

Taking Information Into Accounts

The Data Projection Model

Michel Biezunski

Infloom
mb@infloom.com
<http://www.infloom.com>

Michel Biezunski

- IT Consultant, dba Infoloom. Innovator, based in New York.
- Created the Topic Maps paradigm. Initiator of the ISO standard.
- Major current project: TaxMap, an electronic delivery tool for IRS publications and forms.
- Starts getting involved in Drupal applications.

Outline

- The Data Projection Model
 - Accounting and Information Management
- DPM, XML and the Semantic Web
- Application Examples

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?

Initial Goal

- Audit our own application by recording processes as well as semantic relations.
- Based on the realization that any relation can be expressed as a set of elementary binary relations.
- The Data Projection Model is about decomposing, and recomposing for viewing purposes.
- It is a “semantic microscope”.

What is the DPM about?

- Tracking background of any information item for:
 - Research purposes: where does a piece of information come from?
 - Search engines: Why a given item is on top of the hits?
 - Identity Theft Investigation: where is the leakage?
 - Privacy: Who knows what?
 - Managing targeted funding: How is this grant being spent?
 - Integrating heterogeneous sources.
 - Expressing multiple perspectives
 - Upgrading information systems
 - Etc.

Information and Accounting

- DPM similar to Double Entry Bookkeeping.
 - Information items are connected by links.
 - Money flows between accounts.
- Microscope:
 - Any relation can be expressed as a set of binary relations.
 - Accounting amounts to describe transactions between accounts.
- Every information item can be seen as an account.
 - “Account Statement”: all links to and from it.

Views

- Information networks, once decomposed as binary transactions, are extremely verbose and detailed.
(Too much information)
- Filtering what's useful for a given usage is a matter of perspective.
- Multiple perspectives can co-exist over the same information base.

Flattening the world

- Perspectives are defined according to projections.
- Perspectives express ways 3-dimensional space is rendered into 2-dimensional, i.e. projections.

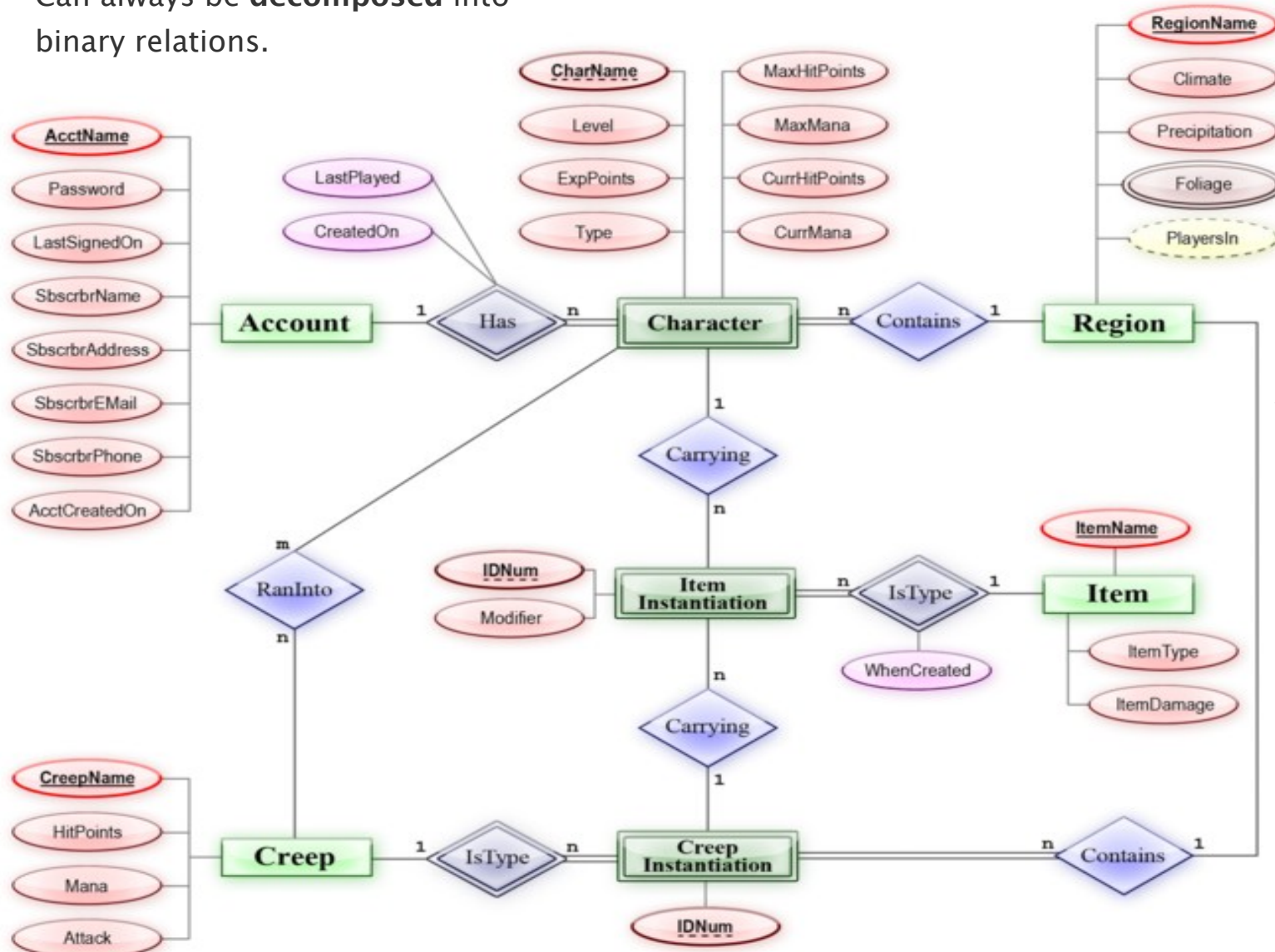


Real World Information:

- 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**.
- **Perspective** is what accompanies projection from n-ary relations to binary relations.

Entity-Relationship Model

Can always be **decomposed** into binary relations.



A simple entity relationship model.
[Http://en.wikipedia.org/wiki/Entity-relationship_model](http://en.wikipedia.org/wiki/Entity-relationship_model)



Relation with Accounting

- DPM:
 - No piece of information exists without at least one other piece of information connected.
 - Meaning comes from connections.
- Accounting:
 - Double Entry Bookkeeping: based on transactions between accounts.

Bookkeeping

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 is an increase in one account matched with a decrease in another account. (Accounts are balanced.)
- No money amount unaccounted for.

Double Entry Bookkeeping

***gnucash-demo - Checking Account**

File Edit View Transaction Actions Business Reports Tools Windows Help

Save Close + - Split

Accounts Travel Air Checking Account

Date	Num	Description	Transfer	R	Deposit	Withdrawal	Balance
10/13/2009		Opening Balance	Equity:Opening Balances	n	10,000.00		10,000.00
10/13/2009		Plane ticket to San Francisco	Expenses:Travel:Air	n		450.00	9,550.00
10/13/2009		Hotel in San Francisco	Expenses:Travel:Hotel	n		1,000.00	8,550.00
10/13/2009		October 2009	Income:Salary	n	5,400.00		13,950.00
10/13/2009	Num	Description	Transfer	n	Deposit	Withdrawal	Balance

Present: \$13,950.00 Future: \$13,950.00 Cleared: \$0.00 Reconciled: \$0.00 Projected Minimum: \$13,950.00

Tuesday 13 October 2009

Double Entry Bookkeeping & DPM

***gnucash-demo - Air**

File Edit View Transaction Actions Business Reports Tools Windows Help

Save Close [Icons] Split [Icons]

Accounts Travel **Air**

Date	Num	Description	Transfer	R	Expense	Rebate	Balance
10/13/2009		Plane ticket to San Francisco	:Current Assets:Checking Account	n	450.00		450.00
10/13/2009	Num	Description	Transfer	n	Expense	Rebate	Balance

Plane Ticket to SFO

2009-10-13

Checking Account

PTSFO
091013
\$450.00

Air

Travel

Expenses

Double Entry Bookkeeping & DPM

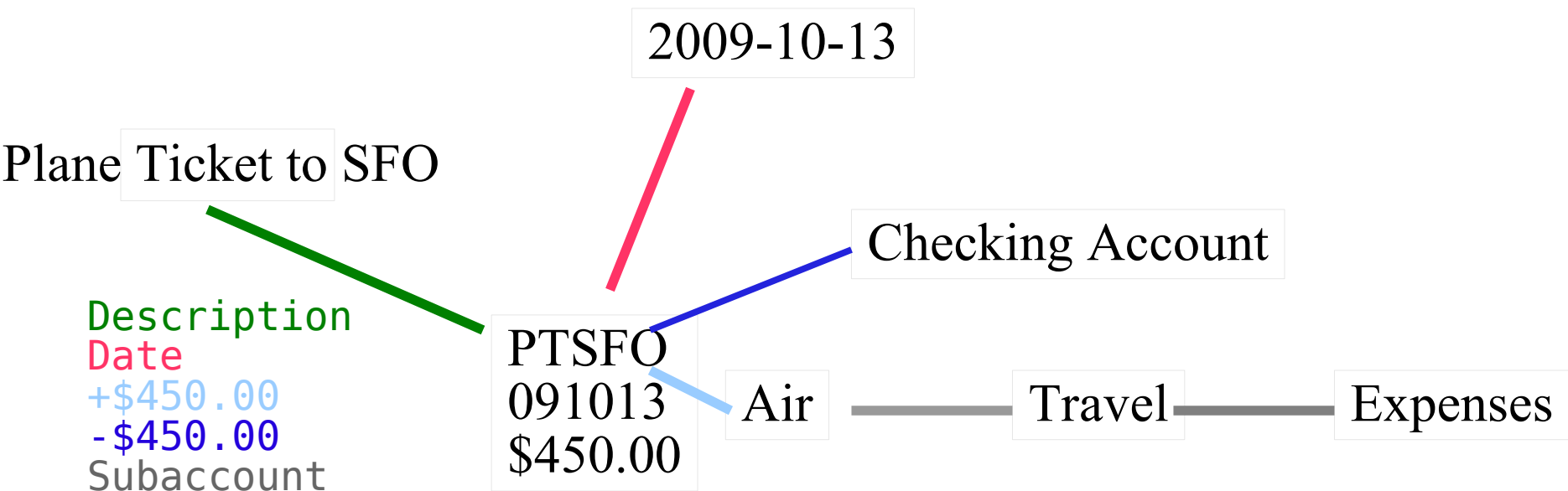
***gnucash-demo - Air**

File Edit View Transaction Actions Business Reports Tools Windows Help

Save Close [Icons] Split [Icons]

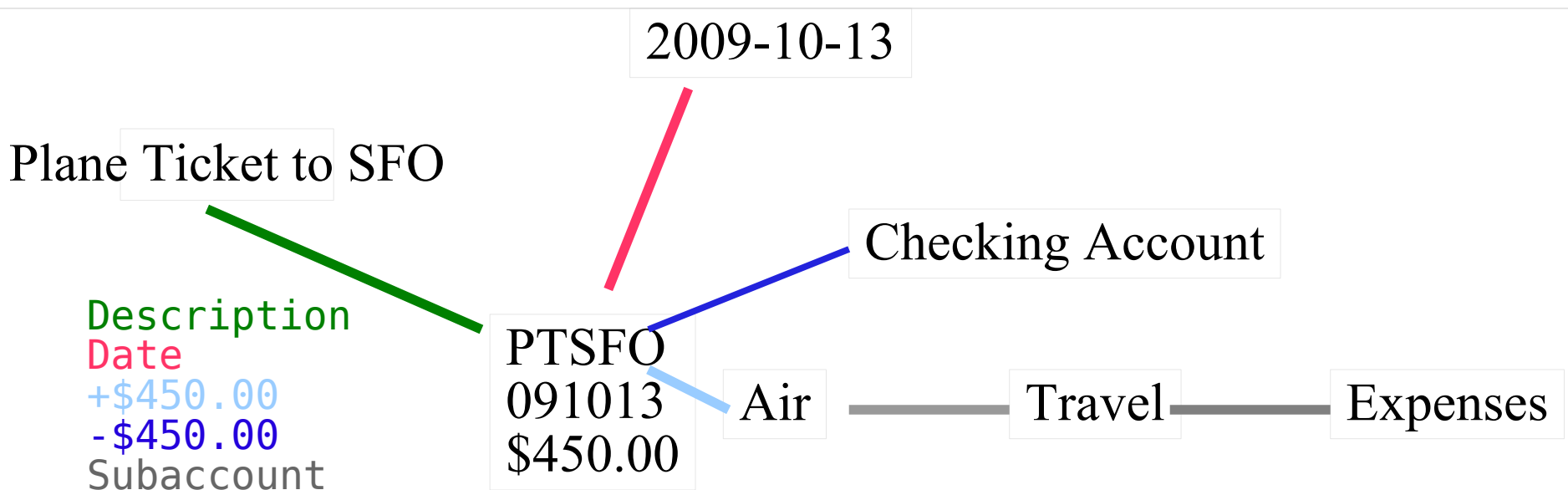
Accounts Travel **Air**

Date	Num	Description	Transfer	R	Expense	Rebate	Balance
10/13/2009		Plane ticket to San Francisco	:Current Assets:Checking Account	n	450.00		450.00
10/13/2009	Num	Description	Transfer	n	Expense	Rebate	Balance



DPM Notation

< PTSFO_09-10-13_450.00	date	2009-10-13	>
< PTSFO_09-10-13_450.00	description	Plane ticket to SFO	>
< PTSFO_09-10-13_450.00	450.00	Air	>
< PTSFO_09-10-13_450.00	-450.00	Checking Account	>
< Air	subaccount of	Travel	>
< Travel	subaccount of	Expenses	>



Perspector

A “**perspector**” is notated:

$$< x \mid o \mid y >$$

x and y are operands (order matters).

o is an operator.

A perspector can represent a semantic relation, for example:

$$< \text{New York} \mid \text{is a} \mid \text{city} >$$

(This is an instance/class relationship)

or $< \text{city} \mid \text{added to the system by} \mid \text{MB} >$

(This is usually considered metadata).

DPM and XML: Angle brackets

< Washington | is a | city >

can be written:

```
<perspector id="p1">  
<x>Washington</x><o>is a</o><y>city</y>  
</perspector>
```

This is XML.

DPM & XML: Conceptual

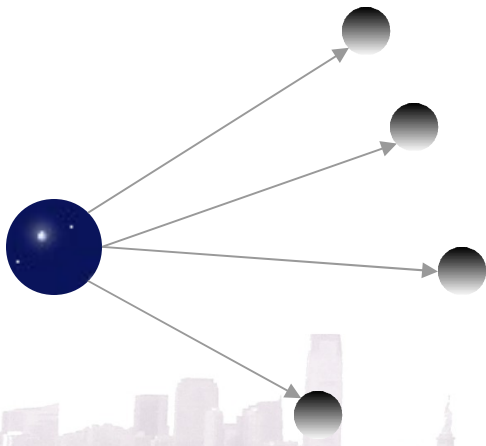
- Binary Relations:
Decompose into pieces / Recompose into views
- SGML & XML
Separate presentation from content. Apply styles afterwards.

From Multiple to Multiple Via One

SGML / XML

- One source,
- Multiple outputs

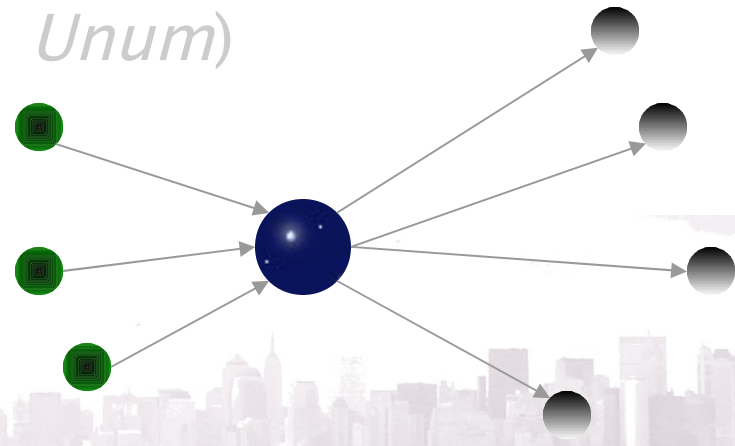
(Ex Uno Plures)



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

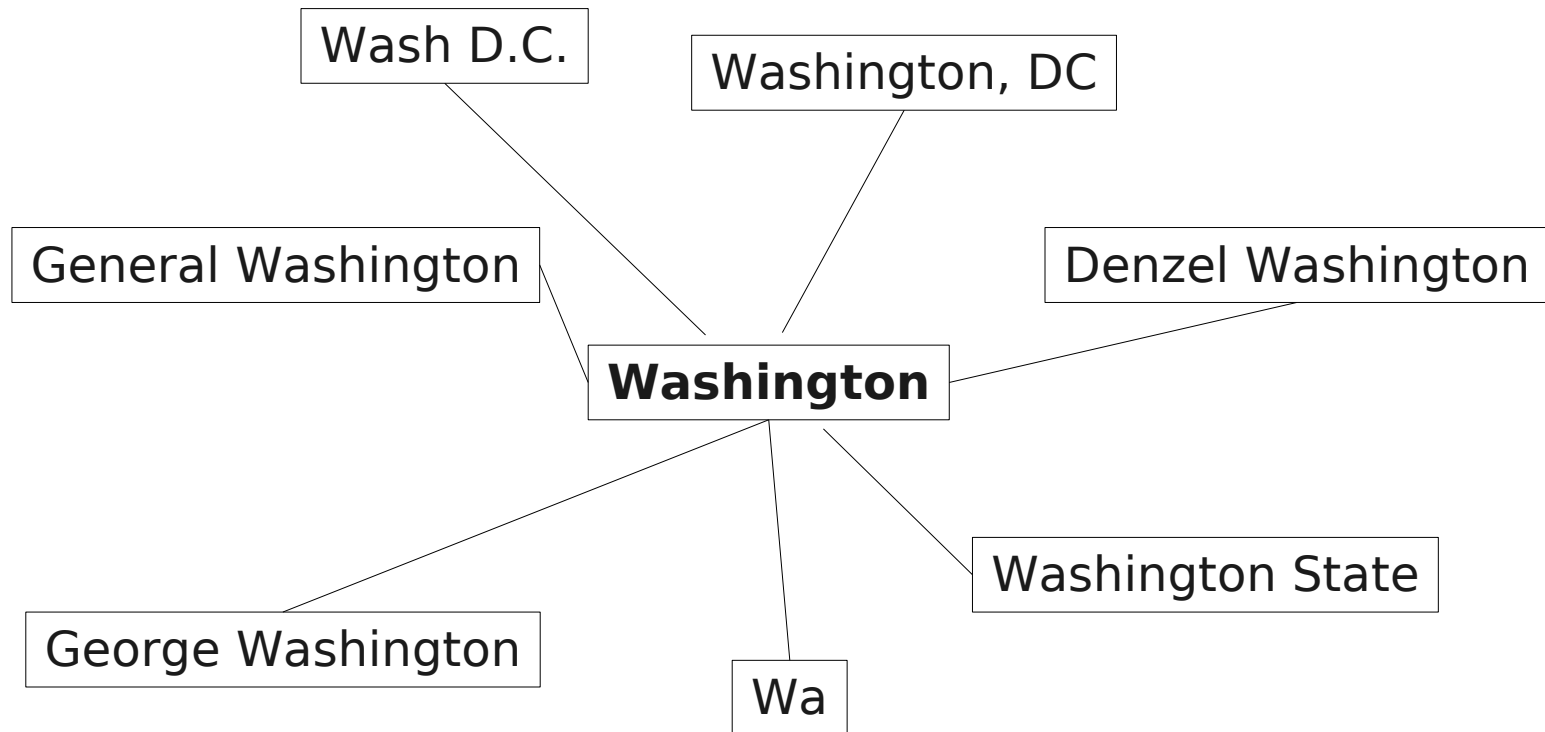
- Topic Maps is a Navigation system using topics as nodes for representing subjects.
- Names, Types, Occurrences are topics connected through specific relationships.
- DPM is a Navigation system based on nodes
- All nodes are related with other nodes.
- Topic Maps can be considered an application of 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.

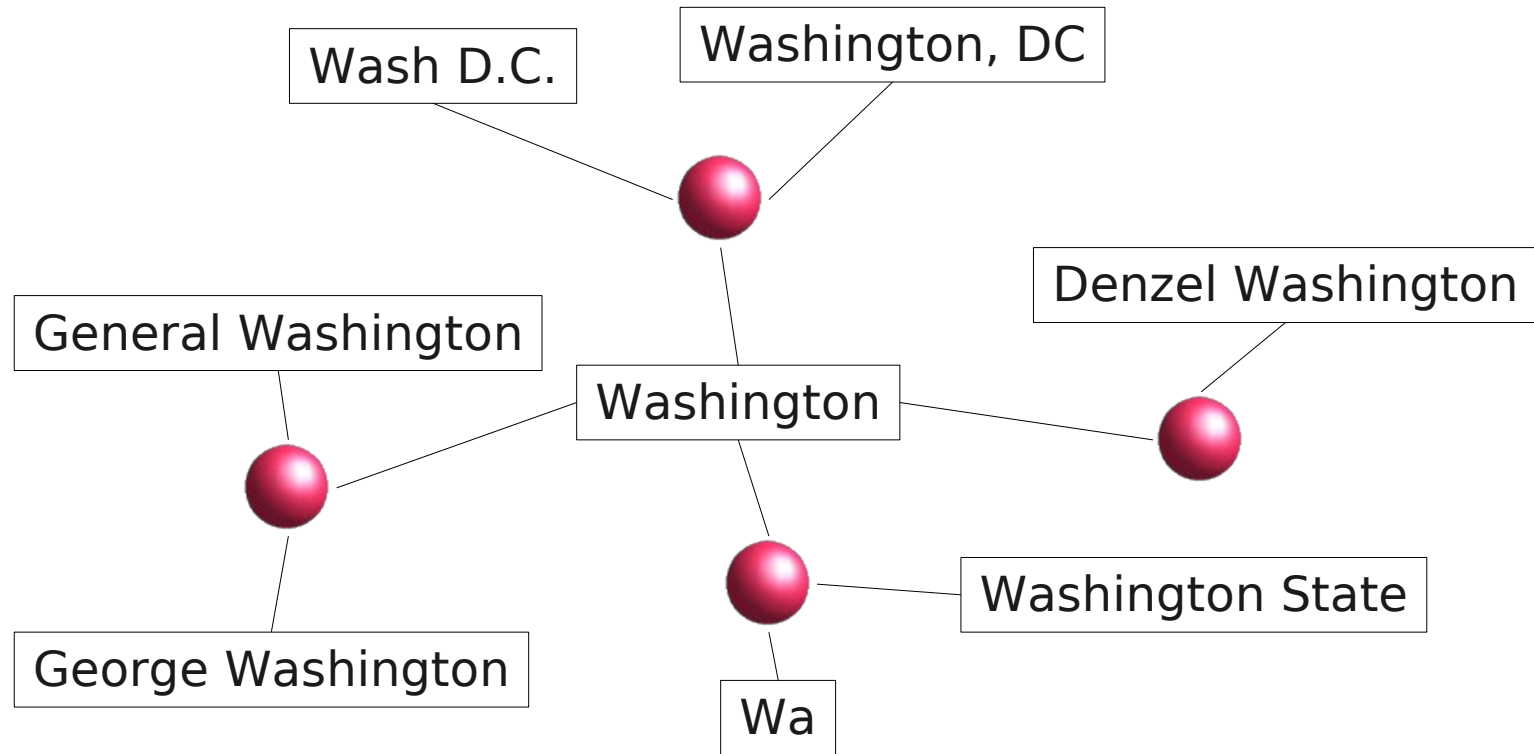
Names



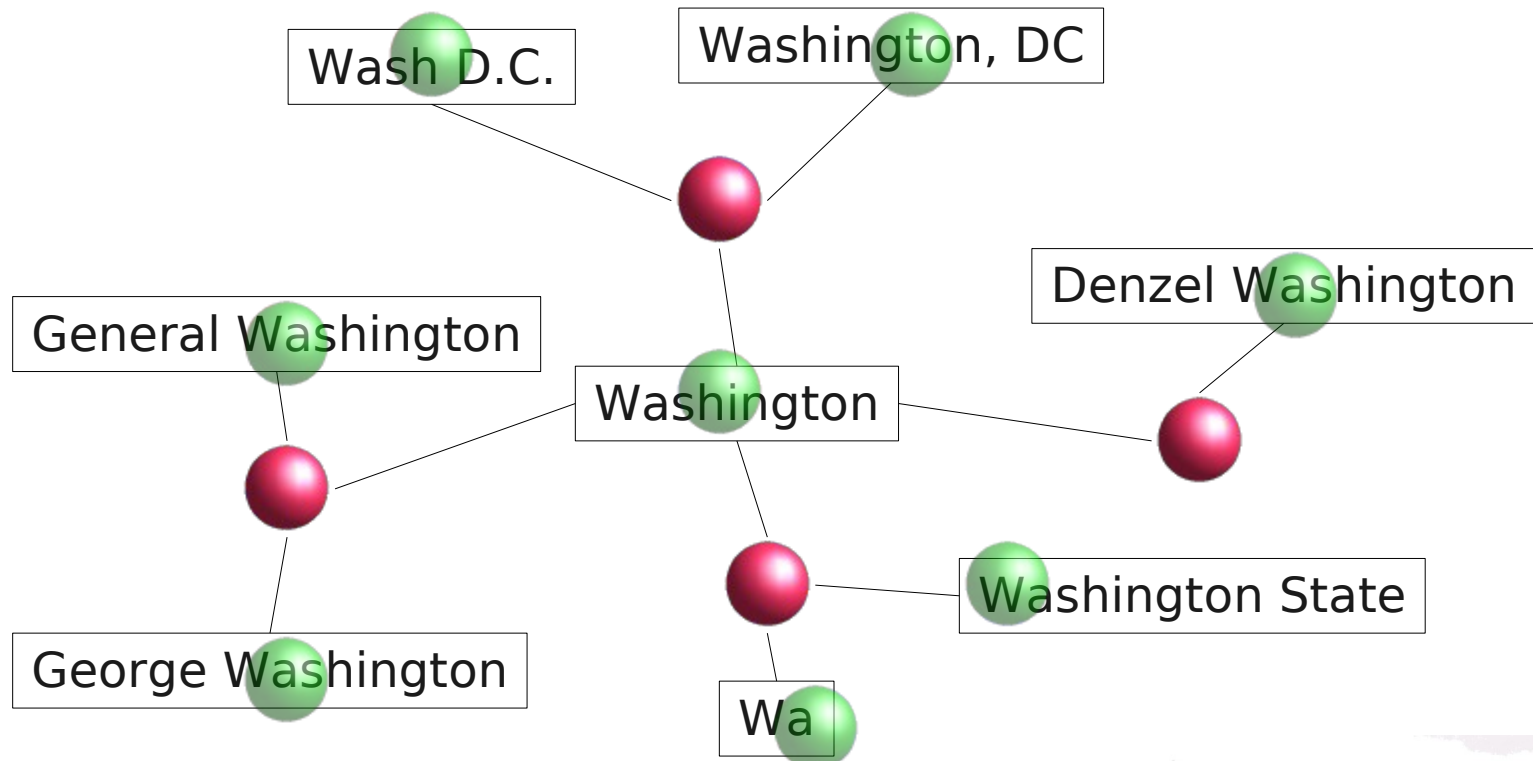
Names

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

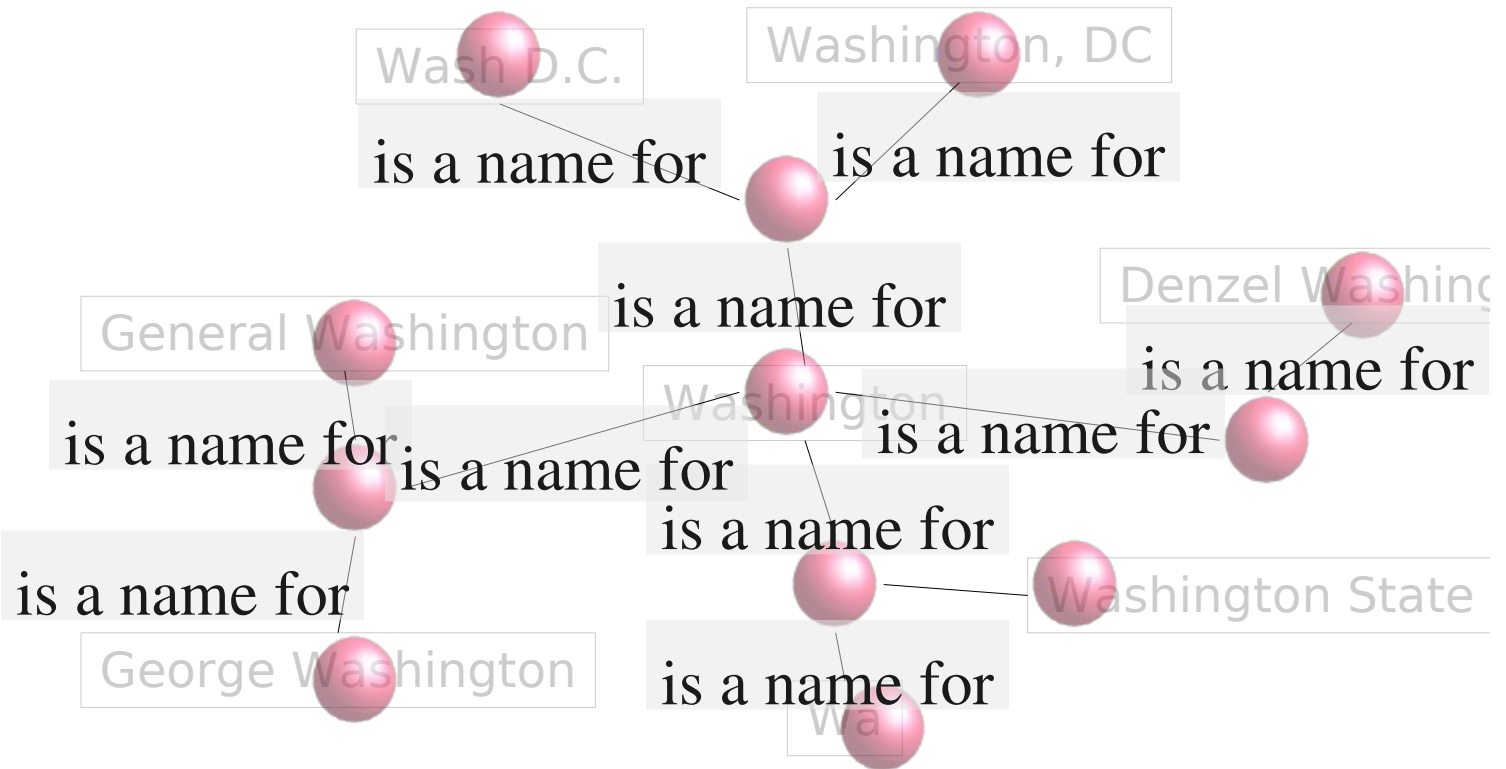
Emerging Subjects



Strings Become Subjects



Generalization



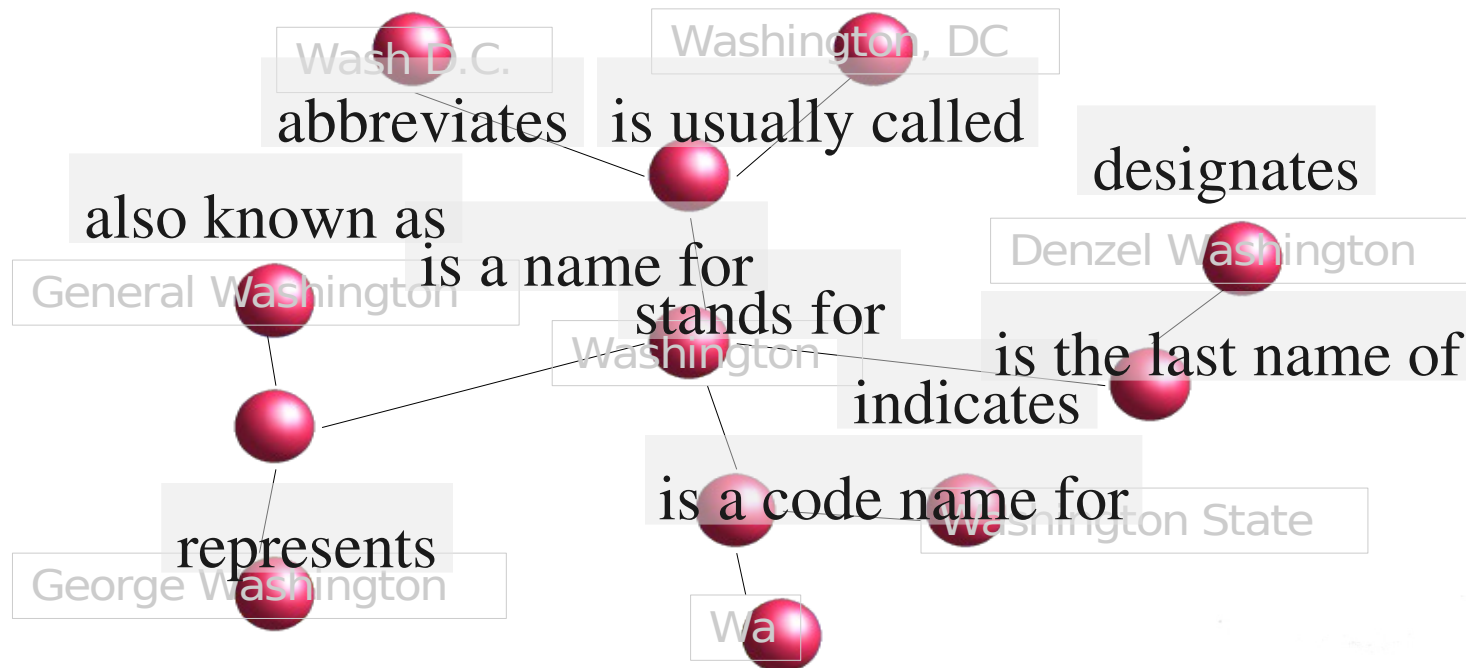
Names and Subjects

< Washington		is a name for		_city_of Washington >
< Washington DC		is a name for		_city_of_Washington >
< Wash. D.C.		is a name for		_city_of_Washington >
< Washington		is a name for		_General_G Washington >
< General Washington		is a name for		_General_G_Washington >
< George Washington		is a name for		_General_G_Washington >
< Washington		is a name for		_Washington_State >
< Wa		is a name for		_Washington_State >
< Washington State		is a name for		_Washington_State >
< Washington		is a name for		_Denzel Washington >
< Denzel Washington		is a name for		_Denzel_Washington >

Strings as Subjects

< Washington		is in character set		UTF-8	>
< Washington		is a name for		_city_of Washington	>
< Washington		is a name in the		_language English	>

Integration



Diversity

< city of Washington
< Washington DC
< Wash. D.C.
< Washington
< General G Washington
< George Washington
< 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 >
General G Washington
General Washington >
General G Washington
Washington State >
Washington State >
Washington State >
Denzel Washington >
Denzel Washington >

Perspective on Naming

< _city_of Washington	← is named →	Washington >
< W̄ashington DC	is a name for	_city_of Washington >
< Wash. D.C.	is a name for	_city_of Washington >
< Washington	is a name for	_General G Washington >
< General G Washington	← is named →	General W̄ashington >
< George W̄ashington	is a name for	General G Washington >
< Washington	is a name for	_Washington State >
< Wa	is a name for	_Washington State >
< Washington State	is a name for	_Washington State >
< Washington	is a name for	_Denzel Washington >
< Denzel Washington	is a name for	_Denzel Washington >

Multidimensional Information

< New York	is a name for	New York City	>
< New York	is a name for	New York State	>
< New York	is a name for	New York County	>
< New York	is a name for	Manhattan	>
< New York	is a name for	Wall Street	>
< New York	is an old name for	Manhattan	>
< Nueva York	is a name for	New York City	>
< ניו יורק	is a name for	New York City	>
< New York	is a name in the language	English	>
< Nueva York	is a name in the language	Spanish	>
< New York	is a name in the language	French	>
< English	is a name for	English	>
< English	is a name in the language	English	>
< Anglais	is a name for	English	>
< Anglais	is a name in the language	French	>
< Inglés	is a name for	English	>
< Inglés	is a name in the language	Spanish	>

etc., etc., etc., etc., etc., etc., etc., etc., etc., etc., etc., etc.

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
A B C D E F G H I
J K L M N O P Q R
S T U V W X Y Z #

FAQs
Forms
Publications
Tele-Tax Topics

Comments
About Tax Map
Previous Tax Maps

Income Earned Abroad

☐ Publications

☐ Publications

Links Inside Publications

Publication 593 - Tax Highlights for U.S.
Citizens and Residents
Income Earned Abroad

[Income Earned Abroad You
exclusion from tax of a limi](#)

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 ABROAD	>
5	URL	ts0/tp6717.htm	>

Search All Topics:

GO

Search Help
Navigation Help

Main Topics
A B C D E F G H I
J K L M N O P Q R
S T U V W X Y Z #

FAQs
Forms
Publications
Tele-Tax Topics

Comments
About Tax Map
Previous Tax Maps

Forms Publications FAQs TeleTax

Forms and Instructions

Form 2350 Application for Extension of Time to File U.S. Income Tax Return [MORE](#)

Form 2555 Foreign Earned Income [MORE](#)
2555 Instructions

Publications

Related Publications

Publication 54 Tax Guide for U.S. Citizens and Resident Aliens Abroad [PDF](#) [MORE](#)

Publication 516 U.S. Government Civilian Employees Stationed Abroad [PDF](#) [MORE](#)

Publication 593 Tax Highlights for U.S. Citizens and Residents Going Abroad [PDF](#) [MORE](#)

Links Inside Publications

Publication 17 - Your Federal Income Tax - Filing Information

U.S. Citizens and Residents Living Outside the United States ...

Related Topic Links

- > Alien, Resident
- > Bona Fide Residence Test
- > Employment Abroad
- > Foreign Corporation
- > Foreign Income
- > Income Earned Abroad
- > International Issues
- > Physical Presence Test
- > U.S. Citizen
- > U.S. Citizens and Resident and Nonresident Aliens
- > VISA Status

Where does “Living Abroad” come from?

Living Abroad

Index

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 decedents 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 IRAs	>
864	< Discount on Debt Instruments	IRAs	
865	< Hope and Lifetime Learning Credits	Traditional IRAs	
866	< Health Spa Expenses	Debt	
867	< Health Spa Expenses	Credits	
868	< Exceptions to reporting OID.	Health	
869	< Exceptions to reporting OID.	Expenses	
870	< SSN on correspondence.	Exceptions.	
871	< Capital Expense	OID	
872	< Home Mortgage	SSN	
873	< Home Mortgage	Capital	
874	< Maximum Exclusion	Home	
875	< Income from property given to a child.	Mortgage	
876	< Income from property given to a child.	Exclusion	
877	< Income from property given to a child.	Income	
878	< Medicare Advantage MSAs	Property	
879	< Farm Income, cash method	Child	
880	< Farm Income, cash method	Medicare	
881	< Farm Income, cash method	Farm	
882	< Farm Income, cash method	Farm Income	

Synonyms created by Tax Experts

139 < Gambling	Wagering	>
140 < Gift, Property	Property Received as a Gift	>
141 < Gift, Shares	Shares Acquired by Gift	>
142 < Gifts to Charity	Charitable Contributions	>
143 < HCTC	Health Coverage Tax Credit	>
144 < Help from IRS	Taxpayer Assistance	>
145 < Hiring Employees	Hiring New Employees	>
146 < Hobby Expenses	Activity not for profit	>
147 < Home Mortgage Interest	Itemized Deduction, Home Mortgage Interest	>
148 < Home Mortgage Interest	Mortgage Interest, Home	>
149 < Home Office	Business Use of Home	>
150 < Home Vacation	Vacation Home	>
151 < Hope Credit	Credit, Hope	>
152 < Household Employment	Employment, Household	>
153 < IC-DISC	Interest Charge Domestic International Sales Corporation	>
154 < IRA	Individual Retirement Account	>
155 < IRA	Individual Retirement Arrangement	>
156 < ITIN	Taxpayer Identification Number	>
157 < Important Reminder	Reminders, Important	>

More Information

Demos, other presentations available at:
<http://www.infloom.com>

Michel Biezunski
Infloom

(718) 921-0901
mb@infloom.com