

# **Indexing of Grantor/Grantee Names By Land Recording Offices**

## **A Review and Recommendations**

Approved by the Executive Board of the Property Records Industry Joint Task Force on February 28, 2002.

## Table of Contents

<b>1. Introduction.....</b>	<b>.....</b>
<b>2. Executive Summary .....</b>	<b>.....</b>
<b>3. Objectives of a Grantor/Grantee Notice/Indexing System.....</b>	<b>.....</b>
<b>4. Name Indexing System Problems and Historical Solutions .....</b>	<b>.....</b>
<b>5. Keying Process.....</b>	<b>.....</b>
<b>6. Indexing Process.....</b>	<b>.....</b>
<b>7. Search Process .....</b>	<b>.....</b>
<b>8. Indexing-Related Rules .....</b>	<b>.....</b>
<b>9. Human Name Rules .....</b>	<b>.....</b>
<b>10. Non-Human Name Rules.....</b>	<b>.....</b>
<b>11. All Name Rules.....</b>	<b>.....</b>
<b>12. Special Name Rules.....</b>	<b>.....</b>
<b>13. Indexing from Electronic Recordings .....</b>	<b>.....</b>
<b>14. Other Forms of Indexing.....</b>	<b>.....</b>

## Acknowledgments and Disclaimer

This paper was prepared by a subcommittee of the Standards Committee. The subcommittee was chaired by Carl R. Ernst.

More than 150 participants from the public and private sectors involved in the process of recording real estate transactions were part of the group that reviewed the information presented in this paper, both by email and in person at the March 2001, July 2001 and February 2002 meetings of the Task Force. We thank everyone for their participation and input.

None of the recommendations or opinions stated in this paper should be construed as representing the official position of any individual member, association or company involved in the review process.

# 1. Introduction

Indexing is one of the seven major issues that are of current interest to both the private and public sectors of the land recording industry, as identified by a vote of the Property Records Industry Joint Task Force at the July 2000 meeting in Charlotte, NC.

The Executive Board of the Task Force made a decision to attack the indexing issue in advance of the top item on the list, electronic recording, because indexing is an issue that impacts all recording operations, including electronic recording, which subject is now being worked on by another committee of the Task Force.

There are literally 3,600 name indexing systems in place in recording offices in the US today. Each system is based on series of rules, established and changed over decades since computer systems were available, for deciding what names to index, how to key in names, how to sort the names into an indexable order, how to allow users to search in the index, and how to present information from the index to users. These systems have a lot in common, of course, since they are all based on the same kinds of documents and include the same kinds of names.

However, as the review of existing indexing manuals in this paper shows, there are literally hundreds of possible variations in how an index is put together based on decisions about the scores of individual characteristics of human and organization names that appear in documents.

This paper takes a particular approach to the subject. It points out why name indexing systems in land recording offices look and act the way they do today based on a number of historical technological factors. It also presents a primer about the technologies that are available today to improve how name indexing systems work. It reviews the detailed rules as presented in available recording office indexing standards manuals. Then, based on a comparison of these rules to the technologies available today, the paper indicates recommended rules and methodologies for bringing name indexing systems into closer conformance with the seven fundamental objectives of the indexing system summarized in Chapter 3.

This paper is designed not only to provide recording offices with recommended standards, but also to educate private industry as to the best practices in the use of names in documents, so that document preparers will enter names in ways that can be more accurately indexed in land recording and title plant systems.

## Background

The purpose of recording instruments and notices applicable to real property transactions or encumbrances is to provide an inexpensive legal means to protect the interests of owners and encumbrancers by **notifying** all creditors, subsequent purchasers, and others with an economic interest in a parcel of real property about ownership of and encumbrances against that property.

The mission of a recording office with respect to real property transactions is to provide a trustworthy place where instruments and notices concerning ownership of and encumbrances against real property **can be located** by all creditors, subsequent purchasers, and others with an economic interest the property.

## Indexing—The Key to Real Estate Records

The means by which land recording offices achieve their mission is not only by maintaining original copies of all real estate-related documents required to be recorded in secure, non-alterable form, but also by providing searchers with a means of accessing these documents through indexes to the documents. A primary index, whatever its local designation, contains the names of all grantors and grantees included on the documents recorded, sorted into an orderly listing. This name index is the type of index that is the subject of this paper.

Other types of indexes exist in some states, in some recording offices in other states, and in private company computer databases, such as title plants. Although these indexes are important ways to access real property information, electronic recording is not the subject of this paper.

Name indexing is considered an art as much as a science by most recording offices. The Nebraska manual calls it “the arbitrary art of indexing.”

One of the purposes of this paper is to suggest how recording offices can reconsider their indexing policies and practices in ways that are more scientific and less arbitrary and artful, utilizing computer technology and programming techniques available today.

### **History of this Paper**

This project was approved in July 2000, and the first draft of this paper was issued in December of that same year. As a result of the analysis presented in this paper, the author developed a form of model name indexing system that can be used for a model grantor/grantee indexing system that can be described succinctly, is highly efficient both in keying and searching, and will work in any recording office. This system is the one recommended in this paper, along with the related instructions to users who search the index.

This model system was presented to the Task Force in March 2001. At that meeting the more than 60 recorders agreed that the system was workable, and suggested some changes that have been incorporated in the model system. The model system is outlined in Appendix 12 and the search instructions that accompany it are given in Appendix 13.

Recorders were still skeptical that their old indexing systems could be converted to operate with the model system in a manner that would allow searchers to find index entries that had been keyed according to a variety of rules, including handling of abbreviations, apostrophes, and the many other manual keying rules followed in the past by recording offices. Therefore, between March and July 2001, the author analyzed an old-style index of 120,000 entries from Trinity County, CA. This index is representative of many of the types of problems that other recording office indexes exhibit. The author analyzed the problem entries in the database and presented his findings to the Task Force at its summer meeting in July 2001.

## 2. Executive Summary

A fundamental objective of land recording offices is to provide access to publicly recorded documents related to real estate transactions through an index of the names of grantors and grantees listed in those documents.

The fundamental objectives of a name indexing system are:

- (1) Timeliness
- (2) Accuracy
- (3) Completeness
- (4) Consistency
- (5) Accessibility
- (6) Ease of Use
- (7) Readability

### Legacy Name Indexing Systems

These objectives are realized through a combination of intake practices and technology. In the past technology has been unavailable or underutilized, and the resulting practices have been time-consuming and inaccurate both for key entry operators and for searchers of the resulting name index. Specific factors supporting this statement are as follows:

- (1) Entries in a name index have depended on what the key entry operator decided to key in.
- (2) Key entry operator rules have been either
  - (a) informal, unwritten, traditional practices, or
  - (b) voluminous written instructions that are difficult to remember
- (3) Key entry operator rules have changed over past decades, generating inconsistent name index entries
- (4) Computer name field lengths have been inadequate to hold longer names.
- (5) Name abbreviations, sometimes changing over time, have been used extensively.
- (6) Human names have not been distinguished from organization names.
- (7) Search instructions have been too skimpy, and often nonexistent.

Appendixes 1-10 compare the approximately 60 rules that recording offices have developed to implement the type of key-operator-based name indexing system used today, which have the kinds of problems listed above, and which we call “legacy” name indexing systems.

### Modern Name Indexing System

Modern computer systems and programming can resolve most of the issues standing in the way of more accurate, more consistent, and more useable name indexing systems are available today at reasonable cost.

A model name indexing system is outlined in Appendix 12 and the search instructions that go along with the model system are presented in Appendix 13. The model indexing system solves many indexing problems and is practical and inexpensive to implement.

However, the system is not a panacea. Three related outstanding issues still need to be addressed in order to ensure the greatest possible accuracy and completeness in future name indexing systems in land recording offices:

## 1. Conversion Activities

It is generally recognized that legacy name indexing systems may contain many index entries that are incompatible with a modern name indexing system. In larger recording offices these records may be counted in the millions. As we illustrated in the test conversion of index records in Trinity County, CA, it is possible to convert the vast majority of legacy index records to the format of a modern indexing system. The remaining anomalies can then be noted in the search instructions to the new system.

**However, even if the legacy index entries are not converted, we recommend the immediate implementation of the model name indexing system because it will prevent any further erosion of the accuracy and completeness of the existing indexing system.** Otherwise, the legacy indexing systems in operation today will continue to be prone to changes in rules and inconsistent keying according to overly complex rules. A combination of computer search logic and search instructions will assist searchers in finding entries in both the old and new formats.

## 2. Preparer Responsibility

Document preparers are largely uninformed about the need to print complete, correct, consistent names in real estate documents to be recorded. As a consequence these documents contain printed names that are incomplete, misspelled, inconsistently given within the same document, and run together, and that do not even agree with signatures in many instances. Preparers need to be educated, and legislation may be necessary, to assure that names and signatures are complete and consistent as given in documents to be recorded.

## 3. Electronic Recording

This paper does not deal with the detailed design and implementation issues that must also be resolved in order to create a successful electronic-recording-system/name-indexing-system interface. These issues include, for example, the design of name data fields, edits of entries presented in those fields, education of electronic document preparers to utilize the data fields properly, and post-audits of names electronically added to the name index system. The electronic recording committee of the Task Force will be addressing these issues. Appendix 11 points out characteristics of each of the rules listed in Appendixes 1-10 that may have an impact on automated electronic recording/name indexing interfaces.

### **3. Objectives of a Grantor/Grantor Notice/Indexing System**

We have identified seven objectives that every recording office—and for that matter, title plant—should strive to meet in developing and maintaining its grantor/grantee indexing system.<sup>1</sup> These should not only be the objectives of each individual recording office, but also of recording offices throughout a state, and, eventually, throughout the United States because there is today a national constituency of document submitters and searchers in the majority of recording offices.

#### **1. Timeliness**

An index to provide public notice of recordings is useful only to the extent that it current. Recording offices should strive to have each document received for recording indexed by the end of the same business day as the document is received..

#### **2. Accuracy**

An index to provide public notice of recordings is useful only to the extent that it accurately replicates the grantor and grantee names that appear in documents. Errors and inconsistencies in keying party names can create breaks in the chain of title and hidden liens.

#### **3. Completeness**

An index to provide public notice of recordings is useful only to the extent that it is complete. “Complete” means both that all document recorded are indexed and that all grantors and grantees in documents are indexed.

#### **4. Consistency**

An index to provide public notice of recordings is useful only to the extent that the indexing of names is consistent. Names should be keyed and entered into the index in a consistent fashion. For example, entry of “national” as either “natl” or “national” can create hidden records and liens. Consistency rules are some of the most difficult to develop, yet they are a necessity.

#### **5. Accessibility**

An index to provide public notice of recordings is useful only to the extent that the index is accessible, through on-site terminals, internet web sites, and sales of bulk data to title plants and private vendors who program different ways of searching and using the information.

#### **6. Ease of Use**

An index to provide public notice of recordings is useful only to the extent that access to it is logical and easy for the user to access. The grantor/grantee index should be divided and ordered alphabetically in such a way that searchers can find what they are looking for with a minimum of special instructions.

#### **7. Readability**

An index to provide public notice of recordings is useful only to the extent that the information provided in a search of the index is readable. This involves not only placing in the index information necessary to identify the type and location of the actual document, but, using modern technologies, also making images of documents readily available to the public.

---

<sup>1</sup> These objectives probably apply equally well to tract indexing systems.

## 4. Name Indexing System Problems and Historical Solutions

The design of name indexes that land recording offices use today are the result of a number of historical factors, especially including the ways document preparers enter names on documents and the ways computer systems have been designed in the past to handle keying, indexing and searching chores.

### A. Divining the Names of the Parties

There is no statute that requires preparers of documents to enter different forms of names in a clear, concise way. The closest that any state statutes come to such requirements are such simple obligations as to print the name of the signers beneath their signatures. As a result of this shortcoming in the statutes, the names placed on documents frequently do not take into account that the recording office is supposed to put the names on an index in a manner that allows future searchers to find the document. Some of the common document preparation and signing errors are,

- (1) Multiple names strung together in unusable formats (Mr. And Mrs. John Jones)
- (2) Party name spelled differently in different places on the document, and
- (3) Signature different from printed name.

Assuming such documents are not rejected for these errors, how should such party names be indexed?

In these and other instances, land recording offices are left to divine the actual names of the parties that should be indexed. In the Nebraska indexing guidelines, for example, the following rules are set out:

- (1) If the signature is legible, it governs.
- (2) If the printed version of a human name spells out a name that is only initials in the signature, and the printed name is not in conflict with the signature, the printed or typed name governs.
- (3) When one or more words in the signature is legible, it governs for that word.
- (4) If none of the words in the signature are legible, the printed or typed name governs.

The Nebraska guidelines then remind the recording offices that they may reject any document on which the name is not typed or legibly printed below the signatures or if the signatures themselves are not legible!

**Without statutory guidance to document preparers, land recording offices will continue to deal with problems of divining the proper name to put into their name indexes.** This paper does not purport to address this problem. Some decisions will remain subjective. However, this paper will discuss methods available today to ameliorate the shortcomings of indexing systems that are due to these kinds of problems.

### B. Legacy Name Indexing Systems

Most, if not all, the name indexing systems extant today are legacy systems that were originally developed earlier in the history of data processing, in the 1960's-1980's.

These legacy indexing systems do not take advantage of the explosion in technology and implosion in data processing costs that have occurred during the past 20 years, and that have accelerated during the past 10 years.

These legacy systems were developed within the following two constraints that were in effect in the past:

- (1) Costly storage.
- (2) Limited programming power (and program design knowledge).

Because disk storage was very expensive a decade or two ago, system designers got into the habit of keeping name field lengths to a minimum in order to save disk space, which meant to save investment



dollars. Because of this, some name indexing systems even today allocate less than 30 characters to a name. This shortage of characters per name had to be solved in some way. Two methods were used:

- (1) Abbreviation, and
- (2) Truncation.

Most human surnames fit into the limited space, but many non-human names did not. Therefore, abbreviations were used in what has turned out to be a misguided attempt to avoid truncation of non-human names.<sup>2</sup>

The general keying rule that developed was to enter the surname or first name of a non-human name in full, and then to abbreviate other words according to more or less lengthy standard abbreviation lists. In some recording offices, even first words were abbreviated if they were very common, such as National = NATL, or the state name abbreviated to replace the full name of the state.

These rules became more and more complex over time as exceptions were noticed, to the extent that the indexing manuals reviewed in this paper are so complex that it is unlikely that any key operator could actually remember all the rules that they are supposed to supply.

### **The Reversal Test**

Also, abbreviation lists were developed without a standard for determining the difference between a good abbreviation and a bad one. The following kind of test should have been applied to any word that is under consideration for standard abbreviations:

#### **Can the abbreviation be reversed to its original word without confusion?**

Consider the following words and abbreviations:

<b>Word</b>	<b>Abbreviation</b>
1. National	NATL
2. United States	US
3. California	CA
4. Indiana	IN
5. North	NO
6. Colorado	CO

NATL is a pretty good abbreviation because there is no obvious word that it could represent other than “National.”

US and CA may seem like good abbreviations, but they do not really meet the test because some businesses have names that use these abbreviations rather than the whole word. Thus an attempt to reverse the abbreviations would result in erroneous names for some non-human entries in the index.

IN, NO and CO are really bad abbreviations. IN is itself a word, such as in Jack In The Box Inc. NO is both a word and an abbreviation for another common term, “Number.” CO is also the abbreviation for “Company.”

Therefore, certain abbreviations in a legacy indexing system can be reversed without concern (NATL), certain abbreviations can be reversed as long as the original abbreviated version is maintained on the index (CA and US), and no attempt should be made to reverse certain abbreviations that would be just plain confusing (IN, NO and CO).

---

<sup>2</sup> Truncation of names works pretty well since the most important words in any name tend to be the first two, which usually will fit within whatever the field length is defined.

### **C. The Burden on the User**

The proliferation of rules to squeeze names into limited space also placed a burden on the searchers of the name indexes. Usually, an indexing system demands the user follow the same rules as used by the key operators in order to find a name on the index. To search human names, the user must know whether a compound surname is concatenated or not, and how it is placed in the index [de la hoya = DELAHOYA or DE LA HOYA or HOYA OR LA HOYA]. To search non-human names the user must know which words are abbreviated and when they are abbreviated.

In fact, the user is expected to know about each of the approximately 60 rules that are detailed in Appendixes 1-10 to be assured that she is inputting a name in a manner that is likely to get the results she is looking for.

In addition, the user is expected to know that the rules being followed by key operators are not always followed. In an actual case of an index available on the Internet, the rules of the recording office are that all dashes are deleted, but to find the entries for “A-1 Inc.” searches must be done under “A1” and “A-1.”

There are myriad examples of this sort of problem, which has given rise to a class of skilled intermediaries, title abstracters and other professional searchers, who know the vagaries of the indexing system in a recording office.

### **D. Converting from a Legacy Indexing System**

As explained in this paper and illustrated by the model indexing system outlined in Appendix 12, there are ways to reduce the burdens on both key operators and users by designing a name indexing system that takes advantage of what has been learned about computer indexing during the past 30 years, using modern computer technology.

However any upgrade to a legacy name indexing system must deal with problems like the irreversible abbreviation problem explained above. Other potential problems include:

- (1) handling of special characters,
- (2) dealing with previously truncated names in longer field sizes, and
- (3) names that were “normalized”<sup>3</sup> differently under the legacy system.

These sorts of problems may be addressed either

- (1) by finding and rekeying names that have characteristics that should be changed
- (2) by programming the search process to identify legacy search problems based on input and notify the searcher,<sup>4</sup> or
- (3) if all else will not work, by explaining the legacy problems in the user instructions.

### **E. A Typical Conversion**

As part of the project that led to the publishing of this paper, the author conducted a review of an actual legacy name index database, and programmed corrections to bring it into conformance with the model indexing system outlined in Appendix 12. The index reviewed contains 117,000 entries for a county in California.

Although there is technically no such thing as a typical conversion, this conversion did involved some of the kinds of problems that are encountered in any legacy index including,

---

<sup>3</sup> Explained below.

<sup>4</sup> For example, the user inputs the name “California National.” In the legacy system, both words were sometimes abbreviated. The search program suggests to the searcher, “Also search under CA for California and NATL for National.”

- (1) Short name field—Only 40 characters for any type of name.
- (2) Multiple individual names in one index entry—James and Mary Able.
- (3) Bad parsing of names—Jones Jr John J (Should have been Jones John J Jr.
- (4) Punctuation and special characters in some names—No consistency, 32,000 entries with punctuation.
- (5) Individual names not separable from organization names—Individual names do not even have a comma after surname.
- (6) & and AND—It appears that AND had been converted to & except in a very few cases.
- (7) THE at beginning of name—Index had many instance of THE at start of name; impossible to tell if it was deleted sometimes.
- (8) Hyphens—Were used in some surnames and company names.
- (9) Abbreviations—Some standard abbreviations used.
- (10) Estate names—Always entered as Jones John J Est of.
- (11) Trusts—TR or TRUSTTEE or TRUST at end of name.
- (12) Government entity names—CALIF used for California in some instances.

The results of this typical conversion were presented at the July 2001 meeting of the Task Force.

## **Conclusion**

Rather than dwelling further on the complications and shortcomings of the legacy indexing systems in place today, this paper is going to address what a sound indexing system ought to look like. All the manual rules will be discussed later on.

An indexing system is actually composed of three distinct but interrelated parts:

- (1) Key Entry Process.
- (2) Indexing Process (Program logic applied to key entries).
- (3) Search Process.

Each of these processes is discussed in the next chapters.

## 5. Keying Process

### Towards a Fundamental Keying Rule

As indicated in the review of current practices in the appendixes, there is today a great burden placed on the key entry operator in a typical recording office to make decisions about what to enter and how to enter it. Most of the manual rules listed in the appendixes, however, are unnecessary because of the sophistication of programming languages today.

### Programming a Rule

Programming rules to normalize names as written into index entries are, for the most part, not very sophisticated. It is beyond the scope of this paper to give the program logic for each of the rules detailed in the appendixes that can be programmed, but competent programmers should be able to figure out how to handle them.

Here is a typical key entry rule today: “Delete the word “The” if it is the first word in a non-human name.” This rule can be replaced by programmed instructions to do exactly the same thing in the index if the name is keyed as written in a field designated for non-human names. The program just looks at the first word in the field, and deletes it from the index entry if it is equal to “The”.

### Manual Decision-making

Of the 60 rules detailed in the appendixes, the only rules that **cannot be handled programmatically** against a name string are as follows (Numbers are the appendix number followed by the item letter—2D is Appendix 2, Rule D).

Decision to Be Made	Possible Rule
2D—Deciding what names to key is a human decision.	Key all names that appear in document.
2E—Name corrections are subjective.	Enter all forms of name that seem correct, including as printed in the document.
2H—Deciding about spelling errors in general are subjective.	Enter all forms of name that seem correct, including as printed in the document.
4C—Deciding what names to key when signatures and printed names are different is a human decision.	Enter all forms of name that seem correct, including as printed in the document.
4P—Multiple party names given together with “and” may be contorted to split out programmatically.	Enter each name as best you can figure.
4R—Determining whether a name is human or non-human.	Enter both ways.
5I—Determining the intent of words in parentheses and quotes.	Enter non-human name with and without the parenthetical words.
6A—Certain forms of AKA or FKA may not be parsable programmatically.	Enter both names. Do not enter the terms AKA or FKA
6F—Recognizing that a name is a government agency subject to special rules.	Mark as government agency so that special rules can be applied.

In addition, the following rules require a manual decision with respect to how the names are entered into fields, but then can be normalized thru programming.\*

Decision to Be Made	Possible Rule	Program Logic
4G—Determining whether two given names go together, such as Mary Kay.*	Enter both names in first name field.	Index separately as one and two words.
4J—Determining the nature of multiple last names.	Enter all words that may be part of the last name in last name field (e.g., “de al Hoya).	Index all forms of the name separately (e.g., Hoya, la Hoya, de la Hoya).
4L—Determining the nature of multiple word surnames.	Enter the multiple name as given in the last name field.	Index as one word and two words.

**These rules all can be condensed into a few relatively simple keying rules, as suggested above in each table.**

### **Inconsistencies in Names in a Document**

Names are inherently messy objects to deal with.

Deciding how to handle names that appear in document will continue to be a subjective process. Even if there were statutory rules that demanded the entry of names in certain ways, the rules would be either too complicated or incomplete. Even the requirement that the preparer of an electronic document place names to be indexed in designated fields is still subject to mistakes which may need manual review at recording offices.

In order to bring as much objectivity to the decision-making process, this paper takes the position that a combination of the key entry operator and program logic should index a name in all the ways that a searcher might try to find it.

For example, Rules 4G, 4J and 4L require making a decision about multiple-word human surnames or given names. If you key the words into the appropriate human name field (surname or given name), the program can then index the words in all the logical ways by recognizing that the field contains more than one word [SMITH-JONES = SMITH, JONES, SMITH JONES].

Thus, program logic can be objectively applied to the results of a subjective decision by the key operator to a index name in all the reasonably possible ways that a searcher might enter it.

### **Keying Efficiency**

Keying efficiency can be increased in the following ways:

- (1) Reduction in the number of manual rules to be followed by the key entry operator, and
- (2) Use of hot keys to enter frequently repeated information, such as party names (5C) and document types (9A)

Abbreviating words manually during key entry is not an efficiency measure. It was designed to save space, or time. It takes time to remember the abbreviation, it is inherently inconsistent, and it makes searching less accurate. Programming abbreviation rules, if they are necessary at all, is more efficient.

### **Recommendation:**

Fundamental Keying Rules—

- (1) Decide which names, and which variations of names are to be entered into the index.
- (2) If you cannot decide how to enter a name, enter it so that it will be indexed in all possible ways.
- (3) Enter human and non-human names in different field structures.

---

\* Some of these programming problems go away in an automated indexing system tied to electronic recording.

- (4) Follow government agency indexing rule when necessary (explained below).

### **Keying Accuracy**

Keying accuracy is a separate issue. It is common knowledge that unverified key entry error rates range from 3% to 5%. Visual verification reduces this error rate to maybe 1% and double-blind verification reduces the error rate to a minimal 0.1-0.2%.

The actual error rate is arguably higher in name indexing systems in recording offices because of the factors discussed in this paper, including:

- (1) uncontrolled source of input from documents, that is, divining the right name to key, and
- (2) changes to names as they appear in documents according to too many detailed keying rules.

Verification costs time and money. Therefore, the extent of verification is a balancing act between level of accuracy and cost. In an underfunded land recording office, this can be a difficult choice.

Double blind keying requires that a second key operator enter the names from a document without knowing what the first key operator entered. A program then compares the two entries and accepts them if they agree. If there is any discrepancy, a third party reviews the document and decides what the correct entry is to be. This is a verification method that recording offices should strive for.

### **Recommendation:**

The Task Force urges the budget-granting authority in each county (parish, city, town) to provide adequate funding for the following purposes:

- (1) developing an improved grantor/grantee indexing system in line with the recommendations in this paper, and
- (2) implementing a key verification system that maximizes the accuracy and completeness of the index.

## 6. Indexing Process

The second leg of a name indexing system is the indexing process. Every index has a design. The design may be better or worse to the extent that it takes advantage of factors like,

- (1) understanding of the characteristics of names, both human and non-human,
- (2) understanding of how people search for names, and
- (3) available technology.

This chapter discusses index design considerations in terms of the technology available today, compared to the crude design of legacy indexing systems.

### Name Fields

There are three considerations in determining the design of name fields:

- (1) type of fields, and
- (2) length of fields.
- (3) index name fields.

Most indexing systems recognize two types of names, human and non-human.<sup>5</sup> Therefore, at a minimum there should be two types of fields to keep the two types of names separate. In addition, it is recognized in all indexing systems that human names consist of various parts, including surname, given names, and lineage titles. Therefore, a set of fields should be set aside for human names. All indexing systems also recognize a type of non-human name that needs special logic, government agencies. Therefore, non-human names might use at least two fields, one for basic name and one for corporate ending (Inc., etc).

Once upon a time disk space and computer memory cost a lot of money. In 1965 32,000 characters of memory on an IBM computer cost \$55,000. Therefore, every character that could be saved was worth a great deal of "savings." As a consequence the first computer systems designed to index names kept fields lengths to the minimum that the designers could get away with and still be able to enter all names that appear on documents.

Today, although there are operating systems and programming methods that allow for variable length fields, standard database design and programming standards continue to demand fixed length data fields, so the problem of how long to make each name field is still with us. It is also clear that the reasonable required length of data fields depends on the type of name. A business entity name, for example needs more space than an individual name. Given how inexpensive disk storage has become, there is no excuse for not having adequate name field lengths.

The cost of storage was also a consideration in the past concerning how many index entries should be created for one name. Today there is no reason to constrain the number of index entries that arise from applying objective programming to developing a variety of index entries from each name entered.

### Indexed Name Normalization (Name as Entered vs. Name as Indexed)

As implied in the chapter on keying, it is the position of this paper that the name as key-entered is separate from the name as indexed. The indexed name is "normalized," that is, the indexing rules detailed in this paper are applied by a computer program to the name as entered to produce one or more index entries to the document.

---

<sup>5</sup> Even in legacy name indexing systems where human names were not separable by some mark, conversion routines can be programmed that will separate most human names without eyeballing.

One of the most significant consequences of having a normalization program to run against the name as entered is that the same program can then be run against a search name as entered to normalize it in the same way. The advantages of this approach to improving the completeness of a search should be obvious.



## 7. Search Process

The third leg of a name indexing system is the search process, that is, what a user of the index goes through in order to find entries in the index for particular names.

### Four Ingredients in a Search Process

The actual search process depends on the following elements:

- (1) Basic index structure,
- (2) Search logic,
- (3) Search input parameters, and
- (4) Presentation of results.

### 1. Basic Name Index Structures

There are two basic types of name indexing structures:

- (1) Word-based, and
- (2) Concatenated.

Each has benefits and shortcomings.

#### Word-based Structure

The indexing structure in most recording offices is word-based, that is, each word in a name is separated by a blank space. This design demands that certain conventions be used to join or separate certain words. Rules 4H, 4J, 4K, 4L, 4N, 5K, and 5O all deal with the issue of how to index words that may be separate or combined.

Since a space is part of the logical (or sort) sequence of the index, the choices about what to concatenate and what to separate affect how the index needs to be searched. Consider the following series of names:

- (1) John La Sorda, David Lansing

Depending on the rules of the index these names may appear in three different parts of the index:

- (1) Keep compound names separate:

LA SORDA JOHN  
LANSING DAVID

- (2) Concatenate compound names:

LANSING DAVID  
LASORDA JOHN

- (3) Do not count compound names:

LANSING DAVID  
SORDA JOHN LA

- (2) John Smith-Jones, David Smithfield

Depending on the rules of the index these names may appear in two different parts of the index:

- (1) Replace hyphen with blank or keep dash:

SMITH JONES JOHN or SMITH-JONES JOHN  
SMITHFIELD DAVID

- (2) Deleted hyphen:

SMITHFIELD DAVID

## SMITHJONES JOHN

These illustrations may seem innocuous as given, but can be searching nightmares in indexes where there are hundreds or thousands of intervening names like David Lansing or David Smithfield.

One of the best solutions to the problem of deciding whether to concatenate or separate words, especially in surnames, is to index more than one form of name, as discussed elsewhere in this paper. The most complete index possible would have each name indexed in all the logical ways that it might appear to a future searcher. It is possible to index one name in many ways using up-to-date programming techniques, as discussed above.

### **Concatenated Structure**

This structure, used in a number of proprietary databases, deletes all the blank spaces for the words to be indexed. For example, in a system where all special characters are always deleted:

Carl R. Ernst, Jr. = ERNSTCARLRJR  
Ernst Publishing Co., LLC = ERNSTPUBLISHINGCOLLC  
Jane Simpson-Smith =SIMPSONSMITHJANE  
2 To One Pace-Maker.com, Inc. = 2TOONEPACEMAKERCOMINC

The advantage of this type of structure is that neither blanks nor special characters impact where the name is indexed. Therefore, searchers know that neither spaces nor special characters will get in the way of finding the name entered.

### **Soundex Variation**

A very special form of concatenated index structure is the basis for what is called Soundex logic. A field is generated from the concatenated name, or part of the name such as the surname only, that eliminates all vowels and considers certain consonants to be equal in value. Without going into the details, this form of index field will give equivalent results for name variations such as the following:

Ernst, Ernest, Earnest  
Volpe, Folpe  
Hernandez, Hernandes

Because of this characteristic of the matching power of Soundex, it is especially useful in matching human surnames, and is less useful in matching non-human names. One negative of this variation is that it tends to give too many hits for certain names.

## **2. Search Logic**

The subject of search logic is separate from, although dependent on, the structure of the name index. There are two basic types of search logic:

- (1) Left-to-right, and
- (2) Word-based.

### **Left-to-Right Search Logic**

This form of search logic is based on the ASCII sort sequence used by most computers today, that is blank space,<sup>6</sup> followed by most special characters, followed by numbers, followed by capital letters. This form of search logic is fast because the index field can be sorted in advance and placed in a separate index file. The name to be searched can be found in the index file by matching letters in the search name to the index file starting with the leftmost letter. Thus, the search is on the letters in the name, “left-to-right.”

---

<sup>6</sup> Remember that a blank space is a “letter” for the purposes of sorting.

Left-to-right search logic may be applied to either type of index structure, word-based or concatenated. The blank space in a word-based index is just another letter.

### **Word-based**

Word-based logic matches input words (or a set of letters) to all the words (or letters within words) in a name. It requires a lot more computer power than left-to-right, but is valuable when words may not be in the right order. For example, in an Asian name like Lee Ng, most of use have no idea which word is the surname. Word-based logic that checks for two words would find both NG LEE and LEE NG.

## **3. Search Input Parameters**

### **Too Little vs. Too Many**

All search logic, whether manual or computer programmed, strives to balance between getting too many hits and getting too few. If you enter the name SMITH in most any name index when you want to search for Xavier Smith, you will waste significant time paging through the list of all the other Smith's before you get to those whose name starts with an "X."

In order to limit results, the system may contain restrictions on the input of names into the search system. For example, name input limitations today include the following:

- (1) no wildcards in surnames, and
- (2) no less than three characters.

The wildcard limitation is used by some human name searching systems. Only full surnames are found based on the letters entered. For example, ERNS would not find ERNST. The minimum characters limitation is designed not to allow just reading through all the entries that start with a letter by just putting in that letter.

Whether or not there are limitations placed on the entry of names, all legacy name indexing systems are programmed to find the first match in ascending sequence on the letters input, and then to apply a format to the presentation of the results..

## **4. Presentation of Results**

Once computer search logic is applied to the letters entered by the user, the system returns the results of the search in some standard, programmed fashion. Three forms of presentation are

- (1) Two-Way Directory,
- (2) One-Way Directory, and
- (3) Matches Only.

The first two formats are common to legacy name indexing systems. Essentially the search logic places a pointer at the first word that matches the letter as input in ascending sequence, that is, entry of ABC will point to the fist company starting with the name ABC, and ERN in a human name index will point to the first surname, which might be ERNEST, that begins with those letters. From this starting point, the indexing system then either allows the searcher to got backward and forward in the index (two-way) or only forward (one-way).

Under Revised Article 9 of the Uniform Commercial Code, state central filing offices are expected to develop a new search strategy that will require the entry of the full name of the debtor, for which only the exactly matching entries in the index will be returned. This is a form of "Matches Only" indexing logic that is not appropriate to grantor/grantee name indexes.

## **Search Name Entry Strategies**

An unstated assumption in all legacy name indexing systems is that only part of a name should be entered in order to perform a search of the name index., Only a fool, some might say, would enter the whole name because the more letters you enter, the less likely you will find a match.

No matter what the input parameters and the presentation methods, both the one- and two-way type presentation systems require the user to form some sort of search strategy in order to find all the matches to the actual name being searched. Discussion of these strategies is not within the scope of this paper.

However, good functional name index system design will assist users to find what they are looking for by decreasing dramatically the burden of matching input to the system peculiarities so they can concentrate on their search strategy.

## **User Search Instructions (2B)**

There are two keys to a user being able to perform accurate searches of an index:

- (1) The design of the indexing process, including keying and name normalization, and
- (2) The user instructions.

**Possibly the greatest flaw in legacy indexing systems is in the combination of complex keying and searching rules with no written user instructions.**

### **Recommendation:**

1. Complete and accurate user instructions will be posted conspicuously wherever users can access the name index.
2. Instructions will note all non-natural name parts that may appear in the index (abbreviations, placement of “The”, etc.)

This includes sorely lacking instructions for searching on-line/Internet-based indexes.

## **Indexing Design Rules**

To summarize:

Fundamental Index Design Rules—

- (1) Keep to a minimum the number of exceptional conditions in the index that need to be known by users.
- (2) Allow index user to enter names as given. Adjust search names to index design by programming.
- (3) Key human names to be searched into separate fields for each part of the name.

## 8. Indexing-Related Rules

Some of the rules used in legacy indexing systems have been mentioned already. In the following chapters, each of these rules as used in existing indexing standards manuals is discussed in the amount of detail necessary to indicate whether a rule is necessary to maintaining a name indexing system and which version of the rule might be preferable as a standard.

It is important to understand that no matter how a name indexing system is designed for the purposes of notifying the public about documents archived in land recording office, most of these rules continue to be necessary. The point of this paper is that most of these rules can be interpreted by computer programming rather than by key entry operators and index searchers.

### Using the Appendixes

Each of the detailed rules from the indexing manuals reviewed have been summarized in Appendixes 2-9.

The purpose of this categorization is to allow discussion of the rules in any orderly fashion. It is not the only ways indexing rules could be categorized, but it is hoped that this order will be useful to provide a model for those recording offices and state associations that wish to develop indexing standard manuals in the future.

The rules are categorized as follows:

- (1) General Rules—Rules that affect the indexing of all debtor names in a general way.
- (2) Name Field Lengths—Any indication of the technical constraints on name fields and solutions to the constraints.
- (3) Human Names—Indexing rules that apply to human names.
- (4) Non-human Names—Indexing rules that apply to non-human names.
- (5) All Names—Rules that apply to both forms of names.
- (6) Special Names—Rules that apply to unusual forms of names.
- (7) Search/Presentation Elements—Any rules that apply to how to search or present the results of searches.
- (8) Other Subjects—Miscellaneous rules, notations, etc. along with the indexing rules.

In the following chapters, we discuss each of the rules in each of these appendixes. There is general agreement as to the handling of many of the rules. In most other instances, there is no particular negative impact of choosing one of the alternatives mentioned over another. However, in some instances, alternative rules have consequences that make them less preferable than others. Based on the discussion of alternative rules below, recommendations are advanced with respect to preferred alternatives.

### General Rules

The general rules have already been discussed in the context of developing a name indexing system, except as follows.

#### **2C—Separate or Combined Grantor/Grantee Index**

Both methods are used and each has an advantage. This paper takes no position on the matter.

#### **2F Single Party documents**

In some systems the name in a single party document is indexed as both grantor and grantee. This makes no difference in a combined index.

#### **2I—Information Included in the Index**

Along with the basic information needed to locate the document, the index may contain other useful information including:

- (1) property identification number,
- (2) reverse party name,
- (3) reference to original documents in subsequent recordings.

## 9. Human Name Rules

### Name Structure (4A, 4D, 4E, 4F, 4G, 4P, 4Q, 4R, 4S, 4U)

All US indexing systems assume that human names are written in the American form, that is, first name followed by middle names, followed by surname, followed by lineage suffix.<sup>7</sup> The name is then reordered for placement in the index into

- (1) surname,
- (2) given names (4G),<sup>8</sup> and
- (3) lineage suffix (4Q)

Surname problems are discussed below. Given names may contain abbreviations (4I), which are entered as they appear in the document. Acceptable lineage suffixes are sometimes listed in the indexing manuals, including Jr., Sr., III or 3<sup>rd</sup>, II or 2<sup>nd</sup>. [John Martin Jones III = JONES JOHN MARTIN III]

Most prefixes, such as Dr. and Mrs. Are excluded from the name except in creation circumstances, as follows:

- (4) The prefix is the only given name (4E, ), such as Mrs. Smith, [SMITH MRS]
- (5) No given gender specific name is available (4E), such as Mrs. John Smith [SMITH JOHN MRS],  
or
- (6) No apparent given name (4F), such as Cardinal O'Connor [OCONNOR CARDINAL] or Captain Jones [JONES CAPTAIN].

Other information may follow these three standard parts, as follows:

- (7) Capacity/Title (4S),
- (8) Associated name (), and
- (9) Association Abbreviation

All the indexing manuals reviewed included capacity words, which may also be known as titles such as Trustee, generally replaced with standard abbreviations (Appendix 10A lists capacity terms that may have standard abbreviations). Professional titles, such as CPA, MD and Esq., are not entered in human name forms (4R).<sup>9</sup> [John Jones, Trustee = JONES JOHN TR; Dr. John Jones MD = JONES JOHN]

Items 5 and 6 of name information were mentioned in one or more indexing manuals. Generally, husband and wife names are indexed separately without any designation or association (4P). In two cases, the given name of an associated spouse was placed after the name of the other spouse (4P). In these two cases, the associated name was separated from the party name in such a way as to not get in the way of the index sequence. In two instances the term "FKA" was listed after the name if the document indicated the name was, for example, a maiden name (4D, 4P) [Jane Jones (formerly Smith) = JONES JANE and SMITH Jane /FKA].

Recommendation:

If it is desired to include information about name associations in the index, put the information in a separate, non-indexed field.

---

<sup>7</sup> The only state to add a rule to deal with variations from the American system is a California rule to deal with Asian names (4U) by not altering the sequence of a name in which all words are attached by dashes.

<sup>8</sup> In the Nevada indexing manual, multiple word given names are supposed to be concatenated (Mary Kay = MARYKAY). This is a degree of speculation that most recording offices would not agree with.

<sup>9</sup> However, if it is unclear whether the name is a human name or a non-human name, the non-human form may also be indexed as written on the document.

## **Surnames (4H, 4J, 4K, 4L, 4M, 4N, 4O, 4P, 4T)**

Most of the rules for entering human names are required on account of problems with parsing out a name as written, because there is no law in any state that requires a human name to be formatted to make clear what is the surname.<sup>10</sup> The following problems are addressed in the indexing manuals reviewed:

- (1) Apostrophes (4H, 4K),
- (2) Mc, Mac (4H),
- (3) St., Ste., Saint (4H),
- (4) One word names (4O), and
- (5) Multiple words in last name (4J),
- (6) Vda. de and V. de (4T),
- (7) Multiple name last names (4L, 4N, 4T),
- (8) More than one person's name (4P), and
- (9) Names in parentheses or quote marks (4M).

Apostrophes are generally deleted from surnames [O'Brien = OBRIEN], except in the Nevada manual. Standards for entering surnames with Mc or Mac in them are split down the middle in the indexing manuals reviewed. Considering the inconsistency with which people print such names however, the better argument would be to merge them with the following word. Some manuals also apply the space deletion logic to the terms St., Ste. and Saint followed by another name. These terms, however, are clearly distinguishable from Mc and Mac because they are rarely concatenated in practice, so that concatenation should not be considered as a standard. In all these cases, computer logic can handle the rules for concatenating or not concatenating in these word forms.

One-word last names, such as Cher, are not mentioned in any of the manuals reviewed. Logic suggests that they be entered as written.

Multiple word last names present a particularly nasty problem to key entry operators. The problem comes in two forms. First, there are names like La Sorda, Van der Kampen, and de la Hoja. Is the prefix part of the surname or not? All the indexing manuals reviewed take the view that the suffixes are part of the surname, but none of the manuals includes any kind of comprehensive list of the types of surname prefixes. The standard is, in a word, subjective. The manuals usually include an admonition to use common sense, or to enter both ways if it is not clear to the key entry operator what the last name is. Also, the manuals are split as to whether to enter the surname with or without spaces [Van de Kamp = VANDEKAMP or VAN DE KAMP]. As with Mc and Mac, some writers will concatenate while others will separate. Therefore the appropriate conclusion is to concatenate for consistency.<sup>11</sup> These variations leave searchers with the problem of second guessing the key entry operator as well. These problems can be resolved objectively as follows:

Recommendation:

- (1) Enter all forms of multiple word surnames names in the index. [John de la Hoja = HOJA JOHN DE LA, LAHOJA JOHN DE, DELAHOJA JOHN]
- (2) Concatenate all the parts of the surname as entered into the index.

The term "Vda. de" in front of a surname is considered a special case in the California manual. However, following the recommendation will assure that the actual surname will be indexed.

---

<sup>10</sup> Although such a statute would be helpful to improving the accuracy of indexes, statutory recommendations are beyond the scope of this paper. One benefit of certain forms of electronic recording is a requirement to divide the parts of human names into logical fields.

<sup>11</sup> Of course, it is also easy to program to enter the names into the index in both the concatenated and unconcatenated forms.



These rules can be programmed as long as the key operator enters all the words into a surname field.

The second form of multiple last name problem comes from the use of what are called hyphenated surnames. If the surname is actually hyphenated, the problem is minor since the indexing rule can easily accommodate indexing the person under all three possibilities [Jane Simpson-Jones = JONES JANE SIMPSON, SIMPOSON JANE, SIMPSON JONES JANE]. This leaves the question of whether to include the hyphen in the surname to be discussed below. One indexing manual suggested that the key entry operator guess whether a non-hyphenated name was a multiple surname; this is not recommended because it is highly subjective.

Finally, the indexing manuals address what to do with names in parentheses or quotes within a human name. Some examples are,

- (1) Theodore “Ted” Jones [=JONES THEODORE, JONES TED],
- (2) Jane Smith (formerly Jones) [SMITH JANE, JONES JANE],
- (3) Jane Smith (Jones) [SMITH JANE, JONES JANE], and
- (4) John (Jack) Smith [SMITH JOHN, SMITH JACK]

Although what to do with each of these seems obvious the human key entry operator, they are all examples of name forms that computer programs have difficulty with. Each indexing manual properly suggests sorting out the name in key entry according to common sense, as indicated above.

### **Character Set (4B)**

The character set for human names is generally just the 26 letters of the capitalized alphabet. Apostrophes and dashes (hyphens) are the only special characters that can occur in human names. Apostrophes are generally deleted (4K) without harm.

Dashes usually appear in compound surnames, although they can occur in given names as well. Handling of dashes as they appear in multiple word surnames needs to be considered carefully. There are three options:

- (1) Keep the dash,
- (2) Replace the dash with a space, and
- (3) Delete the dash.

None of these rules is perfect from the point of view of a searcher, although confusion is more likely in a first name that a last name [Mary Kay = MaryKay =Mary-Kay]. If the searcher is admonished not to enter full first names that might be compound, the problem can be confined to surnames. It is less likely that there will be a concatenation in a surname [Smith Jones may = Smith-Jones, but is unlikely to = Smithjones].

If a hyphenated name is given in a document without a hyphen, it is impossible to objectively determine that the two words should go together. If, on the other hand, the name is hyphenated, there is no good case for replacing with a blank either. Finally, deleting a blank will create a rather funny-looking name in many instances.

As a further consideration, how is a dash in a non-human name handled? Does it make sense to handle a dash in human names one way and in non-human names another way?

### **Recommendation:**

Enter hyphenated names into the index in all three possible ways.

## 10. Non-Human Name Rules

### Character Set (5A) and Special Characters (5B)

The indexing manuals reviewed show a number of different approaches to including in the index characters other than the 26 letters and 10 numbers. Following is a list of all the special characters that sort before capitalized letters in the normal ASCII sort sequence. The characters up to the slash (/) sort before numbers 0-9 and the characters from colon (:) on sort after the numbers 0-9.<sup>12</sup> The list indicates whether any of the manuals reviewed includes this character in its index, and whether the character also gets other special treatment.

Special Characters	Indexed
!	Yes
“	No
#	Yes
\$	Yes
%	Yes
&	Yes (also used to represent “and”)
‘	Yes
(	No
)	No
*	Yes
+	Yes
,	No
-	Yes
.	Yes (for dot-com’s)
/	Yes (primarily for dates and fractions)
:	No
:	Yes, in time (10:00)
<	No
>	No
?	Yes
@	Yes

Some indexing methods replace special characters with blanks whereas others delete the character altogether. In one instance, special characters are supposed to consistently be replaced with an asterisk (\*) by the key operator. In another instance the word “and” is always replaced with &, which is a very suspect rule. The dot-com era has seen at least one indexing manual make a special case of the period (.). This rule is also suspect because the period was probably deleted in company names before the dot-com era, which would increase inconsistency in the index.

Words surrounded by quote marks or parentheses are handled under special key entry rules in certain manuals (5H).

---

<sup>12</sup> This is a rarely appreciated sort-sequence problem if your index includes these later characters.

The extent of variations in treatment of special characters in non-human names is troubling. Extrapolating usages in those states that do not have any formal indexing standards suggests that from one recording office to the next a searcher can have no comfort how to handle special characters in forming a search query.

The question we suggest asking when designing an indexing system for non-human names is,

**Does the deletion of all special characters from non-human names compromise or promote any of the fundamental objectives of a name indexing system?**

On the one hand, deleting all special characters does no harm to finding an entry; nor does it generate any inconsistencies in the presentation of the indexed entries. On the other hand, if the deletion of special characters is transparent to the user of the index, searching will be easy and straight-forward, whether or not the computer search program normalizes names or not. If the search system does not normalize names entered for searching, the only user instruction needed is,

When entering a non-human name, do not enter any special characters in the name. Only enter letters and numbers.

**Recommendation:**

?

**Frequently Indexed Party Names (5C)**

A couple of the manuals suggest that the names of parties that appear on the index frequently, such as major lenders, be tabled in such a way that a hot key can call up the name. For example, function key F! might stand for Bank of America. This is a keying efficiency idea that promotes the fundamental objectives of a name indexing system.

**Special Words and Terms (5E, 5G, 5J, 5K, 5L, 5M, 5O, 5P, 5R, 5Q)**

Certain words and terms get special treatment in many of the indexing manuals, including:

- (1) "The" as first word (5E),
- (2) "A" as first word (5E),
- (3) "And" (5M),
- (4) "\*&" (5L),
- (5) "In" "At" "Of" (5O),
- (6) Single character series (5K),
- (7) Alphabetical and numeric forms of numbers (5J),
- (8) Directionals (5O),
- (9) Roman numerals (5P),
- (10) "Mount" vs. "Mt." (5R),
- (11) Fractions (5Q), and
- (12) Business endings (5G),

"The" is unanimously deleted from the front of non-human names, and some manuals have the word placed at the end of the name. On the other hand, the manuals are equally unanimous in not altering non-human names starting with "A."

Half the manuals replace the word "and" with an ampersand (&). This is an interesting rule. It promotes simplicity and ease of use because users do not need to figure out which form a non-human name takes. The user instructions must be clear never to use the word "and" in a search unless the search program does the conversion to ampersand.

There are problems with handling “&” as it appears in non-human names. There are three possibilities:

- (1) Enter with surrounding blanks,
- (2) Enter without surrounding blanks, and
- (3) Enter with or without blanks, as it appears in the document.

The manuals favor entering as is between words, just as with the word “and.” There are variations in usage if the ampersand is between free-standing letters. Standardization on deleting surrounding blanks in this instance seems preferable to the other alternatives.

The California manual makes a special case of geographic pronouns, that is, “in,” “at,” and “of,” which it allows to drop. Although this saves a little space, it can create confusing entries as well as confusing user instructions, and is not recommended.

The indexing manuals generally agree that the way to treat single character series is to eliminate spaces, period and hyphens and concatenate the letters [A B C = ABC, A.B.C. = ABC, A-1 = A1 and A T & T = ATT&T]. Variations include J.C. = J C but J C = JC and keeping the dash in A-1. Neither of these exceptions is recommended as they are both inconsistent with other rules.

Certain other types of words are mentioned in the manuals. All the manuals agree that the numeric and alphabetic form of numbers are not to be changed. The California manual calls for abbreviation of single and compound directional words (North, etc.) in some cases. This practice does not meet the reversal test explained above. Other words and terms that are not changed as a matter of practice in all the manuals include roman numerals and Mount and Mt.

Fractions are handled in two different ways in the manuals that mention them:

- (1) Enter as is with a space before a preceding whole number (1 1/2), and
- (2) Replace the slash with a blank [1/2 = 1 2].

Although the first of these two rules appears to be reasonable, it would make a special case of the slash in one usage. Instructing a user about this tiny exception would not accord with the fundamental objectives of an indexing system. Therefore, we see no reason not to delete the slash in this as in all their cases.

A majority of the indexing manuals make an attempt to normalize business endings such as Inc. and Company, using standard abbreviations. Such a list makes sense as the ending is rarely if ever critical to finding the name in the index. The user instructions should just say never to enter business endings as part of the search string.

## **Human Names That Are Part of Business Names**

The manuals generally agree that human names that are part of a business name (Carl R. Ernst Publishing Co.) are to be entered as is, with exceptions in two cases:

- (1) St. James entered as STJAMES (California manual), and
- (2) Do not concatenate Mac Donald (three manuals).

The first exception follows the usage for human names, which has been recommended against above. The second exception creates differing usage for human and non-human names.

### **Recommendation:**

Human names that are a part of business names should be changed only to conform them to surname usage?

## **Extraneous Words in a Non-Human Name (5H, 5N)**

Three cases are mentioned in the indexing manuals:

- (1) Names with embedded parentheses (5H),
- (2) Names with embedded quotation marks (5H), and
- (3) Combined business and division names (5N).

Dealing with the first two situations is subjective unless the special characters are removed.

The third situation is the equivalent of combined human names like John and Mary Smith. A typical example is Quaker Oats Division, General Mills Corporation. Like combined human names, the manual that deals with this example separates the name into two index entries, one for the corporation and one for the division, followed by the modifier "DIV." Entering both names in the index makes sense, whether or not a modifier is included..

## 11. All Name Rules

Both human and non-human names have some characteristics in common. Among these characteristics are the following six:

- (1) AKA (6A),
- (2) Apostrophes (6B),
- (3) Et al (6C),
- (4) Name abbreviations (6D),
- (5) Ordinal numbers (6E), and
- (6) Fictitious names (6F).

AKA rules have been discussed already. In a non-human name, the two names would be separated. Apostrophes are generally deleted in both human and non-human names. Et al is generally deleted. Name abbreviations are used in many instances for non-human names, such as US, for some other words in order to shorten them, such as NATL for National, and for modifiers (Trustee) or business endings (Inc). Ordinal numbers are printed together [John Jones 3<sup>rd</sup> = JONES JOHN 3<sup>RD</sup>].

Fictitious names are traditional names that replace persons unknown, such as John and Jane Doe and Doe Corporation. Two manuals mention these names and do not index them.

## 12. Special Names

Certain forms of names are so special that separate rules apply to them. The forms are,

- (1) Estate/Decedent names (7A),
- (2) Trustee, guardian, and other names with capacity (7B),
- (3) Deed of trust trustee names (7D),
- (4) Power of attorney names (7E),
- (5) Partnership names (7G),
- (6) Trust names (7C),
- (7) Government agencies (7F), and
- (8) Names on plats (7H).

Estate names are generally entered as human names followed by a modifier such as EST or DECD.

Trustee, guardian and other names with indicators of capacity are generally indexed with an abbreviated modifier. Trustee on deeds of trust are indexed in some deed of trust states in the same way as trustees of other types of trusts. The name of a person signing under a power of attorney is mentioned in only one manual, where the rule is not to enter the name.

Partnership and trust names are generally entered along with the names of the trustees or partners. In two instances the names of the beneficiaries are indexed.

Government agency names are indexed in a wide variety of ways. This diversity suggests that standard formats are needed to keep these names straight in the index while at the same time indexing them in a manner that can be easily explained to users. For example, separate fields might be used as follows:

- (1) Name of jurisdiction (US, Arizona, Phoenix, Bedford),
- (2) Level of government (US, State, County, City, Town, Village),
- (3) Name of agency (Transportation, Forest Service, IRS)
- (4) Type of agency (Department, Court)

The manuals agree that the names of officials signing documents should not be indexed.

Two manuals mention indexing of owner names on plats.

### 13. Indexing from Electronic Recordings

There are two ways to generate an index entry from an electronic recording:

- (1) keying from a textual print of names found in the electronic document, or
- (2) indexing directly from the electronic document data fields.

Of course, manual keying of index entries would create a bottleneck in an electronic recording system that would offset some of the potential benefits, such as accuracy and instant turnaround time. Therefore, the grail of electronic recording—to accept a recording without human intervention—cannot be attained without electronic indexing.

The primary problem with fully automated indexing from electronic documents is the low level of understanding by preparers about essential elements of name indexing, such as,

- (1) printing the name the same way throughout a document,
- (2) assuring that the printed names agree with the names as signed, and
- (3) parsing individual names properly into separate fields for first, middle, and last names, and separating other lineage and role identifiers.

A fully automated indexing system will require a combination of preparer training, program and database design, and computer edits to be effective.

It may be wise to check on the accuracy and consistency of index entries that result from the automated system through periodic manual eyeballing of automated index entries.



## 14. Other Forms of Indexing

### Tract Indexes

A tract index is one that is ordered by the designation of property location. This type of index may be based on short legal description, such as section, lot and block or on a parcel number. Creating and maintaining a tract index demands that all recorded documents that refer to a specific property contain the shorthand information necessary to post such an index. In larger counties in states where recording offices do not have the capability of developing a tract index, private industry has stepped in to fill the void in the form of title plants.

Only a few states require their recording offices to maintain this type of index. We believe these states include Nebraska, North Dakota, Oklahoma, South Dakota, Utah and Wyoming. In addition, a few other states permit recording offices to maintain tract indexes. We believe these states include Kansas, Minnesota, Ohio, and Wisconsin.

In addition, more than 800 counties in the US use some form of property identification number (PIN), which must appear on certain or all documents submitted for recording.