MICHEL BIEZUNSKI, AI CAMP, UN NEW YORK
JULY 12, 2016

TAKING INFORMATION INTO ACCOUNTS

The Data Projection Model

INTRO

- Information is abundant.
- Information is messy.
- There are multiple ways to view the same information.
- To be trusted, information needs to be accounted for.
- … Introducing the Data Projection Model

Michel Biezunski, Infoloom Inc., July 12, 2016. AI Camp, UN, New York

NAMES ARE A MATTER OF PERSPECTIVE

- What's in a name?
  - New York is a state, a city, a county
  - A world in itself. But that's another story…
  - New York City, The City of New York are synonyms of New York (as city).
  - NY is a synonym of New York (as state).
  - New York, New York synonym of Manhattan. But New York, New York is also a movie, and is also several songs, including from the movie New York, New York, the other one from On the Town
  - For old Brooklynites, "New York" means "Manhattan".
  - For contemporary Brooklynites, "The City" means "Manhattan".
    - However, "The City" is not really a synonym of Manhattan. It also means the financial district in London, etc.

INTRODUCING THE DATA PROJECTION MODEL

- Information is multidimensional.
- Computers use “flat information”.
- Flat information means information described as a set of binary relations.
- The process of flattening information is called “projection”
- The information resulting from a projection is viewed as a perspective.

Michel Biezunski, Infoloom Inc., July 12, 2016. AI Camp, UN, New York

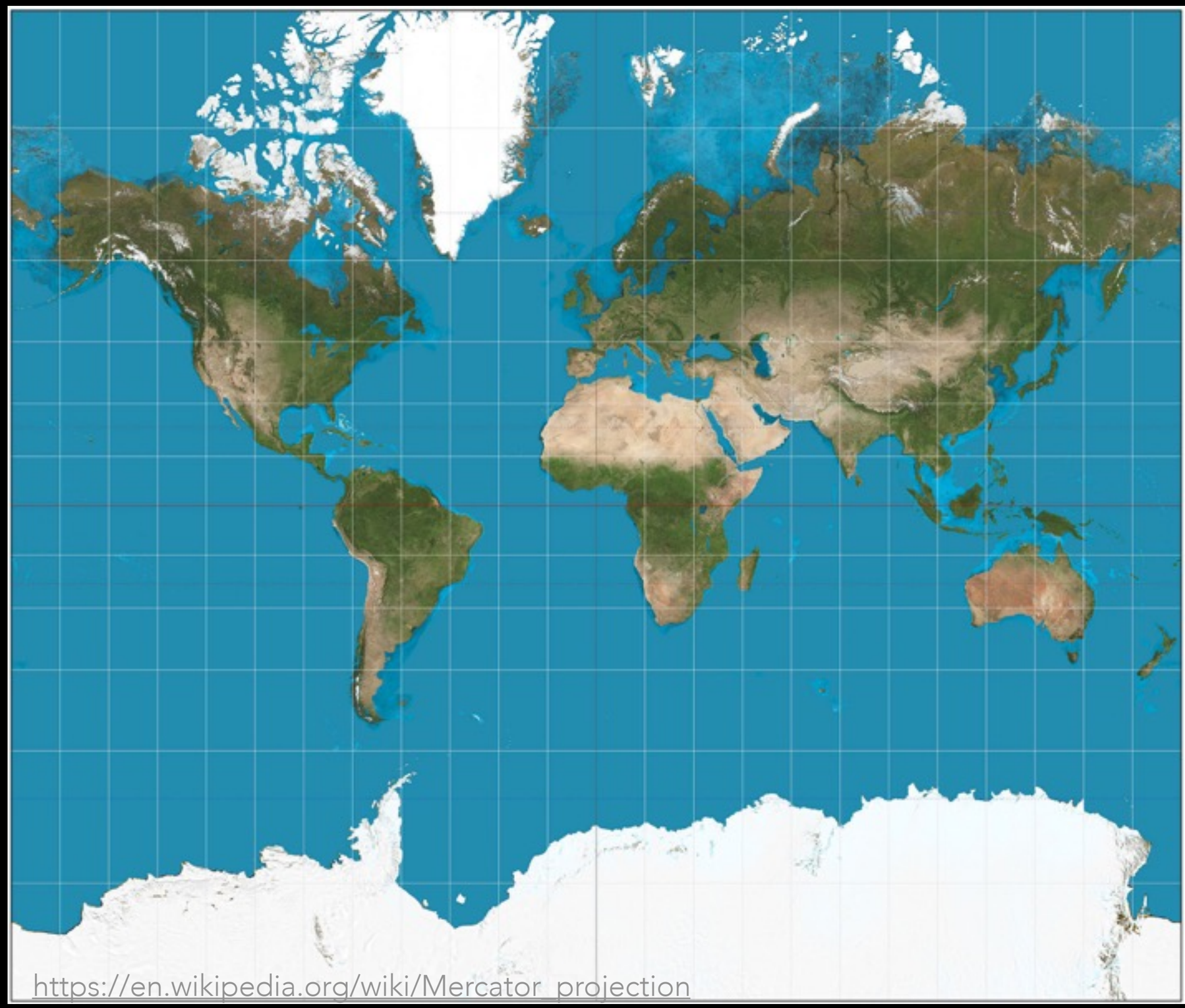
NOTHING NEW HERE...

- A projection is a method to describe a 3-dimensional world into a 2-dimensional representation.

- A perspective is a view that corresponds to a given projection.

- The world can be seen from multiple perspectives.

FLATTENED WORLD SPHERE



https://en.wikipedia.org/wiki/Mercator_projection

Mercator projection of the world between 82°S and 82°N

PERSPECTIVE

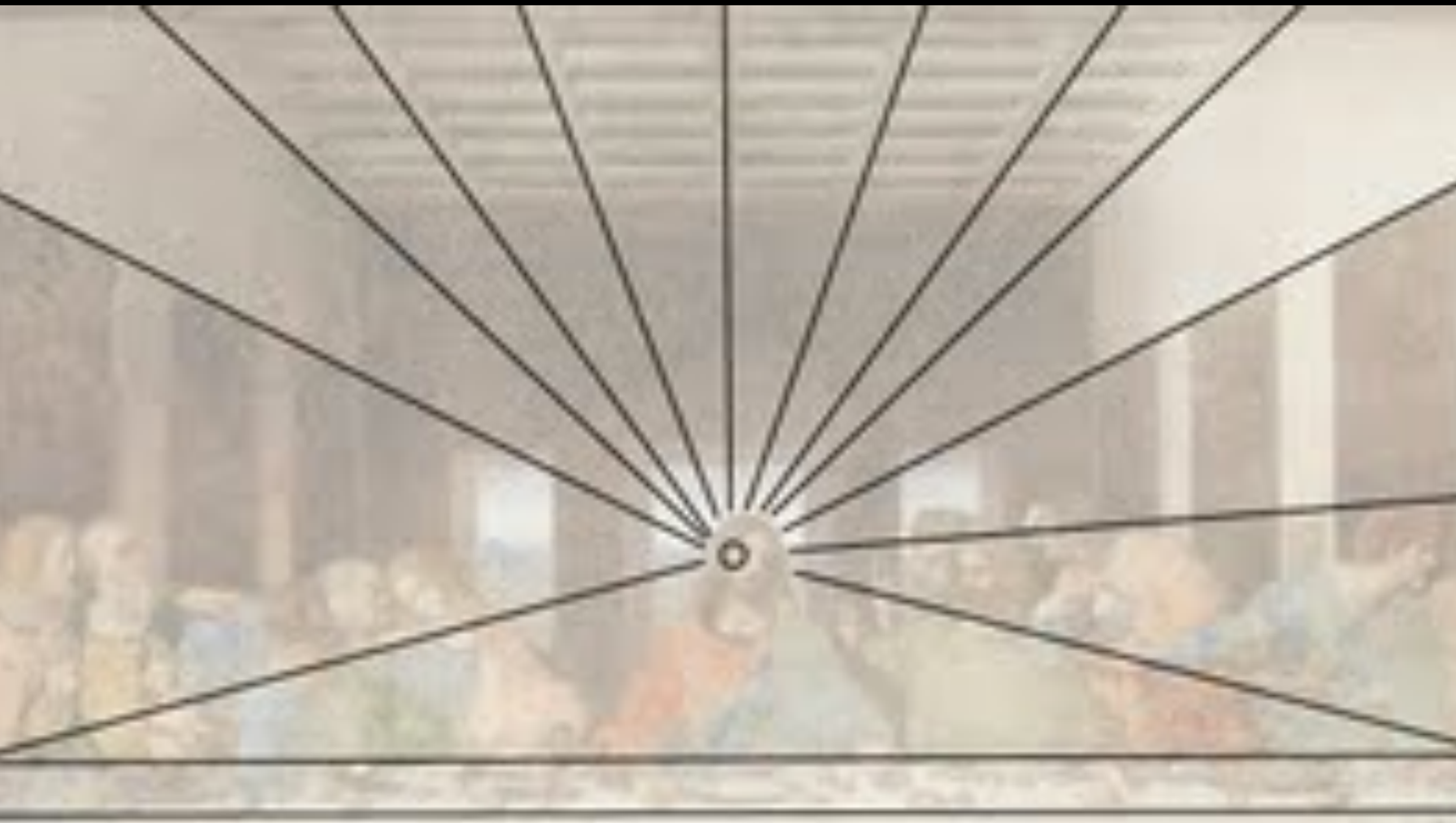
- Perspective is the art and mathematics of realistically depicting three-dimensional objects in a two-dimensional plane.

  Perspective -- from Wolfram MathWorld

  mathworld.wolfram.com/Perspective.htmlMathWorld

- Whether **Manhattan** should be called **New York**, or **The City**, or something else, is a matter of perspective.

Leonardo da Vinci was one of the innovators in perspective.



https://mathsimulationtechnology.wordpress.com/perspective/

Luca Pacioli

LUCA PACIOLI (C.1447-1517)

- Taught Geometry to Leonardo da Vinci.

  Worked on Perspective, and was on Leonardo's side in the controversy.

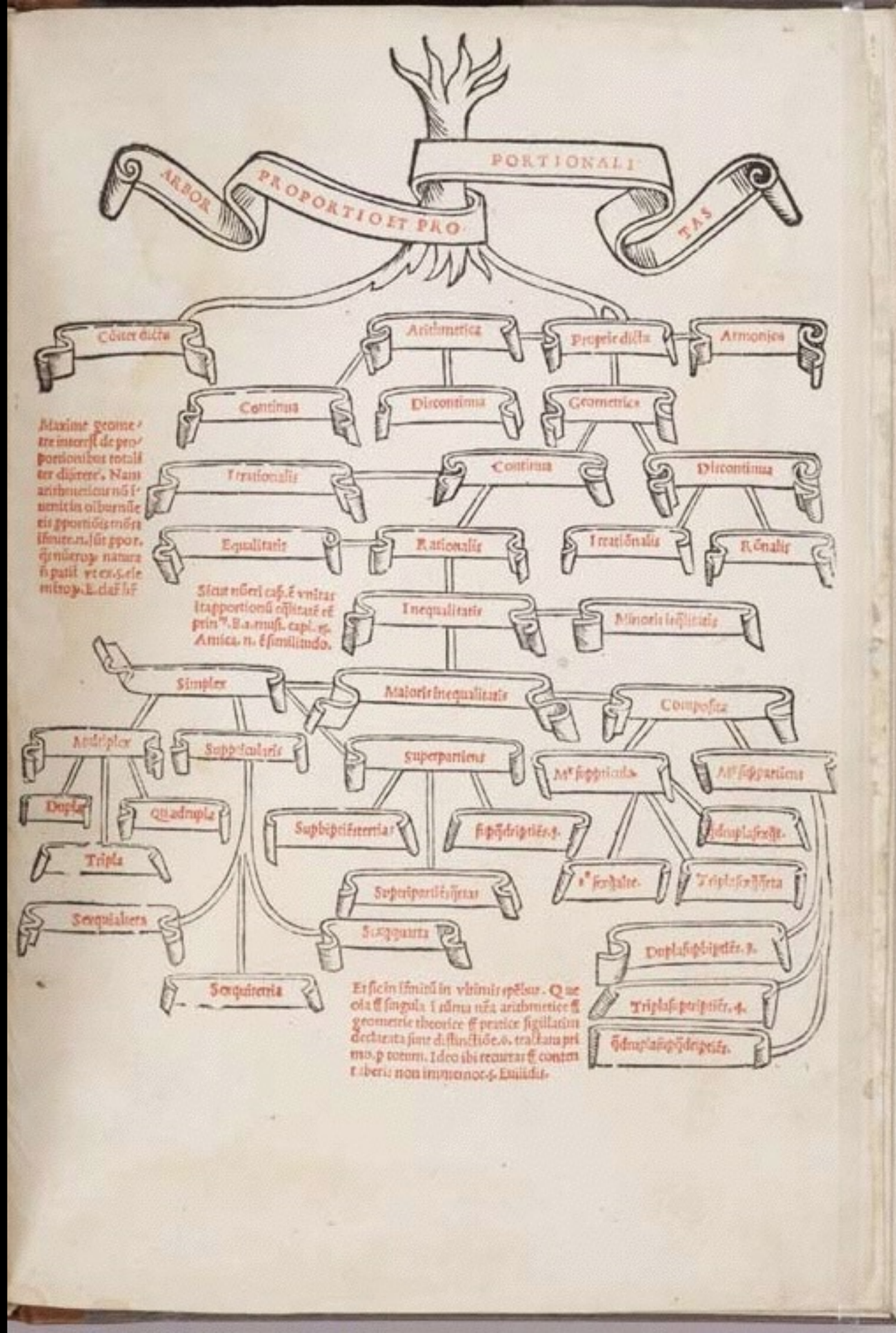
- Wrote on the ethics of accounting.

  "Invented" and popularized Double Entry Bookkeeping

THE TREE OF PROPORTION AND PROPORTIONALITY

- Taxonomy of mathematical concepts.

  Luca Pacioli, *De Divina Proportione*, 1509



DOUBLE ENTRY BOOKKEEPING

- Pacioli named “the father of accounting and bookkeeping.
- First published description of the practice of Venetian merchants.
- Every entry in an account has a corresponding entry in another account.
- Debit in one account corresponds to credit in the other account.
- Trust is based on accountability.

Michel Biezunski, Infoloom Inc., July 12, 2016. AI Camp, UN, New York

DATA PROJECTION MODEL

- Applying to Information the Double Entry Model used in Accounting.

- Multidimensional information is flattened using projections and resulting views are perspectives.

DATA PROJECTION MODEL: FLATTENING

- Any thing is an information object:
    - including:
        - names to represent things.
        - processes used to perform transformations.
        - any qualifier used
- In other words, metadata is no different from data.

Michel Biezunski, Infoloom Inc., July 12, 2016. AI Camp, UN, New York

DATA PROJECTION MODEL: ACCOUNTABILITY

- No information is isolated.
  For example, an information item has a name, a creator, information is created at a certain date.
- An information object resides in an "account" (aka "topic" in Topic Maps). An information account is related to many other information accounts.
- N-ary relations between accounts can be decomposed into a number of binary relations.

INFORMATION SYSTEM EXPLODED

- Expressing information objects in terms of binary relations with other information objects creates zillions of information objects. It's similar to seeing matter as its atomic components (or elementary particles).
- Complex relationships are simplified, but number of relations is exploding.

EXAMPLES

- Numbers

2 is the price of a drink.
2 is in US$
10 is the price of a meal
10 is in US$
2 plus 10

Strings described
as composition process

"N" is encoded in ASCII.
"N" is followed by "e"
"e" is followed by "w"
"w" is followed by "space"
etc.

- Semantic relationships

"Brooklyn" is located in "New York"
"Brooklyn" is part of "New York City"
"New York City" is in "New York"
"New York" is a "state"
"New York" is a "city"

DATA PROJECTION MODEL: PERSPECTORS

A **perspector** is notated:

`p = < x | o | y >`

`x` and `y` are operands (order matters).

`o` is an operator.

A perspector can represent a semantic relation, for example:

`< New York | is a | city >`

or a process connected with workflow:

`< city | added by | MB >`

(This is usually considered metadata).

EXAMPLE: NAMES VS. SUBJECTS

- A name does not identify a subject
    - Multiple names may be used to designate the same subject:
        - Synonyms
        - Typographical variations
        - Multiple languages
- One name may identify several subjects

NAMES

Washington DC

DC

Denzel Washington

Washington

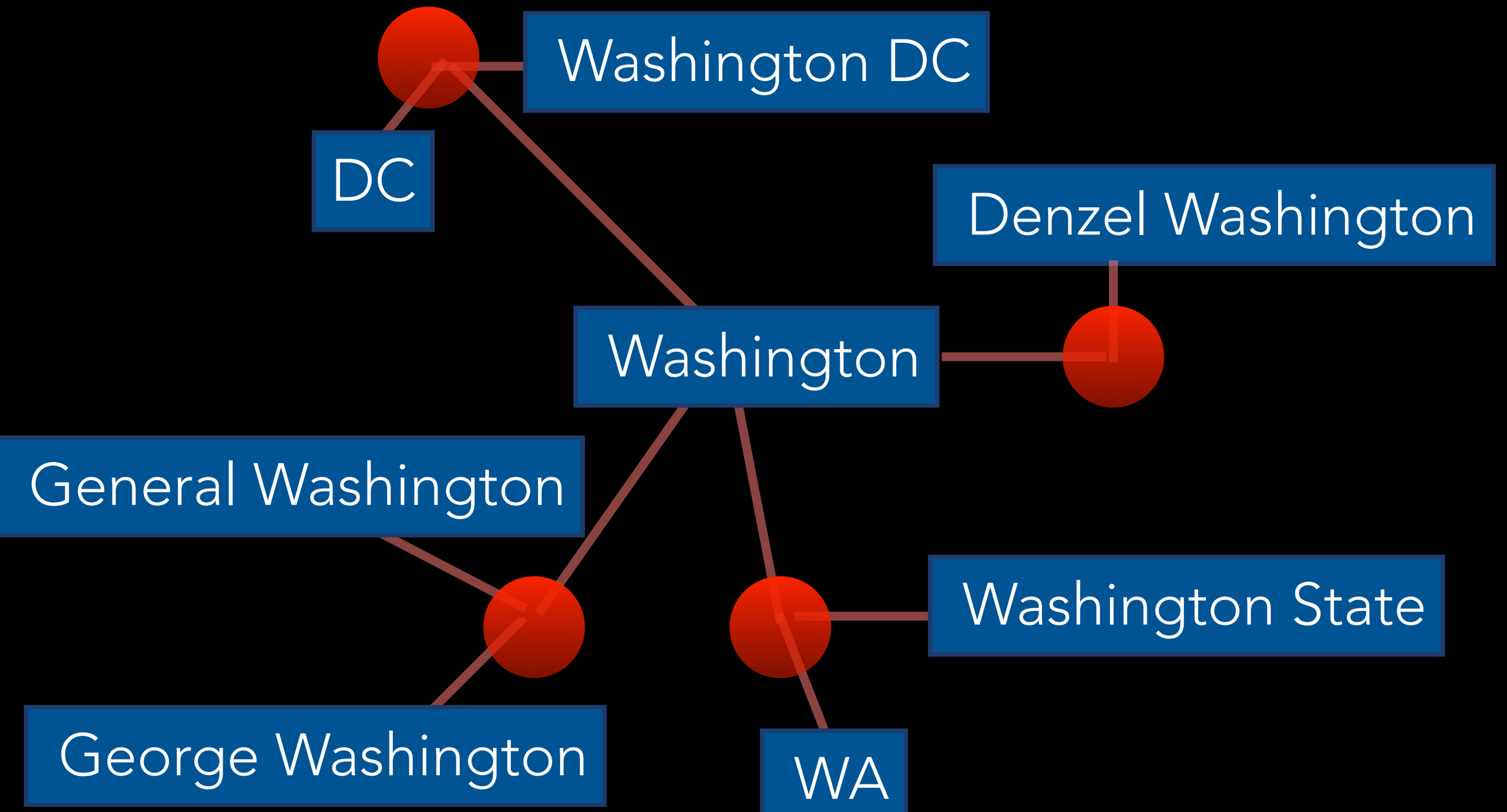
General Washington

Washington State

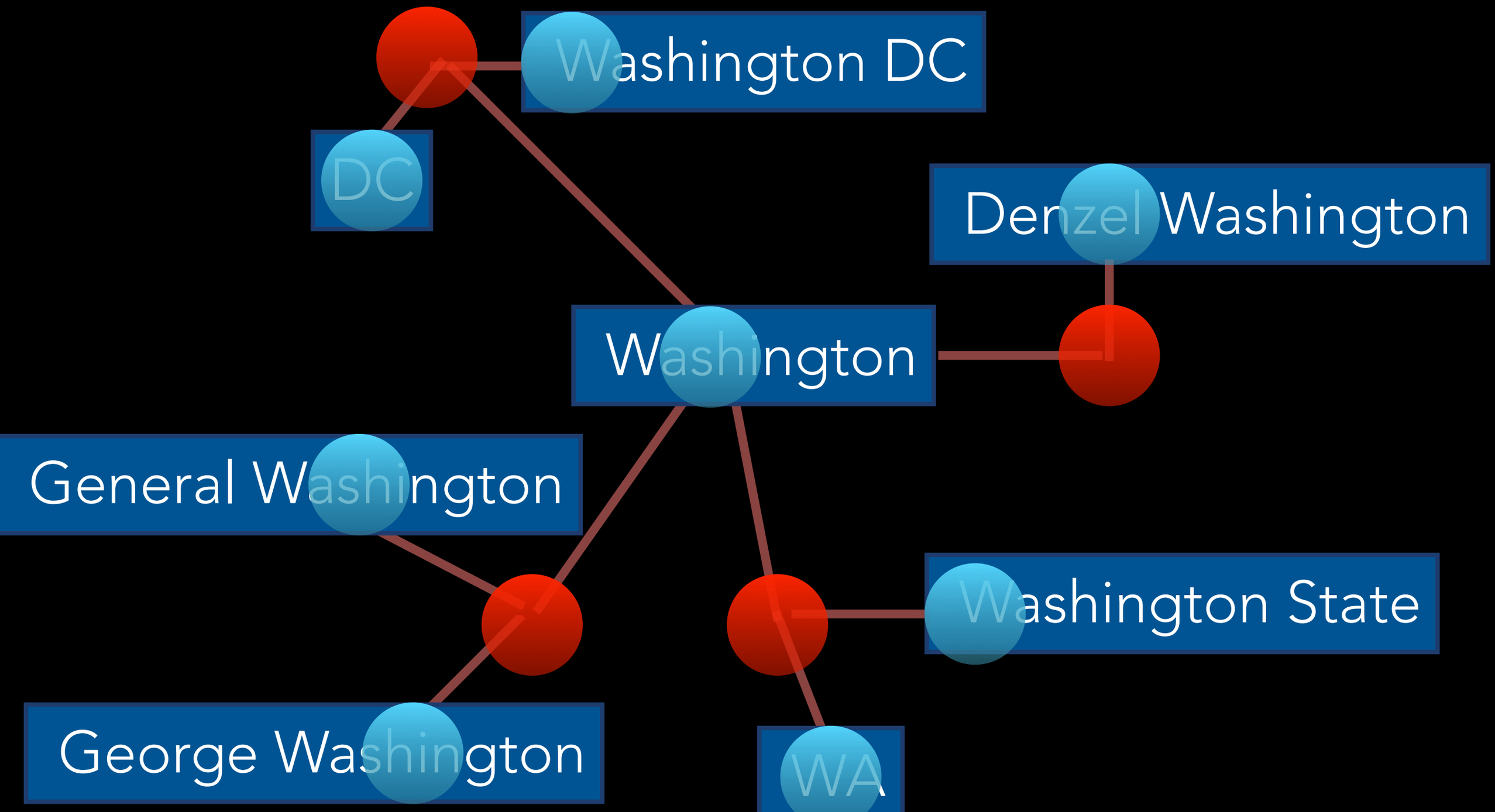
George Washington

WA

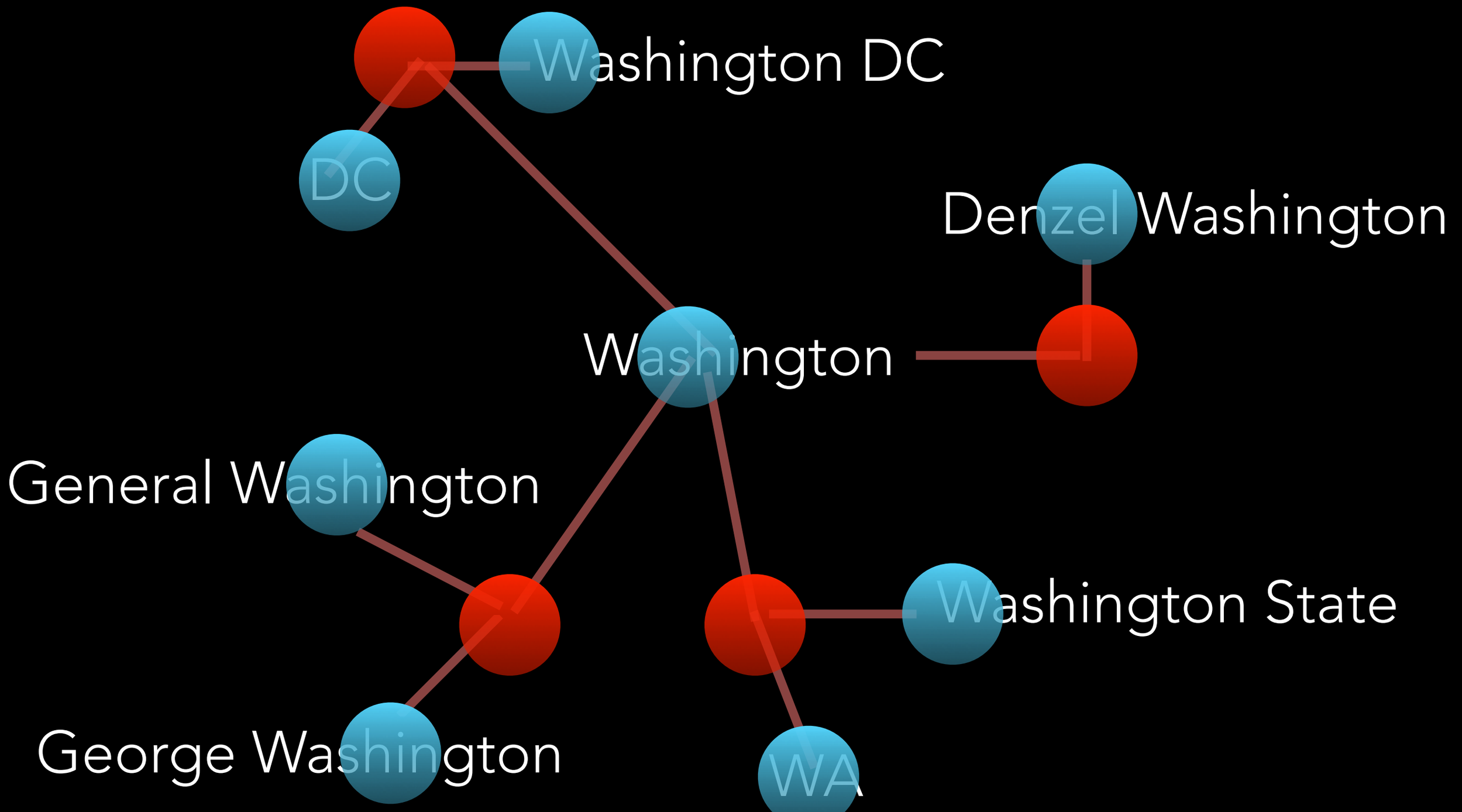
EMERGING SUBJECTS



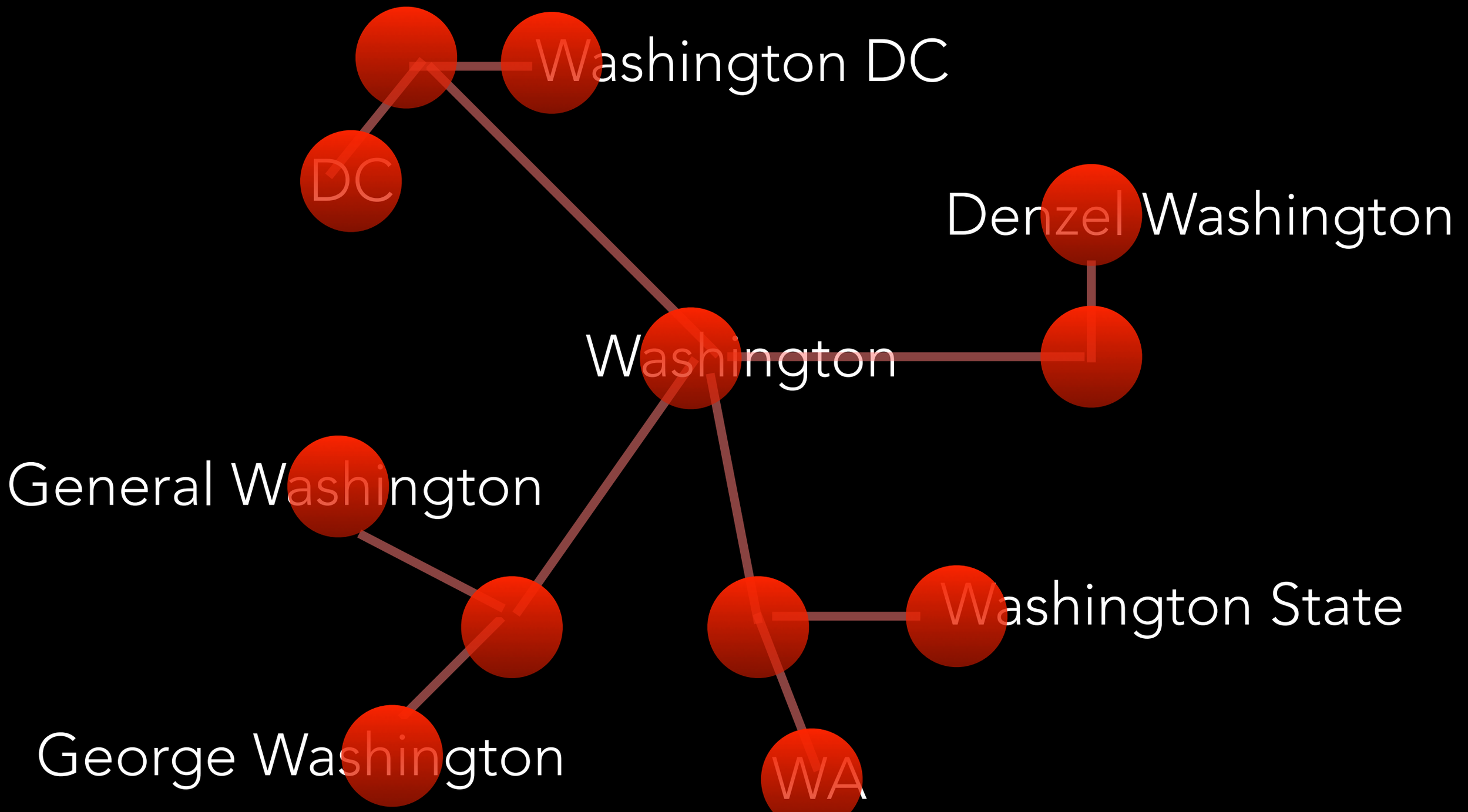
STRINGS BECOME SUBJECTS



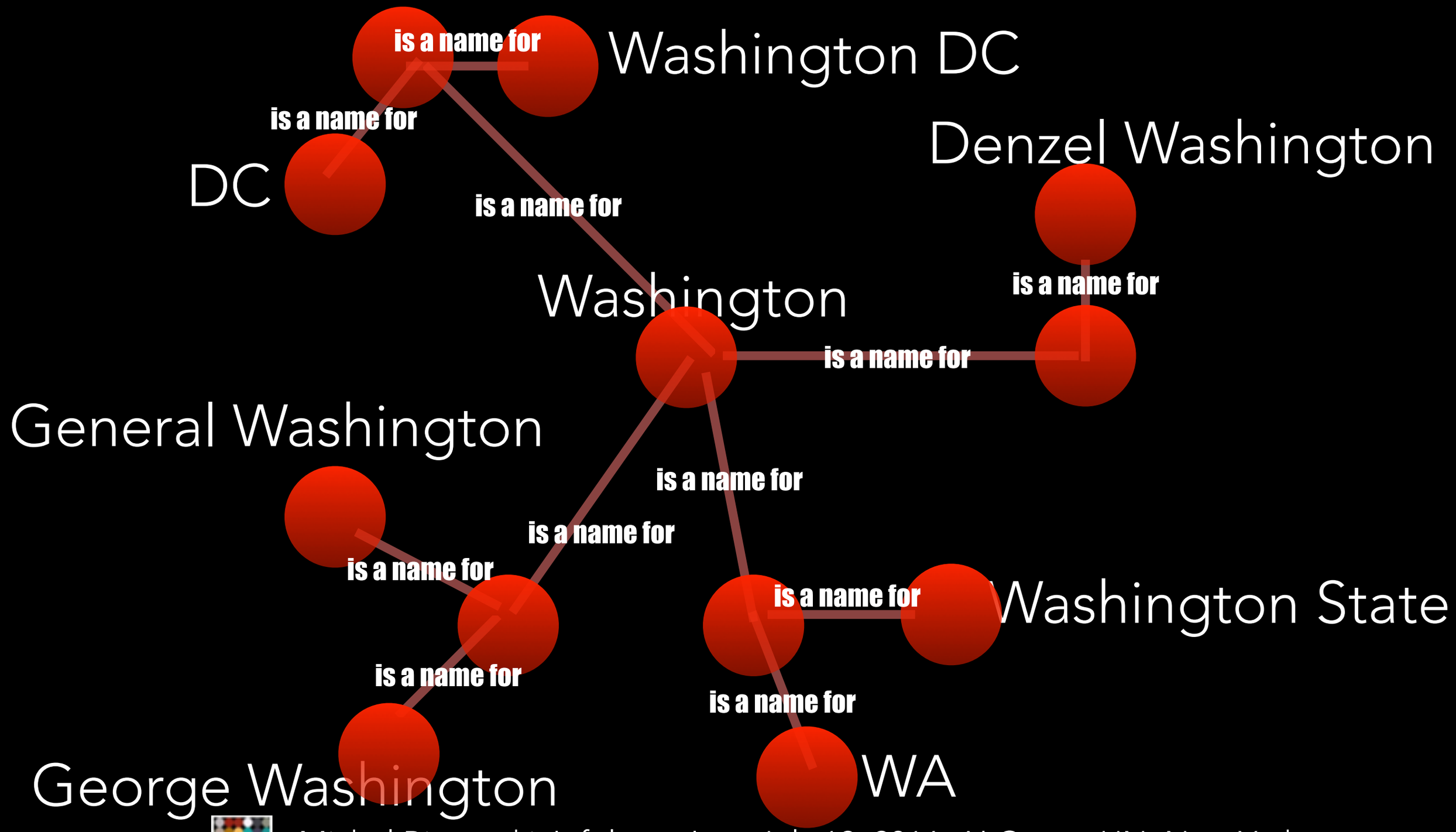
STRINGS BECOME SUBJECTS



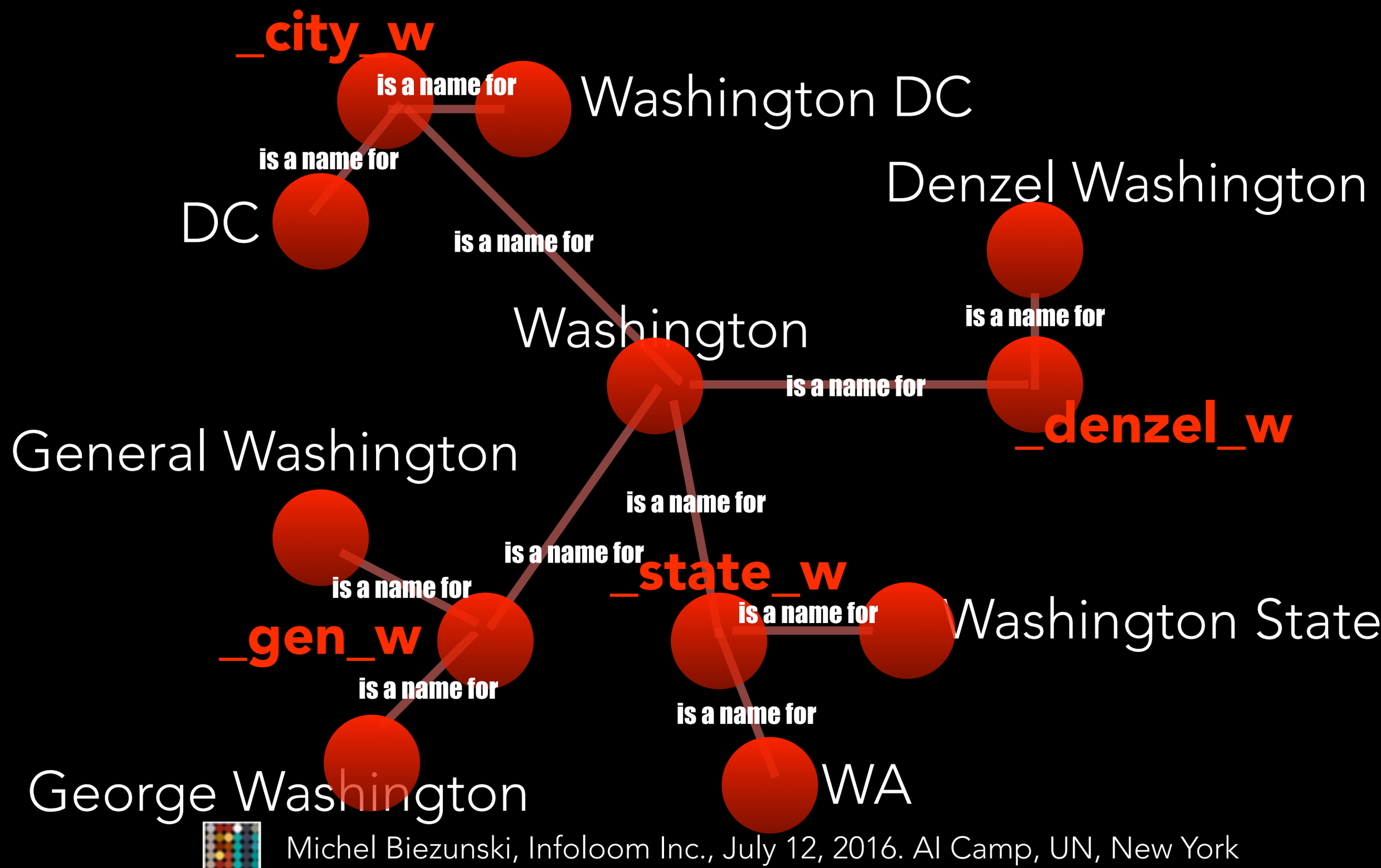
STRINGS BECOME SUBJECTS



RELATIONS BETWEEN SUBJECTS



FLATTENING


_city_w
is a name for
Washington DC
is a name for
DC
is a name for
Denzel Washington
is a name for
Washington
is a name for
_denzel_w
General Washington
is a name for
is a name for
_state_w
is a name for
Washington State
is a name for
_gen_w
is a name for
is a name for
George Washington
WA

PERSPECTORS

<	Washington	|	is a name for	|	_city_w	>
<	Washington	|	is a name for	|	_gen_w	>
<	Washington	|	is a name for	|	_state_w	>
<	Washington	|	is a name for	|	_denzel_w	>
<	Washington DC	|	is a name for	|	_city_w	>
<	DC	|	is a name for	|	_city_w	>
<	Denzel Washington	|	is a name for	|	_denzel_w	>
<	Washington State	|	is a name for	|	_state_w	>
<	WA	|	is a name for	|	_state_w	>
<	General Washington	|	is a name for	|	_gen_w	>
<	George Washington	|	is a name for	|	_gen_w	>

DIGGING DEEPER

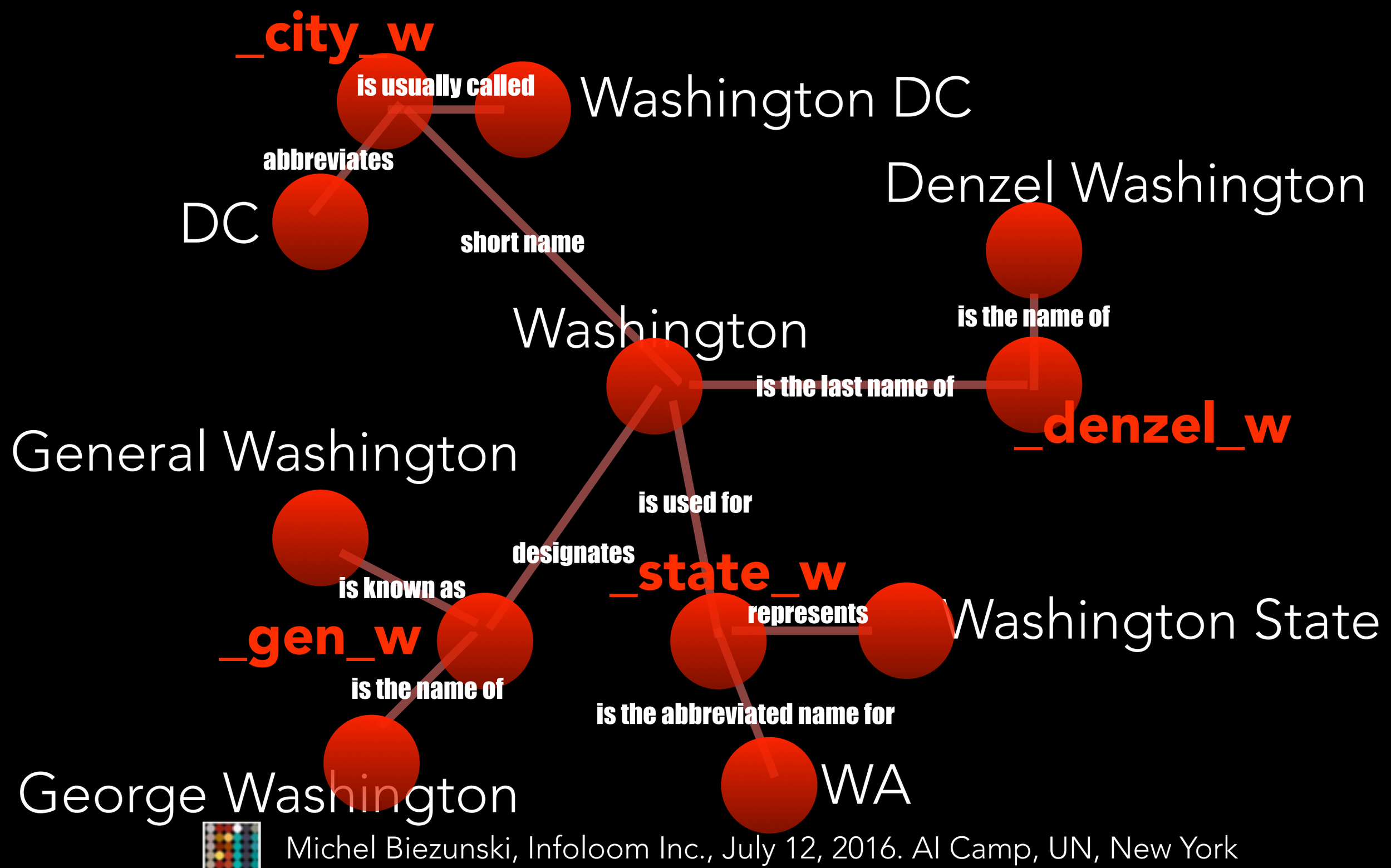
- Why should strings be subjects?
  - Because we may want to express some of their properties:

< Washington | length | 10 >

< Washington | encoding | UTF-8 >

< Washington | language | English >

MULTIPLE PERSPECTIVES: DIVERSITY



PERSPECTORS: DIVERSE OPERATORS

<	Washington	|	short name	|	_city_w	>
<	Washington	|	designates	|	_gen_w	>
<	Washington	|	is used for	|	_state_w	>
<	Washington	|	is the last name of	|	_denzel_w	>
<	Washington DC	|	is usually called	|	_city_w	>
<	DC	|	abbreviates	|	_city_w	>
<	Denzel Washington	|	is the name of	|	_denzel_w	>
<	Washington State	|	represents	|	_state_w	>
<	WA	|	is the abbreviated name for	|	_state_w	>
<	General Washington	|	is known as	|	_gen_w	>
<	George Washington	|	is the name of	|	_gen_w	>

PERSPECTORS: DIVERSE OPERATORS

<	Washington	|	short name	|	_city_w	>
<	Washington	|	designates	|	_gen_w	>
<	Washington	|	is used for	|	_state_w	>
<	Washington	|	is the last name of	|	_denzel_w	>
<	Washington DC	|	is usually called	|	_city_w	>
<	DC	|	abbreviates	|	_city_w	>
<	Denzel Washington	|	is the name of	|	_denzel_w	>
<	Washington State	|	represents	|	_state_w	>
<	WA	|	is the abbreviated name for	|	_state_w	>
<	General Washington	|	is known as	|	_gen_w	>
<	George Washington	|	is the name of	|	_gen_w	>

UNIFIED PERSPECTIVE

<	Washington	|	is a name for	|	_city_w	>
<	Washington	|	is a name for	|	_gen_w	>
<	Washington	|	is a name for	|	_state_w	>
<	Washington	|	is a name for	|	_denzel_w	>
<	Washington DC	|	is a name for	|	_city_w	>
<	DC	|	is a name for	|	_city_w	>
<	Denzel Washington	|	is a name for	|	_denzel_w	>
<	Washington State	|	is a name for	|	_state_w	>
<	WA	|	is a name for	|	_state_w	>
<	General Washington	|	is a name for	|	_gen_w	>
<	George Washington	|	is a name for	|	_gen_w	>

PERSPECTIVE AS MAPPING

short name	is a name of
designates	is a name of
is used for	is a name of
is the last name of	is a name of
is usually called	is a name of
abbreviates	is a name of
is the name of	is a name of
represents	is a name of
is the abbreviated name for	is a name of
is known as	is a name of
is the name of	is a name of

WRAP-UP

- Once flattened, information treats semantic units and processes the same way.
- Design choices to assimilate or distinguish are documented as projection methods.
- Multiple projection methods are possible, and therefore multiple views (aka perspectives).
- Information is a gigantic graph in which each node is an account (similar to accounting).
- It’s possible to traverse the graph to gather forensic data along the way about provenance and other relevant item of interest.

Michel Biezunski, Infoloom Inc., July 12, 2016. AI Camp, UN, New York