)

❏ Purge/Delete (Optional)

- Date
- Authorization

5.4.8.2 Use History Profile Creation and Update

❏ Create a Use History Profile at the time the intellectual property application is filed and log the "filing" event metadata in the Use History Profile.

❏ Log each access to an electronic dossier in the Use History profile.

❏ Log to the Use History Profile each event that receives, creates, updates, disposes of or is otherwise material to the prosecution and maintenance of an intellectual property dossier.

❏ Log each copy, reformat and transfer event in the Use History Profile

❏ Log the disposition action and date to the Use History File when retention schedules are applied and dossiers are disposed.

❏ Log the types of dossiers transferred, the dossier count transferred and the date transferred when dossiers are transferred to archival storage.

5.4.8.3 Link to Other Information System Tracking or Event Logging Systems

❏ Link to IPO workflow or process tracking information systems to extract or capture information related to maintenance events, such as reassignments of a patent from one party to another.

5.4.9 Vital Records Backup and Recovery

Intellectual property dossiers are considered vital records, that is, they must be available for purposes of reconstructing the business of an IPO should a disaster occur that damages the primary electronic dossiers.

To ensure that this vital information is backed up and available, specific policies must be defined and specific procedures implemented and practiced.

A vital records backup copy of all electronic dossiers and associated metadata must be made and available for disaster recovery purposes. A backup copy of all information systems and operating software used to manage the electronic dossiers should also be available for disaster recovery.

- ❑ Develop automated backup policies and procedures for electronic dossiers, metadata and associated information systems and operating software.
- ❑ Create a vital records backup copy of all electronic dossiers and records, and related metadata.
- ❑ Create a backup copy of all operating software and information systems that process and maintain electronic dossiers, records and related metadata software
- ❑ Store all vital records backup copies in a geographically separate location from the primary media.
- ❑ Provide electronic and physical (where appropriate) labeling that allow ready identification of the location of and access means to the backup media.
- ❑ Develop procedures and implement automated and manual capabilities, as required, that allow for full recovery of electronic dossiers and records and related metadata in the event of a disaster.
- ❑ Maintain metadata and/or file directory that allow access to both the primary and backup media for each electronic dossier and records, as well as for any associated hard copy records.
- ❑ Apply retention management, including disposal of electronic dossiers and records and related metadata, consistently for the backup media and the primary media.
- ❑ Secure the vital records backup media and protect it from environmental and other potential harms, including: a) ordinary hazards, such as fire, water, mildew, rodents, and insects; b) human-made hazards, such as theft, accidental loss, sabotage, and commercial espionage; c) disasters, such as fire, flood, earthquake, wind, and explosion; and d) unauthorized use, disclosure, and destruction.

5.4.10 Records Retention

IPO's must ensure that electronic dossiers and records are retained as long as they are needed. Retention procedures should include provisions for scheduling the disposition of all electronic records and their related documentation and indices. The information in electronic records systems should be scheduled as soon as possible but no later than one year after system implementation. Provision should be made for transferring a copy of the electronic records and related documentation and indices to archival storage. It is also important to establish procedures for regular recopying and reformatting and the performance of other necessary maintenance to ensure the retention and usability of electronic records throughout their life cycle.

Maintaining the integrity of intellectual property records is essential for ensuring the protection of the intellectual property rights and the value of the inventor's business assets. For these reasons, checks and balances must be in place to assure the preservation of the integrity, authenticity and trustworthiness of the IPO's records over time. This may be accomplished by managing and protecting the electronic records from any loss, alteration, removal, or premature destruction. Since evidence of record tampering may not be as readily identifiable with electronic records, as it might be with paper records, it is even more important that the controls are in place to adequately protect and preserve the record.

5.4.10.1 Dossier Retention

For the scheduling of dossier retention, the complete electronic intellectual property dossier is considered to be a single entity, and, as such, the complete dossier can be managed and preserved using a single retention period.

5.4.10.2 Retention Period

As stated in Section 3.5 Record Retention, retention periods vary based on the status of the dossier, i.e., whether the patent is issued or abandoned, and also based on the period of time during which the application was filed. The determination of retention period for any given record is a matter of business need, legal requirement, or historical value. In other words, is the record needed to conduct business? If yes, it should be retained so long as that need exists. It is more obvious and less of an intuitive decision if a legal requirement or historical value determine retention period.

5.4.10.3 IPO Dossier Retention Schedules

Retention periods for intellectual property dossiers and records, whether electronic or paper-based, can be described as "event driven." For example, a retention period does not start until the prosecution of a pending patent application is completed—when a patent has been granted or

abandoned. The duration of the retention period is also dependent on whether the patent was granted or abandoned.

Suggested retention periods are provided below.

For patent dossiers showing the prosecution of applications for, and the granting of, a patent:

- Files include the original application, the patent drawing, and all materials relating to the prosecution of the application and subsequent actions by the IPO. Includes patent files for reissues.
- Closed (granted) patent dossiers selected by the IPO Director are permanent and transferred to archival storage after 40 years.
- All other closed (granted) patent dossiers are destroyed 40 years after closure.

For abandoned patent applications that do not result in the grant of a patent:

- Abandonment occurs when the applicant; fails to pay fees or submit documentation requested by the examiner within the allowed time; when claims made for the invention are not patentable or were previously patented; or when another applicant has filed an application for the same invention and can demonstrate an earlier date for the conception of the invention.
- Applications retained because they are referred to in another application that was granted are disposed of with the patent dossier in which it is cited.
- Abandoned applications are destroyed 23 years after closure.

5.4.10.4 Determine Retention Period

The business rules for determining the retention periods for intellectual property dossiers, records and associated metadata are relatively straight forward. The information system that captures the event that triggers the start of a retention period, such as the granting of a patent, should populate the appropriate Disposition metadata fields in the Dossier Profile.

- ❏ The information system that captures an event that triggers the beginning of a retention period should automatically determine the values for and populate the Disposition attributes “Instruction Code” and “Action Date” in the Dossier Profile with the appropriate retention information. --Instruction Code relates to an action, such as “disposal,” “transfer” or “hold.” --Action Date is the date that the “instruction” is to occur.

- ❏ Any modifications to disposition attributes should be restricted to personnel that are authorized by the IPO Records Officer or equivalent/designate, and should be tracked as part of the Use History Profile metadata for each dossier.
- ❏ Under authorization by the IPO Records Officer or equivalent/designate, provide for the capability to perform selective, mass updates to the Disposition attributes, such as when retention schedules are modified. For example, if the transfer period to permanent archival storage for granted patents was changed from 40 years to 30 years, allow for the mass change of the transfer "Action Date" to permanent archival storage for all affected dossiers.

5.4.10.5 Records to be Retained

- ❏ All documentary materials received or created by an IPO and working files that require preservation should be stored and the integrity preserved for the full retention period.
- ❏ Only one legal record copy, plus a vital records disaster backup copy, of each electronic dossier and record, including related metadata, should be retained.
- ❏ Ensure that the content and structure of the record is preserved so that a human readable display or print of the data can be reconstructed over the life of the record.
- ❏ Preserve the context (the circumstances under which the record was generated and stored) over the full retention period.
- ❏ Ensure that the chosen record format(s) will provide for record availability for the full retention life of the record, such as an industry standard or widely supported defacto standard file format.

5.4.10.6 Disposal

IPOs should preserve permanent records and promptly dispose of or retire temporary (non-permanent) records. The requirements for disposal of electronic dossiers and records, and associated metadata and backup copies are summarized below.

- ❏ Provide for access to and screening of the disposition action for dossiers by authorized personnel using the Disposition Action metadata in the Dossier Profile.
- ❏ Periodically produce electronic or hard copy lists for approval of dossiers whose action codes and dates indicate a pending disposal, produced either automatically based on predetermined criteria or upon manual initiation by authorized personnel.
- ❏ Provide for electronic approval of disposal actions, either using disposal lists or via manual entry by personnel authorized by the IPO Records Officer or equivalent/designate.
- ❏ Provide for a "confirmation" action prior to actually begin the disposal of dossiers and records.
- ❏ Provide for the electronic initiation of the disposal process for a specified list of dossiers, either using an information system or employing a separate utility specifically designed for disposal actions.
- ❏ Disposal audit trail—extract and save, for a to-be-specified period of time, a subset of the Use History Profile metadata documenting the disposal action taken for each dossier.

- ❑ Disposal of a dossier should include the complete contents of the dossier, all metadata associated with the dossier, and the vital records backup of the dossier.
- ❑ Disposal should include complete physical destruction (overwrite with unintelligible characters or disposal of a unit of media) of all dossier information, not just logical deletion of metadata and pointers to the dossier.
- ❑ A sampling of attempted accesses to disposed dossiers should be conducted in order to assure that the disposal action has been accurately and completely carried out.
- ❑ If a dossier or record is hyperlinked to another dossier, such as an abandoned patent dossier referenced by an issued patent, the abandoned dossier should be deleted only when it is eligible for deletion by the referencing dossier.

5.4.10.6.1 Disposal of Working Files

Working files are managed outside of the electronic dossier. However, they should be disposed of at the earliest opportunity. Working files should be deleted when the prosecution of the intellectual property application is completed (issuance, registration or abandonment), or earlier when they have no expected reference value and are not subject to a hold order.

- ❑ Provide a method for authorized users to dispose of working files that are not required for preservation as records.
- ❑ Deletion should include complete physical destruction (overwrite or disposal of a unit of media) of all dossier information, not just logical deletion of pointers to the dossier.
- ❑ Do not track the disposal activity of working files in the Use History Profile audit trail.

5.4.11 Migration

Electronic intellectual property dossiers must remain accessible and transferable despite changes in information technology. The content, structure and context of the electronic records must be able to be displayed, printed or otherwise reproduced, as they originally were captured, for the complete life cycle of the record.

5.4.11.1 Copy, Reformat and Transfer

Technology is certain to change and advance with a frequency that will require multiple migrations of electronic intellectual property dossiers and records over most retention periods. This will include the copying and reformatting of the dossiers to new media and/or the transfer to new information systems or new hardware and software.

- ❑ Develop a strategy and plan, and allocate future funding, to accomplish the copying, reformatting or transfer of electronic dossiers in a timely and accurate manner, since a break in the migration cycle may make the electronic records effectively irretrievable.
- ❑ The content, structure and context should be accurately preserved through copy, reformat and transfer migrations.
- ❑ Copy and reformat, if required, electronic dossiers at the time they are moved from one information system or one software and or hardware system environment to another.
- ❑ Reformat electronic dossiers, as required, when new storage devices or media are utilized.
- ❑ When using magnetic tape for either the primary or vital records backup, annually read a statistical sample of all reels of magnetic tape to identify any loss of data and to discover and correct causes of data loss.
- ❑ When using magnetic tape for either the primary or vital records backup, copy the electronic dossiers when the annual readability sample of magnetic tape discloses ten or more temporary or read errors.
- ❑ When using magnetic tape for either the primary or vital records backup, copy electronic dossiers every ten years for the primary and the vital records disaster backup.
- ❑ Transfer electronic intellectual property dossiers when the current software is upgraded or a new or upgraded electronic records file management system is installed.
- ❑ Ensure the reliability and integrity of reformatted, copied, and transferred electronic patent dossiers by employing a strict quality control procedure that may include bit/byte comparisons and comparisons of hash digests and Cyclical Redundancy Checks (CRC).
- ❑ Document fully all actions taken when reformatting, copying and transferring electronic intellectual property dossiers and include this information in the Use History Profile metadata associated with each dossier.
- ❑ Provide for vital records backup and disaster recovery by creating two copies of electronic intellectual property dossiers at the time of reformatting, copying, or transfer and store one copy at a separate geographical location.

5.4.11.2 Media Management

The manufacturer's specifications for both the *"pre-written"* media life and the *"post-written"* archival life must be followed.

- ❑ Determine and follow the media manufacturer's specifications for pre-written and post-written life. The *pre-written* media life is the time period from the date of manufacture until the date after which the *initial* writing of information to the media is not recommended. The *post-written* archival life is the period from the date of manufacture until the media should be copied to a new unit of media.
- ❑ Store the physical media in accordance with the manufacturer's environmental requirements for temperature and humidity controls.

- ❏ Periodically check the media for read-errors and read-error correction rates, even if the manufacturer's specifications for temperature and humidity controls are followed.
- ❏ When read-error correction rates meet or exceed the tolerance level or when the post-written archive life is imminent, copy the digital data to "fresh" media. This practice of copying the records before the media expires is an effective preservation technique only as long as the existing media is readable.

5.4.12 Transfer to Permanent Archival Storage

IPOs should establish policies and procedures to ensure that electronic records and their documentation are retained as long as needed.

- ❏ Transfer only the "selected" electronic dossiers and records and associated metadata to permanent archival storage, as designated by the Instruction Code and Action Code in the Dossier Profile.
- ❏ Generate all required forms to transfer records.
- ❏ Provide for the transfer of dossiers and records into the file formats and media types appropriate for permanent archival retention.
- ❏ Verify the quality of the records and associated metadata being transferred.
- ❏ Update the Transfer Profile metadata as a means of keeping a detailed audit trail of records transferred to permanent archival storage.

5.4.13 Records Hold

Records should be held (not deleted) when litigation, audit or investigation is imminent or would be foreseen by an average person.

- ❏ Identify which records are subject to the hold.
- ❏ Mark the metadata for the records as requiring a hold; identify the specific hold(s) that affect the record; and automate the release of the record, once the hold order is lifted (one record may be subject to multiple hold orders.)
- ❏ Only the IPO Records Officer or equivalent/designate should authorize selected personnel to create, change, or release a dossier hold order.

5.5 METADATA GUIDELINES

This section includes metadata guidelines for reformat, copy and transfer profiles.

5.5.1 Reformat Metadata Profile

This metadata profile is divided into input metadata and output metadata elements. The purpose of these metadata elements is to capture detailed information about the status of a dossier before and after reformatting in order to establish the foundation for its trustworthiness over time. At the time of the first reformatting, many of the input data elements most likely would be extracted from the Dossier or Dossier Records Metadata Profiles. Subsequent reformattings would link backward to the most recent processing, which could be reformatting, copying, or transferring, and extract the relevant metadata elements.

SAMPLE REFORMAT METADATA PROFILE (INPUT)

Date of Reformatting

Reformat Iteration Number

Dossier Identifier

Dossier Record Identifier (If Appropriate)

File Formats Used

- **TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)**
- **JPEG**
- **PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[3]**
- **XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)**
- **SGML**
- **Other (e.g., complex work units)**

Dossier or Record Byte Count

Record Authentication (If Used)

---

[3] PDF may be used if permissible for long-term archival storage by individual IPOs.

CRC

Hash Digest

Storage Media

Vendor

Type (e.g., RAID, 3480 or DLT Tape)

Product Name

Volume ID

Software Used in Reformatting

Name Of Product

Version Number

SAMPLE REFORMAT METADATA PROFILE (OUTPUT)

Date of Reformatting

Reformat Iteration Number

Dossier Identifier

Dossier Record Identifier (If Appropriate)

File Formats Used

TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)

JPEG

PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[4]

XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)

SGML

Other (e.g., complex work units)

Dossier or Record Byte Count

Record Authentication (If Used)

---

[4] PDF may be used if permissible for long-term archival storage by individual IPOs.

CRC

Hash Digest

Storage Media

Vendor

Type (e.g., RAID, 3480 or DLT Tape)

Product Name

Volume ID

Comparison

Byte Count

CRC

Hash Digest

Visual Inspection

Discrepancies (If Any)

Corrections (If Any and Explanations)

Supervisor Review

Physical Storage Location

Primary

Backup

5.5.2 Copy Metadata Profile

The purpose of the metadata elements in the copy format template is to capture detailed information about the status of a dossier before and after copying in order to establish the foundation for its trustworthiness over time. Consequently, the Copy Metadata Requirements are divided into two categories—Input and Output. Copying intellectual property dossiers can occur at the time after closure, after initial reformatting, or in conjunction with transfer (discussed later). At the time of the first copying many of the input data elements most likely would be extracted from the Creation-Use or Reformat Template. Subsequent copying would link backward to the most recent processing, which could be reformatting, copying, or transferring, and extract the relevant metadata elements.

SAMPLE COPY METADATA PROFILE (INPUT)

Date of Copying

Copy Iteration Number

Dossier Identifier

Dossier Record Identifier (If Appropriate)

File Formats Used

TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)

JPEG

PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[5]

XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)

SGML

Other (e.g., complex work units)

---

[5] PDF may be used if permissible for long-term archival storage by individual IPOs.

Dossier or Record Byte Count

Record Authentication (If Used)

CRC

Hash Digest

Storage Media

Vendor

Type (e.g., RAID, 3480 or DLT Tape)

Product Name

Volume ID

Software Used in Reformatting

Name Of Product

Version Number

SAMPLE COPY METADATA PROFILE (OUTPUT)

Date of Copying

Copy Iteration Number

Dossier Identifier

Dossier Record Identifier (If Appropriate)

File Formats Used

TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)

JPEG

PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[6]

XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)

SGML

---

[6] PDF may be used if permissible for long-term archival storage by individual IPOs.

Other (e.g., complex work units)

Dossier or Record Byte Count

Record Authentication (If Used)

CRC

Hash Digest

Storage Media

Vendor

Type (e.g., RAID, 3480 or DLT Tape)

Product Name

Volume ID

Comparison

Byte Count

CRC

Hash Digest

Visual Inspection

Discrepancies (If Any)

Corrections (If Any and Explanations)

Supervisor Review

Physical Storage Location

Primary

Backup

5.5.3 Transfer Metadata Requirements

Like the Reformat and Copy Metadata Requirements, the purpose of the Transfer Metadata Profile is to capture information that documents fully the actions taken in this activity that will help support the trustworthiness of electronic intellectual property dossiers despite changes in technology. The metadata elements in this profile capture detailed information about status of a dossier before and after transfer. It is likely that at the time of the first transfer many of the input data elements most likely would be extracted from either the Reformat or Copy Templates.

Subsequent transfers would link backward to the most recent processing, which could be reformatting, copying, or transferring, and extract the relevant metadata elements.

SAMPLE TRANSFER METADATA PROFILE (INPUT)

Date of Transfer

Transfer Iteration Number

Dossier Identifier

Dossier Record Identifier (If Appropriate)

File Formats Used

- **TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)**
- **JPEG**
- **PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[7]**
- **XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)**
- **SGML**
- **Other (e.g., complex work units)**

Dossier or Record Byte Count

Record Authentication (If Used)

- **CRC**
- **Hash Digest**

Storage Media

- **Vendor**
- **Type (e.g., RAID, 3480 or DLT Tape)**
- **Product Name**

---

[7] PDF may be used if permissible for long-term archival storage by individual IPOs.

Volume ID

Software Used in Transferring

Name Of Product

Version Number

SAMPLE TRANSFER METADATA PROFILE (OUTPUT)

Date of Transfer

Transfer Iteration Number

Dossier Identifier

Dossier Record Identifier (If Appropriate)

File Formats Used

TIFF (V6.0 with Group 4 Compression, Single Strip, Intel Encoded, 200, 300 or 400 dpi, Maximum Size A4 or Letter)

JPEG

PDF (Acrobat V3 Compatible, Non-Compressed Text, Un-encrypted Text, No Digital Signatures, No Embedded OLE Objects, All Fonts Embedded, Standard PS17 or Built from Adobe NM Fonts)[8]

XML (Character Set UTF-8 Encoded Unicode UCS-2 (ISO/IEC 10646:193) or ISO-2022-JP Encoded JIS-X0208 or for PCT Applications, Chinese GB2312 and Korean KSC 5601 also acceptable)

SGML

Other (e.g., complex work units)

Dossier or Record Byte Count

Record Authentication (If Used)

CRC

Hash Digest

Storage Media

Vendor

Type (e.g., RAID, 3480 or DLT Tape)

---

[8] PDF may be used if permissible for long-term archival storage by individual IPOs.

Product Name

Volume ID

Comparison

Byte Count

CRC

Hash Digest

Visual Inspection

Discrepancies (If Any)

Corrections (If Any and Explanations)

Supervisor Review

Physical Storage Location

Primary

Backup

Attachment 6. Acronyms

PKCS	Public Key Cryptographic Standard
PKI	Public Key Infrastructure
SDIF	SGML Document Interface Format
SGML	Standardised Generic Mark-up Language
XML	Extensible Mark-up Language
ePCT	Electronic PCT Application
…	

[Appendix II follows]

APPENDIX II

XML DTDs for IP Document Exchange

Management Summary

This document presents all the DTDs used for the electronic exchange of IP documents as defined in Annex F and Appendix I of the PCT Administrative Instructions.

It also contains details of the methodology adopted in drafting these DTDs.

Table of Contents

Attachment 1. DTDs for XML Documents

1.1 Strategy for further development of DTDs

Although limited at this stage to the initial electronic fling of a new PCT application, this specification is expected to expand in scope over time to include the subsequent formal exchanges between the parties involved. Below is a list of DTDs required for the initial phase of electronic submission of an electronic application. Other DTDs will be required for later phases of the processing of an electronic PCT application (ePCT).

While the immediate goal of this specification is to support ePCT applications, the Trilateral Offices intended to use it as the basis for their own national electronic applications for a variety of intellectual-property types and recommend that it would be the basis for an eventual WIPO standard for use by other Offices.

With that in mind, the DTDs created for ePCT will be constructed from so-called "architectural DTDs" that contain templates for element definitions and from which the Trilateral Offices and others can derive elements and DTDs for their needs in a consistent and compatible manner. The draft ePCT DTDs completed so far will be revised to take into account the architectural forms. The other DTDs that will eventually be required will also be based on the architectural DTDs. The XML name-space facility (XMLNS) will be used to implement architectures in the DTDs supporting this specification.

As an aid to creating the architectural DTDs and forms, a table will be constructed of the elements and structures in the sources of the proposed DTDs and those newly created for this specification, that will illustrate their relationships and their definitions.

1.2 DTD Inventory as of 1999 November 5

1. Electronic PCT application:
   public “-//wipo//dtd epct v0.2 1999-10-30//en”
   in epct-v02.dtd
2. Package header:
   public “-//wipo//dtd epct-ph v0.1 1999-10-30//en”
   In epct-package-header-v01.dtd
3. Confirmation certificate:
   public “-//wipo/dtd epct-cc v0.1 1999-10-30//en”
   In pcte-confirmation-certificate-v01.dtd
4. Ticket:
   public “-//wipo/dtd epct-ti v0.1 1999-10-30//en”
   in epct-ticket-header-v01.dtd

1.3 Future DTDs

5. Post-application, procedural and formal communications of various types
6. RO to IB
7. RO to ISA
8. IB to ISA
9. IB to IPEA
10. ISA to IB
11. IPEA to IB
12. IB to RO
13. DTDs required by other types of intellectual property

1.4 Reusable Components

Given the large number of IP Document Exchange DTDs, they will be built using a number of reusable standard components combined with transaction specific components.

The method for producing and organising these components shall be based on the XML structure called Architectural Forms.

The following components have been defined:

- Request
- Description
- Claims
- Abstract
- Drawings
- Sequence Listing
- PCT Receiving Office data
- Electronic Signature

Attachment 2. Standard Packages and Header Specifications

2.1 Package Header Object

The Header Object Data item is written in XML based on the following DTD:

```
<?xml version="1.0" >
<!—DTD for Package header object -->
<!--  To use this DTD define DOCTYPE declaration like
below. -->
<!-- <!DOCTYPE pkgheader SYSTEM "pkgheader.dtd" >
-->
<!ELEMENT pkgheader
     (ip,package,return-route?,file+)>
<!ELEMENT ip   EMPTY>
<!ATTLIST ip   type       CDATA #REQUIRED
          procedure CDATA #REQUIRED>
<!ELEMENT package   (#PCDATA)>
<!ATTLIST package id      CDATA #REQUIRED>
<!ELEMENT return-route   (#PCDATA)>
<!ELEMENT file      (#PCDATA)>
<!ATTLIST file type      CDATA #REQUIRED>
```

Tags are described below:

(1) **pkgheader**

Content : Route tag. Information goes under this tag.
Attribute : None
Repeat/Omission : 1 time/not possible

(2) **ip**

Content : None
Attribute : *type* (mandatory) Application type
1:Patent
2:Utility model
3:Design
4:Trademark
Expand as necessary
procedure (mandatory) Procedure type
1:Application
2:Amendment
3:Request for examination
Expand as necessary
Repeat/Omission : 1 time/not possible

(3) **package**

Content : Indicate package type
Attribute : *id* (mandatory) Package type
1: ticket request
2: ticket response
3: new application
4: confirmation certificate
Repeat/Omission : 1 time/not possible

(4) **return-route**

Content : Specify the reply method
Attribute : None
Repeat/Omission : 1 time/possible

(5) **file**

Content : Indicate Data Item file name
Attribute : *type* (mandatory) Data item type
1: Header Information
2: IP Documents
3: Ticket Information
Expand as necessary
Repeat/Omission: 1 time or more/ not possible

Here is an example of a Header Object:

```
<?xml version="1.0" >
<!DOCTYPE pkgheader SYSTEM "pkgheader.dtd" >
<pkgheader>
  <ip type="1" procedure="1" />
  <package id="3">New application</package>
  <file type="1">header.xml</file>
  <file type="2">ipdoc.dat</file>
   <file type="3">ticket.dat</file>
</pkgheader>
```

2.2 Confirmation Certificate

The XML version of the Confirmation Certificate must comply with the following DTD:

```
<?xml version="1.0" >
<!-- DTD for Confirmation Certificate -->
<!--  To use this DTD define DOCTYPE declaration like
below. -->
<!-- <!DOCTYPE confcertificate SYSTEM
"confcertificate.dtd" >      -->
<!ELEMENT confcertificate
     (application-number, applicants-reference, date-of-
receipt, action, message-digest, result-code, result-
message?)>
<!ELEMENT application-number   (#PCDATA)>
<!ELEMENT applicants-reference (#PCDATA)>
<!ELEMENT date-of-receipt      (#PCDATA)>
<!ELEMENT action               (#PCDATA)>
<!ELEMENT message-digest       (#PCDATA)>
<!ELEMENT result-code          (#PCDATA)>
<!ELEMENT result-message       (#PCDATA)>
```

Tags are described below:

(1) **confcertificate**
- Content : Route tag. Information goes under this tag.
- Attribute : None
- Repeat/Omission : 1 time/not possible

(2) **application-number**
- Content : Indicate the application number assigned
- Attribute : None
- Repeat/Omission : 1 time/not possible

(3) **applicants-reference**
- Content : The reference used by the applicant
- Attribute : None
- Repeat/Omission : 1 time/not possible

(4) **date-of-receipt**
- Content : Date assigned by the IP Office for receipt of the IP Document
- Attribute : None
- Repeat/Omission : 1 time/possible

(5) **action**
- Content : Type of IP Document sent (e.g. New US Application)
- Attribute : None
- Repeat/Omission : 1 time/possible

(6) **message-digest**
- Content : A hexadecimal representation of the message digest value
- Attribute : None
- Repeat/Omission : 1 time/not possible

(7) **result-code**
Content : A numeric representation of the result of the submission indicating success or failure(0=OK)
Attribute : None
Repeat/Omission : 1 time/not possible

(8) **result-message**
Content : A character string representation of the reason for any failure
Attribute : None
Repeat/Omission : 1 time/not possible

Here is an example of a Confirmation Certificate :

```
<?xml version="1.0" >
<!DOCTYPE confcertificate SYSTEM " confcertificate.dtd">
<confcertificate>
  <application-number>EP98200345</application-number>
  <applicants-reference>MyRef001</applicants-reference>
  <date-of-receipt>19990929</date-of-receipt>
  <action>NEW-EP-Application</action>
  <message digest>12:34:56:af:c3: 12:34:56:af:c3:
12:34:56:af:c3:12:34:56:af:c3</message digest>
  <result-code>0</result-code>
</confcertificate>
```

2.3 Ticket

The Ticket is used to maintain the information received by RO, and written in XML. Comply with the following DTD when writing the Ticket in XML.

```
<?xml version="1.0" >
<!-- DTD for Ticket -->
<!--  To use this DTD define DOCTYPE declaration like
below. -->
<!-- <!DOCTYPE ticket SYSTEM "ticket.dtd" >      -->
<!ELEMENT ticket
     (ticket-number, stamp-date, expire-date, message
digestfile)>
<!ELEMENT ticket-number  (#PCDATA)>
<!ELEMENT stamp-date     (#PCDATA)>
<!ELEMENT expire-date    (#PCDATA)>
<!ELEMENT message-digest           (#PCDATA)>
```

Tags are described below:

(1) **ticket**
Content : Route tag. Information goes under this tag.
Attribute : None
Repeat/Omission : 1 time/not possible

(2) **ticket-number**
- Content : Indicate the Ticket received number
- Attribute : None
- Repeat/Omission : 1 time/not possible

(3) **stamp-date**
- Content : Indicate the Ticket received date as CCYYMMDDHHMMSSTMZ where TMZ is the Time Zone
- Attribute : None
- Repeat/Omission : 1 time/not possible

(4) **expire-date**
- Content : Indicate the Ticket expiration date (Midnight 5 days later)
- Attribute : None
- Repeat/Omission : 1 time/possible

(5) **message-digest**
- Content : A hexadecimal representation of the has value
- Attribute : None
- Repeat/Omission : 1 time/not possible

Here is an example of a Ticket:

```
<?xml version="1.0" >
<!DOCTYPE ticket SYSTEM "ticket.dtd" >
<ticket>
  <ticket-number> 01234567 </ticket-number>
  <stamp-date>19990929112734CET</stamp-date>
  <expire-date>19991228</expire-date>
  <message-digest>12:34:56:af:c3: 12:34:56:af:c3:
12:34:56:af:c3:12:34:56:af:c3</message digest>
</ticket>
```

Attachment 3. DTD for New PCT Applications

```
<?xml version="1.0" encoding="ISO-8859-1"?>
˝
<!--Document type definition for international patent applications
˝
filed electronically under the patent cooperation treaty
ePCT v0.3, 1999 November 5
Reference this DTD as public "-//wipo//dtd epct v0.3 1999-11-05//en"

contacts:
EPO:
JPO:
USPTO: bruce.cox@uspto.gov
WIPO:

********** Revision History **********
Revised 1999 November 5, Bruce B. Cox, USPTO
..Modified name structure to accommodate both persons with organizational
...affiliation and organizations
..Converted all element names to lower case, following XML convention
Revised 1999 October 25, Bruce B. Cox, USPTO
..Moved electronic-signature to party structure
...and changed content model from (#PCDATA) to
...(date,((mark,signature-file?) | (signature-file,mark?))
first draft 1999 October 18, Bruce B. Cox, USPTO
..Based on elements found in
..1) "Proposed sgml specification for pct/easy," version 8, dated 1999 may 14
...(represented by b000 tags and b tags with wo suffix);
..2) st.32 and st.32/us/grant v1.8, 1999 august 26 (uspto's red book)
...(represented by the remaining b tags and most of the non-b tags in uppercase);
..3) USPTO's team DTD for a patent specification, u-specif.dtd, 1999 June 25
...(represented by tag names in lower case; some content models modified
...to remove parochial characteristics).
..Additional tags created for the purpose.
-->
<!DOCTYPE pct-international-application [
<!ELEMENT address  (pobox?,street,city,postcode?,country,telephone*,fax*,email*,
               b7110wo-formatted-address*) >

<!ELEMENT address-1  (#PCDATA) >

<!ELEMENT address-2  (#PCDATA) >

<!ELEMENT appendix-data  (heading,program-reference+) >

<!ELEMENT application-filing-date  (date) >

<!ELEMENT application-reference  (document-number,application-filing-date,country,
               kind-code) >

<!ELEMENT artwork  (drawing-reference-character*) >

<!--the agency or other body issuing the identification
number, or some other indication of the source or
meaning of the number. for example, "ssn" for u.s.
social security number; "uspto registration" for
agents and attorneys registered to practice before
the uspto.-->
```

```
<!ELEMENT authority  (#PCDATA) >

<!ELEMENT b000-wo  (b020wo-receiving-office?,b031wo-record-copy-at-ib-date,
               b040wo-signatory,b060wo-check-data?) >

<!--receiving office elements.
#usage:for use by receiving office only.-->
<!ELEMENT b020wo-receiving-office  (b021wo-drawings-not-received?,
               b022wo-corrected-date-of-receipt?,b023wo-date-of-receipt-pct-11-2?,
               b024wo-search-copy-delayed-until-fee-paid?) >

<!--drawings not received. (pct request no. 10-2)-->
<!ELEMENT b021wo-drawings-not-received   EMPTY  >

<!--corrected date of actual receipt due to later but timely received papers or
drawings completing the purported international application.-->
<!ELEMENT b022wo-corrected-date-of-receipt  (date) >

<!--date of timely receipt of the REQUIRED corrections under pct article 11(2).-->
<!ELEMENT b023wo-date-of-receipt-pct-11-2  (date) >

<!--flag: transmittal of search copy delayed until search fee is paid. (pct request
no. 10-6)
#usage:ELEMENT present when true, absent when false.-->
<!ELEMENT b024wo-search-copy-delayed-until-fee-paid   EMPTY  >

<!--date of receipt of the record copy by the ib.-->
<!ELEMENT b031wo-record-copy-at-ib-date  (date) >

<!--signatory information.-->
<!ELEMENT b040wo-signatory  (b041wo-applicant-agent+) >

<!--applicant / agent information.
#usage:repeat for each person signing the application.-->
<!ELEMENT b041wo-applicant-agent  (party,b042wo-signatory-capacity?) >

<!--capacity (role) of signatory. (pct request ix-x-3)-->
<!ELEMENT b042wo-signatory-capacity  (#PCDATA) >

<!--request.-->
<!ELEMENT b0601wo-request   EMPTY  >

<!ELEMENT b0603wo-claims  (claims) >

<!--abstract.-->
<!ELEMENT b0604wo-abstract  (heading,paragraph) >

<!--drawings.-->
<!ELEMENT b0605wo-drawings  (embedded-image+) >

<!--check data-->
<!ELEMENT b060wo-check-data  (b0601wo-request?,b0610wo-signed-power-of-attorney?,
               b0611wo-general-power-of-attorney?,b0612wo-lack-of-signature?,
               b0613wo-fee-calculation-sheet?,b0614wo-microorganisms?,
               b0615wo-sequence-listing?,b0617wo-lack-of-novelty?,
               b0618wo-translation?) >

<!--flag: separate signed power-of-attorney-->
<!ELEMENT b0610wo-signed-power-of-attorney   EMPTY  >
```

```
<!--flag: general power-of-attorney-->
<!ELEMENT b0611wo-general-power-of-attorney   EMPTY  >

<!--flag: statement explaining lack of signature-->
<!ELEMENT b0612wo-lack-of-signature   EMPTY  >

<!--flag: fee calculation sheet-->
<!ELEMENT b0613wo-fee-calculation-sheet   EMPTY  >

<!--flag: separate indications concerning deposited microorganisms-->
<!ELEMENT b0614wo-microorganisms   EMPTY  >

<!--flag: nucleotide and/or amino acid sequence listing and statement-->
<!ELEMENT b0615wo-sequence-listing   EMPTY  >

<!--flag: statement concerning non-prejudicial disclosures or exceptions to lack of
novelty-->
<!ELEMENT b0617wo-lack-of-novelty   EMPTY  >

<!--flag: translation of application-->
<!ELEMENT b0618wo-translation   EMPTY  >

<!ELEMENT b200-filing-info  (b210-ia-number,b211wo-country,
               b220-international-filing-date,b250-language) >

<!--international application number-->
<!ELEMENT b210-ia-number  (document-number) >

<!--receiving office specified by applicant using st.3 code.-->
<!ELEMENT b211wo-country  (country) >

<!ELEMENT b220-international-filing-date  (date) >

<!--language of filing (ISO 9)-->
<!ELEMENT b250-language  (#PCDATA) >

<!--priority information-->
<!ELEMENT b300-priority-info  (b310-priority-application-number,b320-priority-date,
               b330wo-pct-application,b330-office-of-filing,b340-paris-state,
               b345wo-priority-doc-requested,b346wo-priority-doc-attached) >

<!--priority application number-->
<!ELEMENT b310-priority-application-number  (document-number) >

<!--priority date-->
<!ELEMENT b320-priority-date  (date) >

<!--priority office-of-filing (WIPO Standard ST.3)-->
<!ELEMENT b330-office-of-filing  (#PCDATA) >

<!--flag: pct application-->
<!ELEMENT b330wo-pct-application   EMPTY  >

<!--paris convention state-->
<!ELEMENT b340-paris-state  (country) >

<!--flag: priority document requested from RO-->
<!ELEMENT b345wo-priority-doc-requested   EMPTY  >

<!--flag:priority document attached-->
```

```
<!ELEMENT b346wo-priority-doc-attached   EMPTY  >

<!ELEMENT b500-search-report  (b540-title,b560-search) >

<!ELEMENT b540-title  (b541-language,b542-title) >

<!--title language (ISO 639)-->
<!ELEMENT b541-language  (#PCDATA) >

<!ELEMENT b542-title  (#PCDATA) >

<!--citations, search report data-->
<!ELEMENT b560-search  (b567-place-of-search,b567wo-earlier-search-by,
               b568wo-earlier-search-date) >

<!--place of search (international search authority)-->
<!ELEMENT b567-place-of-search  (country) >

<!--earlier search country or regional office-->
<!ELEMENT b567wo-earlier-search-by  (country) >

<!--earlier search date-->
<!ELEMENT b568wo-earlier-search-date  (date) >

<!--specification and drawings-->
<!ELEMENT b590-drawings  (b598-figures-to-publish*) >

<!--figures (drawings)-->
<!ELEMENT b598-figures-to-publish  (#PCDATA) >

<!--applicant-->
<!ELEMENT b700-parties  (b710-applicants,b720-inventors?,b740-agents?) >

<!--main applicant-->
<!ELEMENT b710-applicants  (b711-applicant-inventor+) >

<!--applicant information-->
<!ELEMENT b711-applicant-inventor  (party,b716wo-capacity-of-legal-rep,
               b717wo-dead-inventor-name?,b711wo-applicant-is-inventor?,
               b715wo-applicant-for-listed-states?,(b712wo-applicant-for-all-states
|
               b713wo-not-us-applicant | b714wo-us-only-applicant)) >

<!--address formatted for printing as a label-->
<!ELEMENT b7110wo-formatted-address  (#PCDATA) >

<!--flag: applicant is also one of the inventors-->
<!ELEMENT b711wo-applicant-is-inventor   EMPTY  >

<!--flag: applicant for all designated states-->
<!ELEMENT b712wo-applicant-for-all-states   EMPTY  >

<!--flag: applicant for all designated states except us-->
<!ELEMENT b713wo-not-us-applicant   EMPTY  >

<!--flag: applicant for us only-->
<!ELEMENT b714wo-us-only-applicant   EMPTY  >

<!--applicant for listed national/regional states-->
<!ELEMENT b715wo-applicant-for-listed-states  (country*,b718wo-regional*) >
```

```
<!--legal representative capacity-->
<!ELEMENT b716wo-capacity-of-legal-rep  (#PCDATA) >

<!--flag: inventor is deceased-->
<!ELEMENT b717wo-dead-inventor-name   EMPTY  >

<!--regional filing-->
<!ELEMENT b718wo-regional  (b719wo-region,country) >

<!--regional office code (WIPO Standard ST.3)-->
<!ELEMENT b719wo-region  (#PCDATA) >

<!--inventor-->
<!ELEMENT b720-inventors  (b721-inventor+) >

<!--inventor name, address, residence-->
<!ELEMENT b721-inventor  (party,b726wo-dead-inventor?,(
               b712wo-applicant-for-all-states | b713wo-not-us-applicant |
               b714wo-us-only-applicant)) >

<!--flag: inventor is deceased (relevant to u.s.).-->
<!ELEMENT b726wo-dead-inventor   EMPTY  >

<!--agent-->
<!ELEMENT b740-agents  (b741-agent+) >

<!--agent-->
<!ELEMENT b741-agent  (party,b742wo-common-representative,b743wo-mailing-address?) >

<!--flag: acting as common representative-->
<!ELEMENT b742wo-common-representative   EMPTY  >

<!--correspondence address-->
<!ELEMENT b743wo-mailing-address  (address) >

<!ELEMENT background-of-invention  (heading,(paragraph | list | section | sequence |
               chemistry | math | table)+) >

<!ELEMENT biological-deposit  (deposit-date,deposit-term,depository,
               deposit-accession-number,deposit-description) >

<!ELEMENT biological-deposit-data  (heading?,biological-deposit+) >

<!ELEMENT brief-description-of-drawings  (heading,(paragraph | figure)+) >

<!ELEMENT chemistry   EMPTY  >

<!--city-->
<!ELEMENT city  (#PCDATA) >

<!ELEMENT claim  (preamble-claim,(claim-step+ | paragraph-claim)) >

<!ELEMENT claim-step  (claim-step | paragraph-claim)+ >

<!ELEMENT claims  (heading,claim+) >

<!ELEMENT colspec   EMPTY  >

<!ELEMENT copyright  (#PCDATA) >
```

```
<!--WIPO Standard ST.3 code for country or other administrative unit.-->
<!ELEMENT country  (#PCDATA) >

<!ELEMENT cross-reference  (#PCDATA) >

<!ELEMENT custom-character   EMPTY  >

<!--yyyymmdd-->
<!ELEMENT date  (#PCDATA) >

<!ELEMENT date-signed  (date) >

<!ELEMENT dependent-claim-reference  (#PCDATA) >

<!ELEMENT deposit-accession-number  (#PCDATA) >

<!ELEMENT deposit-date  (#PCDATA) >

<!ELEMENT deposit-description  (#PCDATA) >

<!ELEMENT deposit-reference  (#PCDATA) >

<!ELEMENT deposit-term  (#PCDATA) >

<!ELEMENT depository  (address-1,address-2?,city,state,postcode,country,telephone*,
               email*,fax*) >

<!--description-->
<!ELEMENT detailed-description  (heading,biological-deposit-data?,(paragraph | list
|
               section | sequence | chemistry | math | table)+) >

<!--document number-->
<!ELEMENT document-number  (#PCDATA) >

<!ELEMENT drawing-reference-character  (#PCDATA) >

<!--the character string which
singifies the signor’s intention to sign.-->
<!ELEMENT electronic-signature  (date-signed,place-signed,((signature-mark,
               signature-file?) | (signature-file,signature-mark?) |
               use-digital-signature)) >

<!--email address-->
<!ELEMENT email  (#PCDATA) >

<!--embedded image file reference.-->
<!ELEMENT embedded-image   EMPTY  >

<!ELEMENT emphasis  (#PCDATA | superscript | subscript)* >

<!ELEMENT entry  (#PCDATA | emphasis)* >

<!ELEMENT family-name  (#PCDATA) >

<!--telefacsimile number-->
<!ELEMENT fax  (#PCDATA) >

<!ELEMENT fiche-count  (#PCDATA) >
```

```
<!ELEMENT fiche-frame  (#PCDATA) >

<!ELEMENT figure  (copyright?,(artwork | sequence | chemistry | math | table)) >

<!ELEMENT given-name  (#PCDATA) >

<!ELEMENT heading  (#PCDATA) >

<!ELEMENT hybrid-claim-reference  (#PCDATA) >

<!--idendification number of an individual or an organization.-->
<!ELEMENT id-number  (authority,number) >

<!ELEMENT kind-code  (#PCDATA) >

<!ELEMENT list  (list-item+) >

<!ELEMENT list-item  (list-item | paragraph | sequence | chemistry | math | table)+
>

<!ELEMENT middle-name  (#PCDATA) >

<!ELEMENT name  ((middle-name?,family-name,name-prefix?,given-name,name-suffix?,
               ref-number?,id-number?,organization-name?,title-at-organization?) | (
               name-prefix?,given-name,middle-name?,family-name,name-suffix?,
               ref-number?,organization-name,title-at-organization?,id-number?)) >

<!ELEMENT name-prefix  (#PCDATA) >

<!ELEMENT name-suffix  (#PCDATA) >

<!--state of nationality-->
<!ELEMENT nationality  (country) >

<!--the identification number.-->
<!ELEMENT number  (#PCDATA) >

<!ELEMENT open-literature-reference  (#PCDATA) >

<!ELEMENT organization-name  (#PCDATA) >

<!ELEMENT paragraph  (#PCDATA | custom-character | emphasis | subscript |
               superscript | application-reference | patent-reference |
               open-literature-reference | deposit-reference | cross-reference |
               sequence-fragment | sequence | chemistry | math | table)* >

<!ELEMENT paragraph-claim  (#PCDATA | hybrid-claim-reference | cross-reference)* >

<!ELEMENT party  (name,electronic-signature?,address?,nationality?,residence?) >

<!ELEMENT patent-reference  (document-number,publication-date,country,kind-code) >

<!ELEMENT pct-international-application  (sdobi-bibliographic-data,sdoab-abstract?,
               sdode-description,sdocl-claims,sdodr-drawings?) >

<!ELEMENT place-signed  (city?) >

<!--post office box-->
<!ELEMENT pobox  (#PCDATA) >
```

```
<!--postal code-->
<!ELEMENT postcode  (#PCDATA) >

<!ELEMENT preamble-claim  (#PCDATA | dependent-claim-reference | cross-reference)* >

<!ELEMENT program-listing  (#PCDATA) >

<!ELEMENT program-listing-deposit  (heading?,program-listing) >

<!ELEMENT program-reference  (fiche-count,fiche-frame) >

<!ELEMENT publication-date  (date) >

<!--reference number provided by the filer (docket number, etc.).-->
<!ELEMENT ref-number  (#PCDATA) >

<!ELEMENT related-documents  (heading,paragraph+) >

<!--state of residence-->
<!ELEMENT residence  (country) >

<!ELEMENT row  (entry+) >

<!--subdocument abstract.-->
<!ELEMENT sdoab-abstract  (b0604wo-abstract) >

<!--subdocument: bib data-->
<!ELEMENT sdobi-bibliographic-data  (b000-wo,b200-filing-info,b300-priority-info,
               b500-search-report,b590-drawings,b700-parties) >

<!--subdocument claims-->
<!ELEMENT sdocl-claims  (b0603wo-claims) >

<!--specification or description of the invention.-->
<!ELEMENT sdode-description  (title-of-invention,related-documents?,
               background-of-invention?,summary-of-invention?,
               brief-description-of-drawings?,detailed-description,
               program-listing-deposit?,appendix-data?,sequence-listing?) >

<!--subdocument drawings.-->
<!ELEMENT sdodr-drawings  (b0605wo-drawings) >

<!ELEMENT section  (heading,(paragraph | list | section | sequence | chemistry |
               math | table)+) >

<!ELEMENT sequence  (sequence-identification) >

<!ELEMENT sequence-fragment  (#PCDATA | sequence-identification)* >

<!ELEMENT sequence-identification  (#PCDATA) >

<!ELEMENT sequence-listing  (heading,sequence+) >

<!--an external file containing either:
an image of a page to which the party's signature
has been afFIXED,
or an image of the party's signature alone.-->
<!ELEMENT signature-file   EMPTY  >
```

```
<!--a text string representing the party's mark.-->
<!ELEMENT signature-mark  (#PCDATA) >

<!ELEMENT state  (#PCDATA) >

<!--street-->
<!ELEMENT street  (#PCDATA) >

<!ELEMENT subscript  (#PCDATA) >

<!ELEMENT summary-of-invention  (heading,(paragraph | list | section | sequence |
              chemistry | math | table)+) >

<!ELEMENT superscript  (#PCDATA) >

<!ELEMENT table  (heading,tgroup+) >

<!ELEMENT tbody  (row+) >

<!--telephone number-->
<!ELEMENT telephone  (#PCDATA) >

<!ELEMENT tgroup  (colspec*,thead?,tbody) >

<!ELEMENT thead  (row+) >

<!ELEMENT title-at-organization  (#PCDATA) >

<!ELEMENT title-of-invention  (#PCDATA | emphasis | superscript | subscript)* >

<!--Flag: a digital signature is used.-->
<!ELEMENT use-digital-signature   EMPTY  >

<!ATTLIST address-1
              text  CDATA    #FIXED "[address: ]" >

<!ATTLIST address-2
              text  CDATA    #FIXED "[address: ]" >

<!ATTLIST application-reference
              hytime  CDATA    #FIXED "hylink" >

<!ATTLIST artwork
              id  ID    #REQUIRED
              source  ENTITY    #REQUIRED
              size (a4 | letter)  "letter"
              color (black-and-white | color)   #FIXED "black-and-white"
               >

<!ATTLIST biological-deposit
              id  ID    #REQUIRED
              type (original | replacement | supplemental)   #IMPLIED
              corroboration  ENTITY    #IMPLIED
              viability  ENTITY    #IMPLIED  >

<!ATTLIST biological-deposit-data
              id  ID    #REQUIRED  >

<!ATTLIST chemistry
              id  ID    #IMPLIED
```

```
              source  ENTITY    #REQUIRED
              display (display)   #IMPLIED  >

<!ATTLIST claim
              id  ID    #REQUIRED  >

<!ATTLIST claim-step
              class  CDATA    #FIXED "paragraph" >

<!ATTLIST colspec
              colnum  CDATA    #IMPLIED
              colname  CDATA    #IMPLIED
              colwidth  CDATA    #IMPLIED
              colsep  CDATA    #IMPLIED
              rowsep  CDATA    #IMPLIED
              align (left | right | center | justify | char)   #IMPLIED
              char  CDATA    #IMPLIED
              charoff  CDATA    #IMPLIED  >

<!ATTLIST copyright
              class  CDATA    #FIXED "primitive" >

<!ATTLIST cross-reference
              hytime  CDATA    #FIXED "hylink"
              anchrole  CDATA    #FIXED "mention target"
              mention  IDREF    #IMPLIED
              target  IDREFS    #REQUIRED
              linktrav  CDATA    #FIXED "e e"
              anchcstr  CDATA    #FIXED "#self #REQUIRED"
              EMPTYanc  CDATA    #FIXED "error error"
              reftype  CDATA    #FIXED
              "target (sequence-listing           |chemistry           |math
|table           "

              ireftype  CDATA    #FIXED
              "target (sequence-listing           |chemistry           |math
|table           "
               >

<!ATTLIST custom-character
              source  ENTITY    #IMPLIED  >

<!ATTLIST dependent-claim-reference
              hytime  CDATA    #FIXED "hylink"
              class  CDATA    #FIXED "reference"
              depends_on  IDREFS    #REQUIRED
              data-format  CDATA    #FIXED "claim-number" >

<!ATTLIST deposit-accession-number
              text  CDATA    #FIXED "[deposit accession number: ]"
               >

<!ATTLIST deposit-date
              text  CDATA    #FIXED "[deposit date: ]" >

<!ATTLIST deposit-description
              text  CDATA    #FIXED "[deposit description]" >

<!ATTLIST deposit-reference
              hytime  CDATA    #FIXED "hylink" >
```

```
<!ATTLIST deposit-term
               text  CDATA    #FIXED "[deposit term: ]" >

<!ATTLIST drawing-reference-character
               id  ID    #IMPLIED  >

<!--ti, type of image:
ad = abstract drawing
cf = chemical formulae
ci = clipped image
cp = computer program listings
dn = dna sequences
dr = drawings
ff = undefined characters
fg = figures
gr = graphs
mf = mathematical formulae
pa = full-page facsimile image
ph = photograph
sr = search report form
tb = table or tabular data
tx = text character [deprecated in us documents]
ui = undefined image [deprecated in us documents]-->
<!ATTLIST embedded-image
               id  ID    #REQUIRED
               he  NMTOKEN    #IMPLIED
               wi  NMTOKEN    #IMPLIED
               file  ENTITY    #REQUIRED
               lx  NMTOKEN    #IMPLIED
               ly  NMTOKEN    #IMPLIED
               imf (st33 | tiff)   #IMPLIED
               ti (ad | cf | ci | cp | dn | dr | fg | ff | gr | mf | pa | ph | sr |
               tb | tx | ui)   #IMPLIED  >

<!ATTLIST entry
               colname  NMTOKEN    #IMPLIED
               namest  NMTOKEN    #IMPLIED
               nameend  NMTOKEN    #IMPLIED
               morerows  CDATA    #IMPLIED
               colsep  CDATA    #IMPLIED
               rowsep  CDATA    #IMPLIED
               align (left | right | center | justify | char)   #IMPLIED
               char  CDATA    #IMPLIED
               charoff  CDATA    #IMPLIED
               valign (top | middle | bottom)   #IMPLIED  >

<!ATTLIST family-name
               class  CDATA    #FIXED "primitive"
               label  CDATA    #FIXED "[family name:] " >

<!ATTLIST figure
               id  ID    #REQUIRED
               view (alternate | exploded | modified | partial | sectional)
                #IMPLIED  >

<!ATTLIST heading
               id  ID    #REQUIRED  >

<!ATTLIST hybrid-claim-reference
```

```
              hytime  CDATA    #FIXED "hylink"
              depends_on  IDREFS    #REQUIRED
              class  CDATA    #FIXED "reference"
              data-format  CDATA    #FIXED "claim-number" >

<!ATTLIST list
              id  ID    #REQUIRED  >

<!ATTLIST math
              source  ENTITY    #IMPLIED
              class  CDATA    #IMPLIED
              style  CDATA    #IMPLIED
              id  ID    #IMPLIED
              other  CDATA    #IMPLIED
              macros  CDATA    #IMPLIED
              mode  CDATA    #IMPLIED
              type  CDATA    #IMPLIED
              name  CDATA    #IMPLIED
              height  CDATA    #IMPLIED
              width  CDATA    #IMPLIED
              baseline  CDATA    #IMPLIED
              overflow (scroll | elide | truncate | scale)  "scroll"

              altimg  CDATA    #IMPLIED
              alttext  CDATA    #IMPLIED  >

<!ATTLIST middle-name
              class  CDATA    #FIXED "primitive"
              label  CDATA    #FIXED "[middle initial:] " >

<!ATTLIST name-suffix
              class  CDATA    #FIXED "primitive"
              label  CDATA    #FIXED "[name suffix:] " >

<!ATTLIST organization-name
              class  CDATA    #FIXED "primitive"
              label  CDATA    #FIXED "[organization name:] " >

<!ATTLIST paragraph
              id  ID    #REQUIRED  >

<!ATTLIST paragraph-claim
              class  CDATA    #FIXED "paragraph" >

<!ATTLIST patent-reference
              hytime  CDATA    #FIXED "hylink"
              data-format  CDATA    #FIXED "patent-number" >

<!ATTLIST pct-international-application
              dtdv  CDATA    #REQUIRED  >

<!ATTLIST program-listing
              id  ID    #REQUIRED
              xml:space (default | preserve)   #FIXED "preserve"
               >

<!ATTLIST program-reference
              id  ID    #REQUIRED  >

<!ATTLIST row
```

```
              rowsep  CDATA    #IMPLIED
              valign (top | middle | bottom)   #IMPLIED  >

<!ATTLIST section
              id  ID    #REQUIRED  >

<!ATTLIST sequence
              id  ID    #REQUIRED
              submission  ENTITY    #IMPLIED  >

<!--the type of signature file.
digital = digital certificate
page = image of page on which the signature has been written
scrawl = image of the signature only-->
<!ATTLIST signature-file
              type (digital | page | scrawl)   #REQUIRED  >

<!ATTLIST state
              text  CDATA    #FIXED "[state: ]" >

<!ATTLIST table
              frame (top | bottom | topbot | all | sides | none)   #IMPLIED
              colsep  CDATA    #IMPLIED
              rowsep  CDATA    #IMPLIED
              pgwide  CDATA    #IMPLIED
              id  ID    #REQUIRED  >

<!ATTLIST tbody
              valign (top | middle | bottom)   #IMPLIED  >

<!ATTLIST tgroup
              cols  CDATA    #REQUIRED
              colsep  CDATA    #IMPLIED
              rowsep  CDATA    #IMPLIED
              align (left | right | center | justify | char)   #IMPLIED  >

<!ATTLIST thead
              valign (top | middle | bottom)   #IMPLIED  >

<!ATTLIST title-at-organization
              class  CDATA    #FIXED "primitive"
              label  CDATA    #FIXED "[title:] " >

<!ATTLIST title-of-invention
              id  ID    #REQUIRED  >

]>
```

[Annex 6 follows]