# Information Made Accountable

The Data Projection Model

Michel Biezunski

Infoloom mb@infoloom.com

http://www.infoloom.com

## Michel Biezunski

- IT Consultant/Innovator, dba Infoloom, based in New York.
- Created the Topic Maps paradigm. Initiator of the ISO/IEC 13250 standard.
- Major current project: TaxMap, an electronic delivery tool for IRS publications and forms.
- Background: History/Philosophy of Science

### Rationale

- The problem
  - Accountability works with financial information.
  - Accountability doesn't work so great with nonfinancial information.
- The fix
  - Use accounting-like approaches for information management.

## **Outline**

- Theoretical Introduction: Why bother?
- The Data Projection Model:
  Double Entry Bookkeeping for Information.
- Application Examples.

#### Luca **Pacioli**

The father of accounting... and of other things



## Luca Pacioli

1445-1514 (or 1517)



Worked on Perspective.

"Invented" Double Entry Bookkeeping.

Wrote on Accounting Ethics and Cost Accounting.

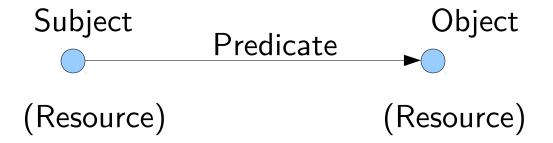
Elementary Algebra

Taught mathematics to Leonardo da Vinci.

Wrote De Divina Proportione.

## RDF, the foundation of the Semantic Web

Based on "triples". Uses URIs to name the relationship between things ("Predicate") as well as the 2 ends of the link.



## Subject - Object

In the *Quattrocento* perspective system, the ideal point of localization is "the one which places as opposite, but parallel, the **subject** and the **object**."

This is called a "one point perspective" or "central perspective". Governed by a single vanishing point.

Several conceptions of perspective:

- Science of Vision
- Technique of Representation
- Technique of the Measurement
- Architecture

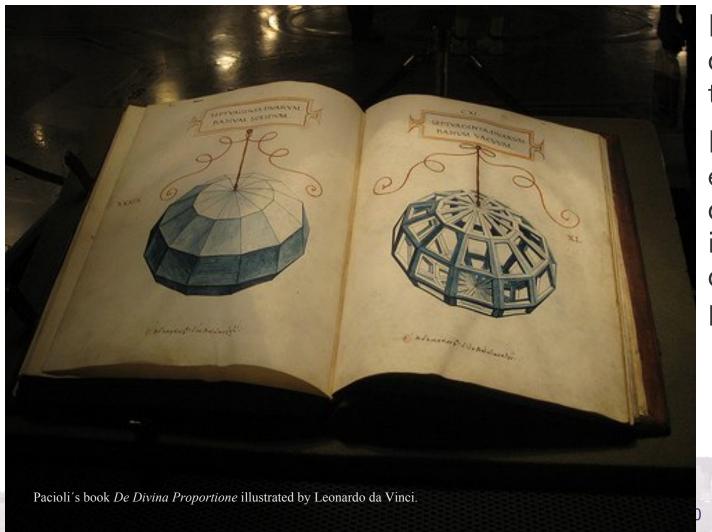
## Perspective as a Science

It is among our senses, the wise men conclude, that the sight is the noblest one. That is why vulgarly said, with reason, that the eye is the first door from which the intellect understands and likes.

(L. Pacioli, De Divina Proportione, f.4r)

The eye, which is said is the window of the soul... (Leonardo da Vinci, *Paragone*)

## Flattening the world



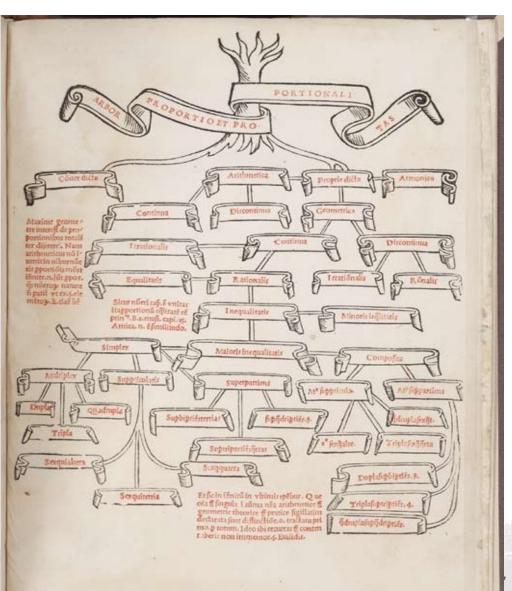
Perspectives are defined according to projections.

Perspectives express ways 3dimensional space is rendered into 2dimensional, i.e. projections.

## **Real World Information:**

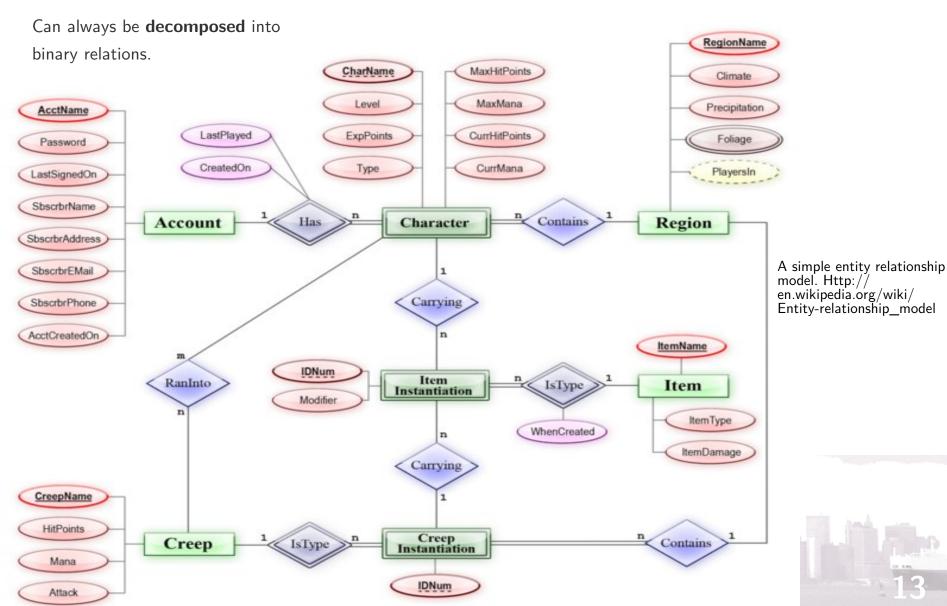
- Is multidimensional.
- Can be flattened to be processed.
- Binary relations correspond to 2D space
- Translating a world of n-ary relations into a world of binary relations is a kind of **projection**.
- The result of projecting is a graph.
- Perspective is what accompanies projection from n-ary relations to binary relations.

## Proportions as a Semantic Network



The tree of proportions and proportionality, by Luca Pacioli.

## **Entity-Relationship Model**



## The limitations of RDF

Or rather, the way it is used

#### The problem

A one point perspective on things.

Not contextual: Who said **SO?** (not automatically reified)

Operates in a closed world.

First order logic works if everything known.

#### The fix

Subject - Predicate - Object: Nodes derive meaning from connectors (provide context)

Enable multiple perspectives.

Every node is an "account".

Logic is not built-in, multiple logics can be super-imposed.

Still RDF, but used differently.

## Introducing the Data Projection Model

• The Data Projection model enables information systems to become auditable. It aims at facilitating system maintenance and knowledge management. The Data Projection Model can be used to integrate information assembled from a variety of sources and to express multiple perspectives on the same information set.

## What is Data Projection for?

Tracking background of any information item for:

- Search engine efficiency: Why a given item is on top of the hits?
- Identity Theft Investigation: Where is the leakage
- Privacy: Who knows what?

Managing funding with strings attached: How is this grant being spent?

Where does the money go? Accounting++

Integrating heterogeneous sources.

Creating diverse views for targeted audiences.

## Data Projection. How it works

Views result from linking data.

Semantic is in the views.

Multiple views are possible:

- Filtering out unwanted information.
- Focusing on details (microscopic views)
- Anything in between.

Views can be created after information is produced.

Different people can have different perspectives of the same information set.

## Information and Accounting

#### Accounting

Transactions occur between accounts.

Account statements: all transactions from or to an account.

#### Information

Any information item is always related to at least another one.

Audit trail: all links from or to an information item.

## Bookkeeping made accountable

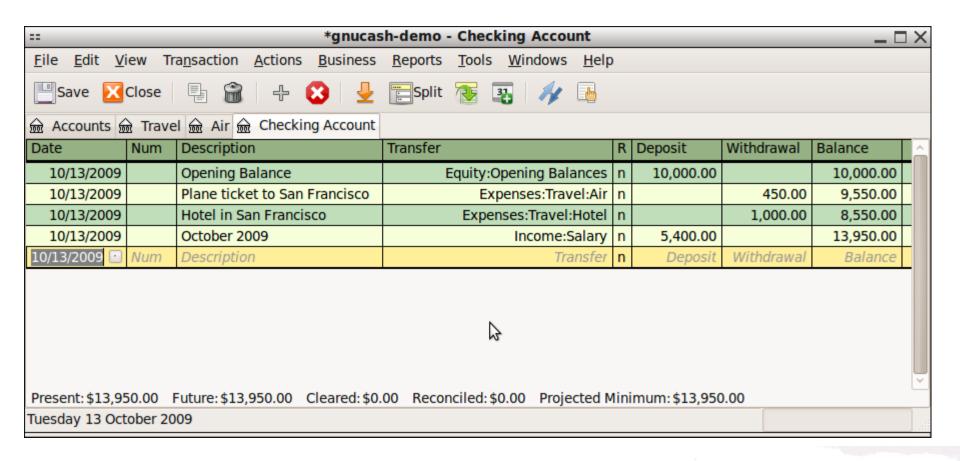
#### **Single Entry**

- List of expenses per category
- List of income sources per category.
- Some money may be unaccounted for (although not desirable).

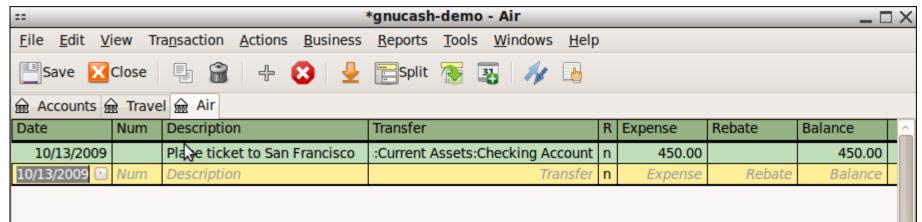
#### **Double Entry**

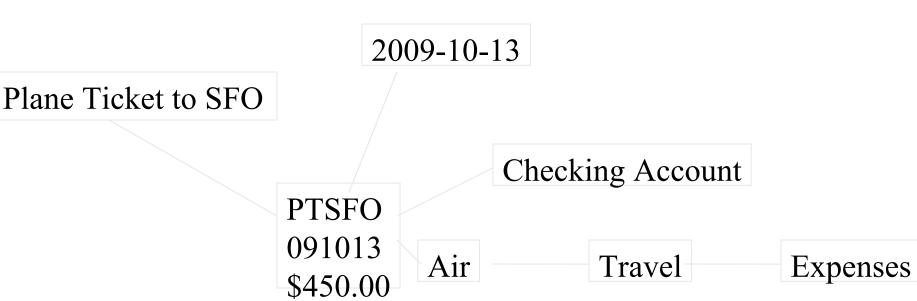
- Organized by accounts.
- Each transaction affects two accounts in a way that keeps the overall system of accounts in balance.
- No money amount unaccounted for.

## Double Entry Bookkeeping

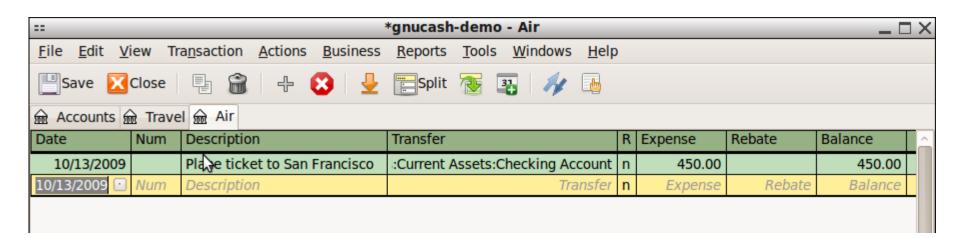


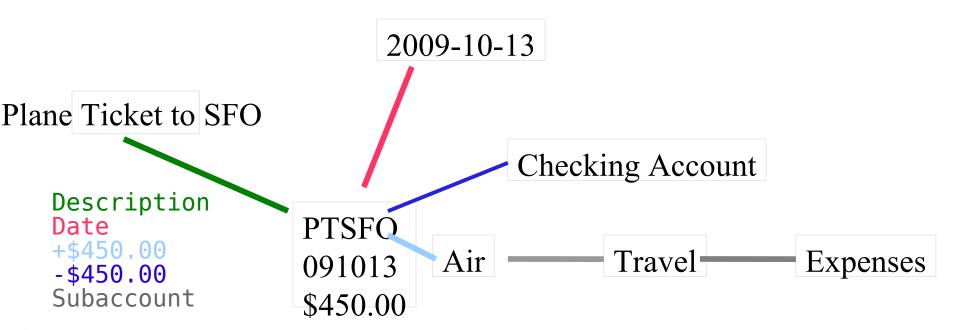
## Double Entry Bookkeeping & DPM





## Double Entry Bookkeeping & DPM





### **DPM Notation**

```
2009 - 10 - 13
  < PTSF0 09-10-13 450.00
                             date
                                               Plane ticket to SF0>
  < PTSF0 09-10-13 450.00
                             description
  < PTSF0 09-10-13 450.00
                             450.00
                                               Air
  < PTSF0 09-10-13 450.00
                              -450.00
                                               Checking Account
  < Air
                                               Travel
                              subaccount of
  < Travel
                              subaccount of
                                               Expenses
                            2009-10-13
Plane Ticket to SFO
                                      Checking Account
   Description
                      PTSFQ
   Date
   +$450.00
                                Air
                                             Travel•
                                                           Expenses
                      091013
   -$450.00
   Subaccount
                      $450.00
```

## Perspector

A "perspector" is notated:

x and y are operands (order matters).

o is an operator.

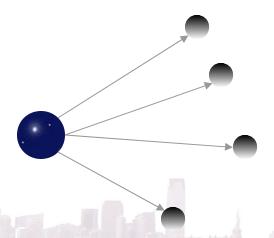
A perspector can represent a semantic relation, for example:

## From Multiple to Multiple Via One

#### SGML / XML

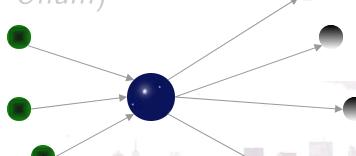
- One source,
- Multiple outputs

(Ex Uno Plures)



#### **DPM**

- Diverse inputs,
- One common representation,
- Multiple outputs
  (E Pluribus Plures Via Unum)



## **DPM** and RDF

- RDF is based on triples that express statements: subject – object – predicate
- RDF connects URIs
- RDF statements are not automatically reified.
- DPM is based on triples that express operations: x operand – operator – y operand
- DPM is not limited to URIs
- DPM perspectors are automatically reified.

## **DPM** and **Topic** Maps

- Topic Maps is a Navigation system using topics as nodes for representing subjects.
- Names, Types, Occurrences are topics connected through specific relationships.

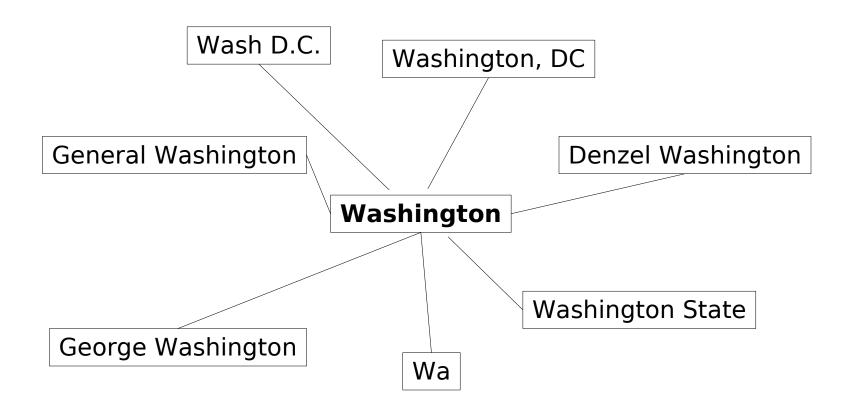
- DPM is a Navigation system based on nodes
- All nodes are related with other nodes.
- Topic Maps can be considered an application of DPM.

## Example: Name versus Subject

#### A Name does not identify a Subject:

- Variant names may be used to designate the same subject.
  - Synonyms
  - Typographical variations
- One name may identify several subjects.

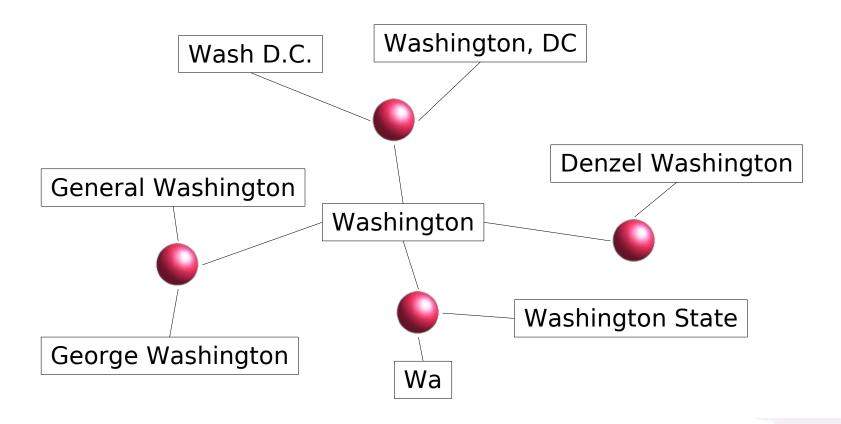
## **Names**



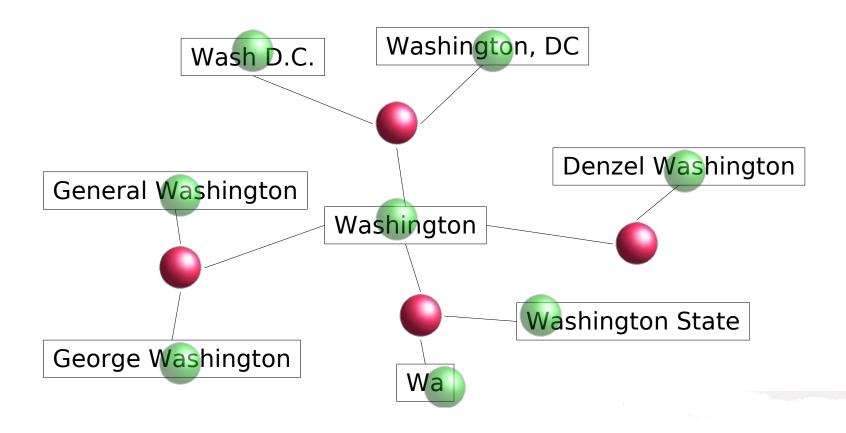
### **Names**

```
< Washington
               is an alternate name for
                                           Wash. D.C. >
 Washington
                                           Washington, DC >
                  an alternate name for
 Washington
                                           General Washington>
               is
                  an alternate name for
 Washington
                                           George Washington >
               is
                  an alternate name for
 Washington
                  an alternate name for
                                           Wa >
< Washington
                  an alternate name for
                                           Washington State >
< Washington
               is an alternate name for
                                           Denzel Washington >
```

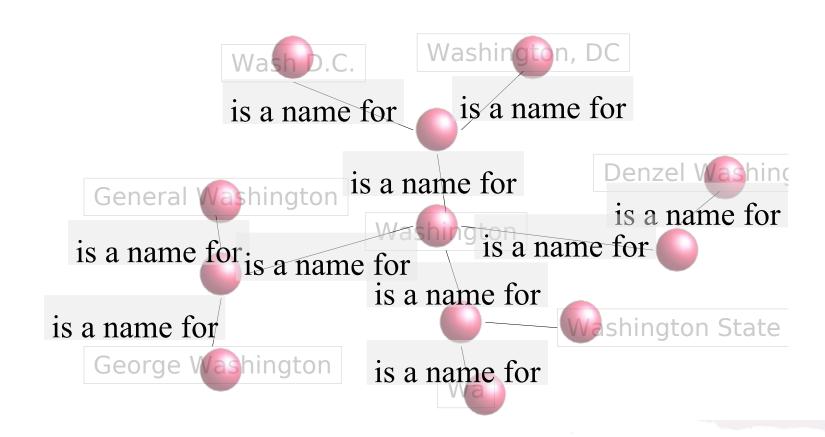
## **Emerging Subjects**



## Strings Become Subjects



## Generalization

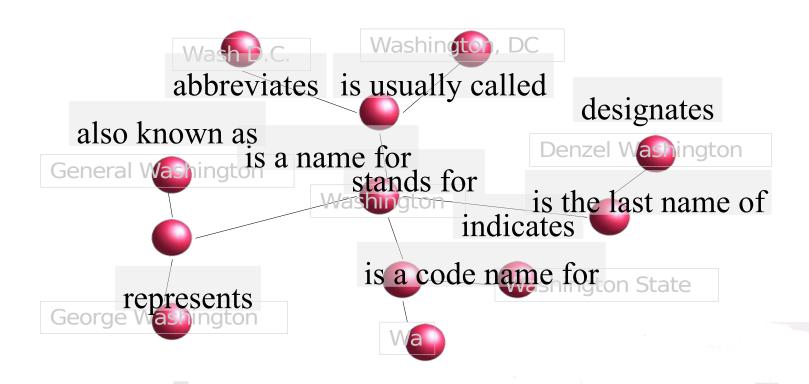


## Names and Subjects

```
< Washington
                   is a name for
                                      city of Washington >
                                      city of Washington >
 Washington DC
                   is a name for
 Wash. D.C.
                                      city of Washington >
                   is a name for
                                           General G Washington >
< Washington
                         is a name for
 General Washington
                                           General G Washington >
                           a name for
< George Washington
                                           General G Washington >
                         is a name for
                                         \overline{\text{Washington}} \overline{\text{S}}tate >
< Washington
                       is a name for
                                         Washington State >
< Wa
                          a name for
< Washington State
                       is a name for
                                         Washington State >
                                          Denzel Washington >
< Washington
                        is a name for
< Denzel Washington
                                         Denzel Washington >
                        is a name for
```

## Strings as Subjects

## Integration



## **Diversity**

- < city of Washington
- < Washington DC
- < Wash. D.C.
- < Washington
- < General G Washington
- < George  $\overline{W}a\overline{s}hington$
- < Washington
- < Wa
- < Washington State
- < Washington
- < Denzel Washington

is usually called indicates abbreviates is a name for also\_known\_as represents stands for is a code name for is a name for is last name of designates

```
Washington >
_city_of_Washington >
_city_of_Washington >
_city_of_Washington >
_General_G_Washington >
_General_G_Washington >
_Washington_State >
_Washington_State >
_Washington_State >
_Denzel_Washington >
_Denzel_Washington >
```

## Perspective on Naming

```
< city_of_Washington
< Washington DC
< Wash. D.C.
< Washington
< General G_Washington
< George Washington
< Washington
< Wa
< Washington State</pre>
```

< Denzel Washington

```
is named
is a name
           for
   a name for
   a name for
is named
   a name for
is
           for
   a name
          for
   a name
   a name for
           for
is
   a name
   a name for
```

```
Washington >
_city_of_Washington >
_city_of_Washington >
_city_of_Washington >
_General_G_Washington >
_General_G_Washington >
_Washington_State >
_Washington_State >
_Washington_State >
_Denzel_Washington >
_Denzel_Washington >
```

< Washington

## Multidimensional Information

```
is a name for
< New York
                                                 New York City
< New York
                                                 New York State
                             for
                   is a name
                                                 New York County
< New York
                   is a name for
< New York
                                                 Manhattan
                     a name for
< New York
                                                 Wall Street
                   is a name for
                   is an old name for
< New York
                                                 Manhattan
< Nueva York
                                                 New York City
                   is a name for
                                              New York City
  נו ׳ורק
                  a name for
                                                 English
< New York
                   is a name in the language
                                                 Spanish
< Nueva York
                     a name in the language
< New York
                      a name in the language
                                                 French
< English
                     a name for
                                                 English
< English
                   is a name in the language
                                                 English
< Anglais
                                                 English
                   is a name for
< Anglais
                   is a name in the language
                                                 French
  Inglés
                                                 English
                   is a name for
< Inglés
                   is a name in the language
                                                 Spanish
```

etc., etc.

## Why was the DPM created?

TaxMap is a topic map-based application, published every other week, containing IRS publications, forms, and FAQs, accessible by topics.

- Topics are extracted from XML document structure.
- Relations between topics are created using rules and by applying a semantic layer authored by tax experts.
- Rendition hides the creation process. Hence the question (among others): why are topics a and b related?

## **Operations on Names**

- TaxMap is built by a combination of automatic and manual processes. Names are added, modified, sometimes deleted, or regarded as synonyms.
- It's hard to know where a topic name comes from.

## TaxMap audited



Search All Topics:

GO

Search Help Navigation Help

Main Topics ABCDEFGHI JKLMNOPQR STUVWXYZ#

> FAQs Forms Publications Tele-Tax Topics

Comments About Tax Map Previous Tax Maps

#### **Income Earned Abroad**

Publications

Publications

Links Inside Publications

Publication 593 - Tax Hi Citizens and Residents Income Earned Abroad

Income Earned Abroad You exclusion from tax of a lim

#### **Related Topic Links**

- > Income
- ➤ U.S. Citizens and Resident

Aliens Abroad

#### **Income Earned Abroad**

Index

1	created	from document p593-001	>
2	related containment	Income	>
3	related PFT	Living Abroad	>
4	normalizes as	INCOME EARNED ABROA	D >
5	URL	ts0/tp6717.htm	>



## Where does "Living Abroad" come from?

ndex			
1	alternate name PFT	American citizens abroad	>
2 < U.S. Citizens and Resident Aliens Abroad	alternate name PFT		
3	privileged name is	U.S. Citizens and Resident Aliens Abroad	>
4 < Individuals Abroad	replaced PFT by		
5 < U.S. Citizen Abroad	replaced PFT by		
6 < U.S. citizens abroad	replaced PFT by		
7 < U.S. Citizens and Residents Living Outside the United States	replaced PFT by		
8 < U.S. Citizens or Residents Living Abroad	replaced PFT by		
9	related PFT	Bona Fide Residence Test	>
10 < Employment Abroad	related PFT		
11	related PFT	F2350	>
12	related PFT	Foreign Corporation	>
13	related PFT	Foreign Income	>
14 < Income Earned Abroad	related PFT		
15 < P516	related PFT		
16 < P54	related PFT		
17 < P593	related PFT		
18	related PFT	Physical Presence Test	>
19 < Retirees who were working abroad.	related PFT		
20 < Survivors of decendents who were working abroad.	related PFT		
21 < VISA Status	related PFT		
22	normalizes as	LIVING ABROAD	>
23	URL	ts0/tp16663.htm	>
24	topic type PFT	Key Topic	>

## **Containment Rule Results**

862	< Roth IRAs and traditional IRAs.	IRAs	>
863	< Roth IRAs and traditional IRAs.	Traditional II	RAs >
864	< Discount on Debt Instruments	Debt	>
865	< Hope and Lifetime Learning Credits	Credits	>
866	< Health Spa Expenses	Health	>
867	< Health Spa Expenses	Expenses	>
868	< Exceptions to reporting OID.	Exceptions.	>
869	< Exceptions to reporting OID.	OID	>
870	< SSN on correspondence.	SSN	>
871	< Capital Expense	Capital	>
872	< Home Mortgage	Home	If and tonic name
873	< Home Mortgage	Mortgage	If one topic name
874	< Maximum Exclusion	Exclusion	is entirely
875	< Income from property given to a child.	Income	contained
876	< Income from property given to a child.	Property	
877	< Income from property given to a child.	Child	into another one,
878	< Medicare Advantage MSAs	Medicare	they get
879	< Farm Income, cash method	Farm	5 0
880	< Farm Income, cash method	Farm Income	- Table 1
881	< Farm Income, cash method	Income	related.
882	< Farm Income, cash method	Cash Method	

## Synonyms created by Tax Experts

139 < Gambling
140 < Gift, Property
141 < Gift, Shares
142 < Gifts to Charity
143 < HCTC
144 < Help from IRS
145 < Hiring Employees
146 < Hobby Expenses
147 < Home Mortgage Interest
148 < Home Mortgage Interest
149 < Home Office
150 < Home Vacation
151 < Hope Credit
152 < Household Employment
153 < IC-DISC
154 < IRA
155 < IRA
156 < ITIN
157 < Important Reminder

İ	Wagering	>
ĺ	Property Received as a Gift	>
İ	Shares Acquired by Gift	>
İ	Charitable Contributions	>
İ	Health Coverage Tax Credit	>
İ	Taxpayer Assistance	>
İ	Hiring New Employees	>
İ	Activity not for profit	>
İ	Itemized Deduction, Home Mortgage	
	Interest	>
	Mortgage Interest, Home	>
	Business Use of Home	>
	Vacation Home	>
	Credit, Hope	>
ĺ	Employment, Household	>
İ	Interest Charge Domestic International	>
	Sales Corporation	_
	Individual Retirement Account	>
	Individual Retirement Arrangement	>
	Taxpayer Identification Number	>
	Reminders, Important	>



## **More Information**

Demos, other presentations available at:

http://www.infoloom.com

Michel Biezunski Infoloom

(718) 921-0901 mb@infoloom.com