

Michel Biezunski

**moving to a**

**new home**

**once it gets**

**???**

**messy**

<https://.fatherly.com/wp-content/uploads/2018/02/messy-house.jpg>



Infloom, Inc.



# A COMMON PATTERN

- Brand New Information System.  
Well designed, well structured, well organized.
- After a while, information doesn't fit any more
- The current system would be too complex to upgrade.  
Rather, companies start from scratch.



# A WAY OUT OF THIS VICIOUS CYCLE

- Graph Data Structures
- Combining Manual Curation with Automatic Processes.



# A CHANGING ENVIRONMENT

Big Data	Does Information still have value?	Yes, but it's the most valuable when it is relevant and accurate. At that point, it is small data.
Artificial Intelligence	Are humans still necessary?	Yes, especially in novelty, tricky situations and when decisions have to be made that need to be accounted for.
Cloud	Should data structures reflect what browsers need?	Yes, but that's mainly for visualization. There is still data in the background.

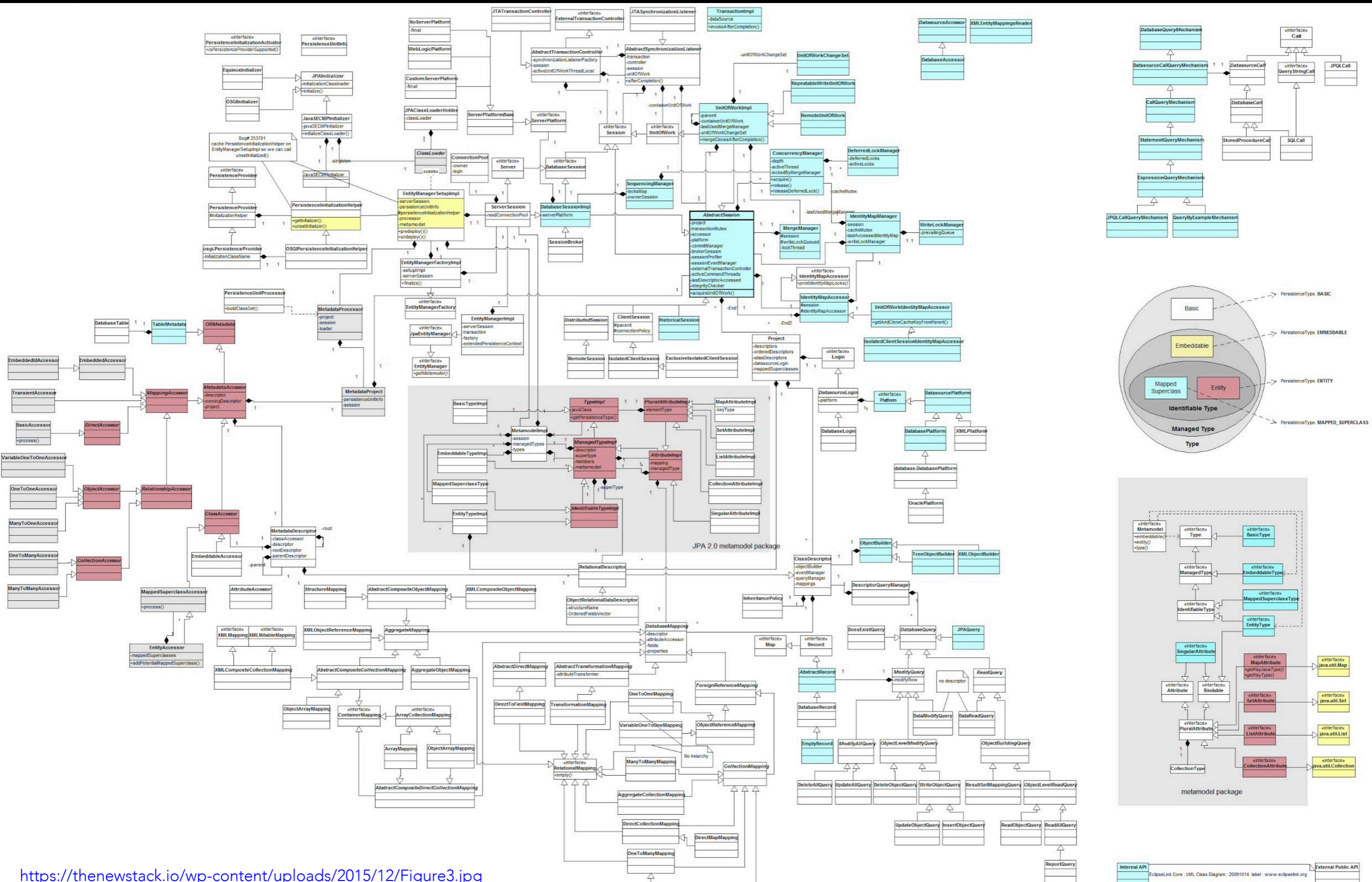


# UNRESOLVED CHALLENGES

- High value data.
  - Where is it? How do we find it?
- Accountability / Security.
  - Who is responsible?
- Presentation vs. Content.
  - What about the internal structure? Long-term preservation?

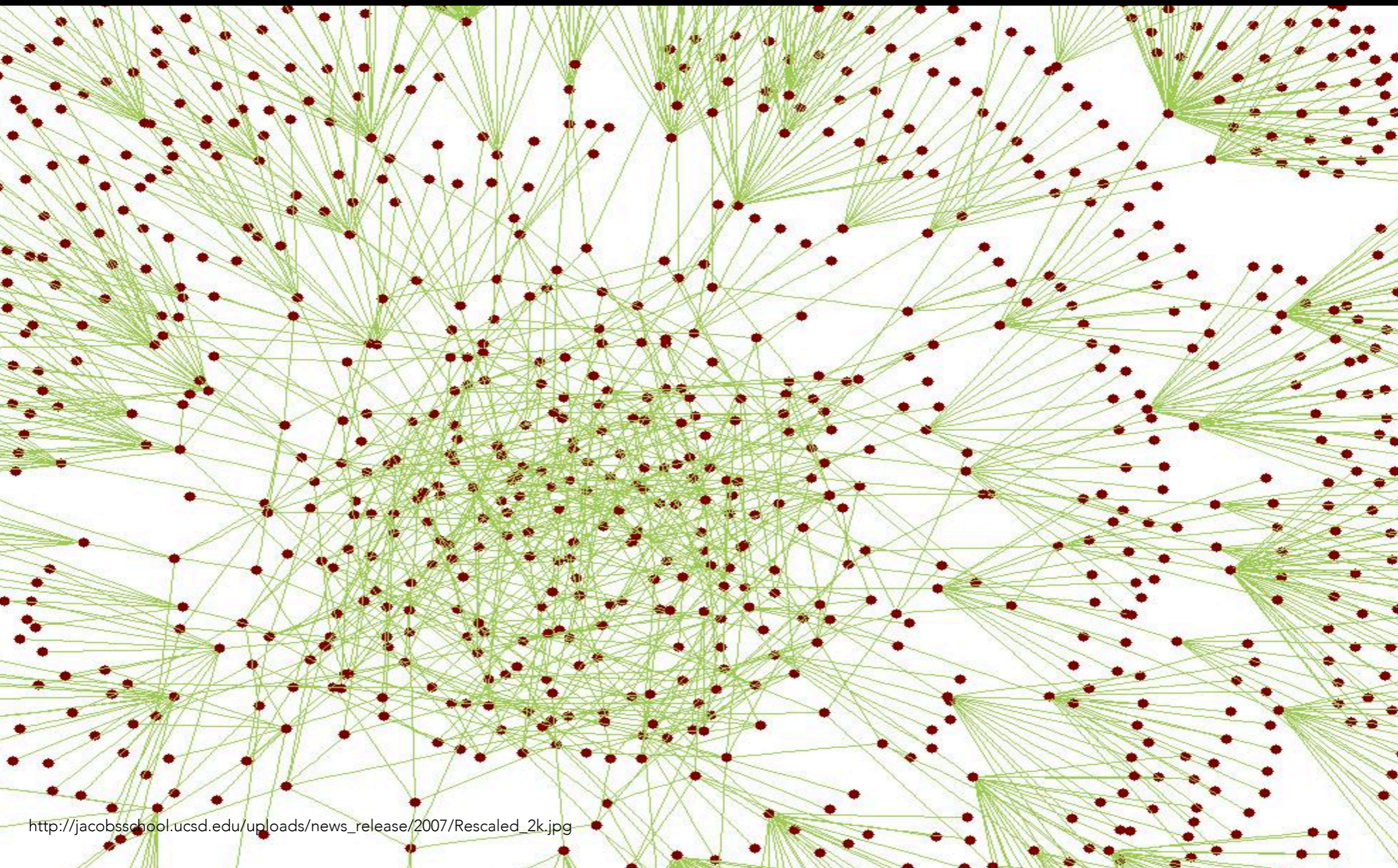


# DATABASE/XML SCHEMAS





# GRAPHS





# LEARNING CURVE

**ASSIGNMENT**

**tuition fee**

**exam**

**€** **\$** **¥**

$x = x_0 + v_0 t + \frac{1}{2} a t^2$

$v_f = v_0 + a t$

$\cos x - \cos y = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$

$\csc(-x) = -\csc(x)$

$\cos(-x) = \cos(x)$

$\sec(-x) = \sec(x)$

$\tan(-x) = -\tan(x)$

$\sin(-x) = -\sin(x)$

$\csc(-x) = -\csc(x)$

$\cos(-x) = \cos(x)$

$\sec(-x) = \sec(x)$

$\tan(-x) = -\tan(x)$

$\cot(-x) = -\cot(x)$

$E=MC^2$

A	v	B
v	0	v
0	0	v
0	0	0

$x = x_0 + v_0 t + \frac{1}{2} a t^2$

$v_f = v_0 + a t$

$a = v^2 / R$

$F = ma = mv^2 / R$

$\tan^2(x) + 1 = \sec^2(x)$

$1/a - 2 - x - 2 / e - 2 = 1$

$\sin^2(x) + \cos^2(x) = 1$

$R_{eq} = R_1 + R_2 + R_3 + \dots$

$\sin x - \sin y = 2 \cos\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$

$\cos x - \cos y = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$

$Mg(NO_3)_2$

$\int_a^b f(x) dx = S$

**ASSIGNMENT**

**tuition fee**

**exam**

**€** **\$** **¥**

$x = x_0 + v_0 t + \frac{1}{2} a t^2$

$v_f = v_0 + a t$

$\cos x - \cos y = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$

$\csc(-x) = -\csc(x)$

$\cos(-x) = \cos(x)$

$\sec(-x) = \sec(x)$

$\tan(-x) = -\tan(x)$

$\sin(-x) = -\sin(x)$

$\csc(-x) = -\csc(x)$

$\cos(-x) = \cos(x)$

$\sec(-x) = \sec(x)$

$\tan(-x) = -\tan(x)$

$\cot(-x) = -\cot(x)$

$E=MC^2$

A	v	B
v	0	v
0	0	v
0	0	0

$x = x_0 + v_0 t + \frac{1}{2} a t^2$

$v_f = v_0 + a t$

$a = v^2 / R$

$F = ma = mv^2 / R$

$\tan^2(x) + 1 = \sec^2(x)$

$1/a - 2 - x - 2 / e - 2 = 1$

$\sin^2(x) + \cos^2(x) = 1$

$R_{eq} = R_1 + R_2 + R_3 + \dots$

$\sin x - \sin y = 2 \cos\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$

$\cos x - \cos y = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right)$

$Mg(NO_3)_2$

$\int_a^b f(x) dx = S$





# DATABASE

	A	B	C	D
1	Title	Author	Publisher	
2	Wishful Thinking	Jim Dulani	GiGi	
3	Just Mercy	Bryan Stevenson	Spiegel & Grau	
4	Where the Crawdads Sing	Delia Owens	G.P. Putnam's Sons	
5	All the Light We Cannot See	Anthony Doerr	Scribner	
6	Send Down the Rain	Charles Martin	HarperCollins	
7	Then She Was Gone	Lisa Jewell	Atria Books	
8	Educated	Tara Westover	Random House	
9	We Were the Lucky Ones	Georgia Hunter	Penguin Group	
10	Ordinary Grace	William Kent Krueger	Simon and Schuster	
11	All the Ugly and Wonderful	Bryn Greenwood	A Thomas Dunne Book for St. Martin	
12	The Shell Seekers	Rosamunde Pilcher	Macmillan	
13	Lilac Girls	Martha Hall Kelly	Ballantine Books	



# ACTUALLY, THE STORY IS...

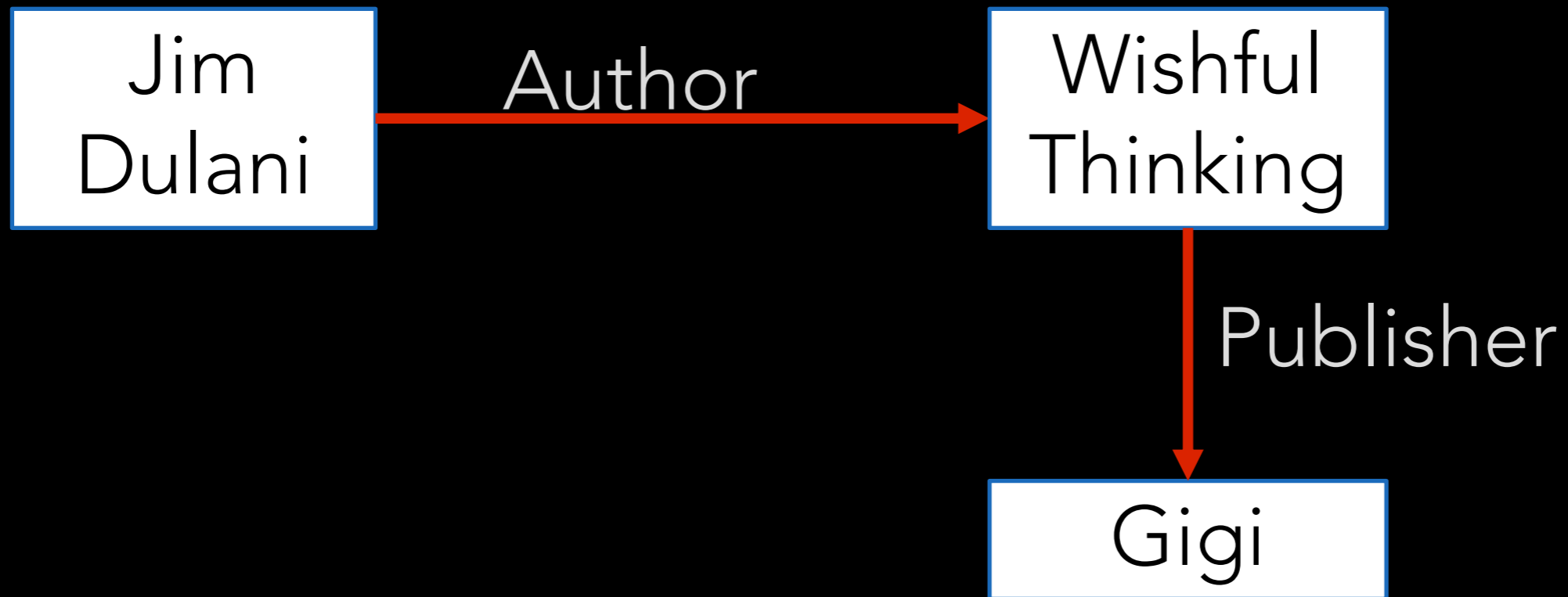
Title	Author	Publisher
Wishful Thinking	Jim Dulani	GiGi

- The publisher has used a simplified name for the author. His real name is James T. Dulãñyi, but they couldn't find how to type these characters when they published the book.
- The publisher is just a web site for self-publishing. There is no copyright, nor ISBN number.
- A movie production company is looking for purchasing the rights to adapt the book but they can't find it.
- They are trying to get in contact with the author, but don't find him either.



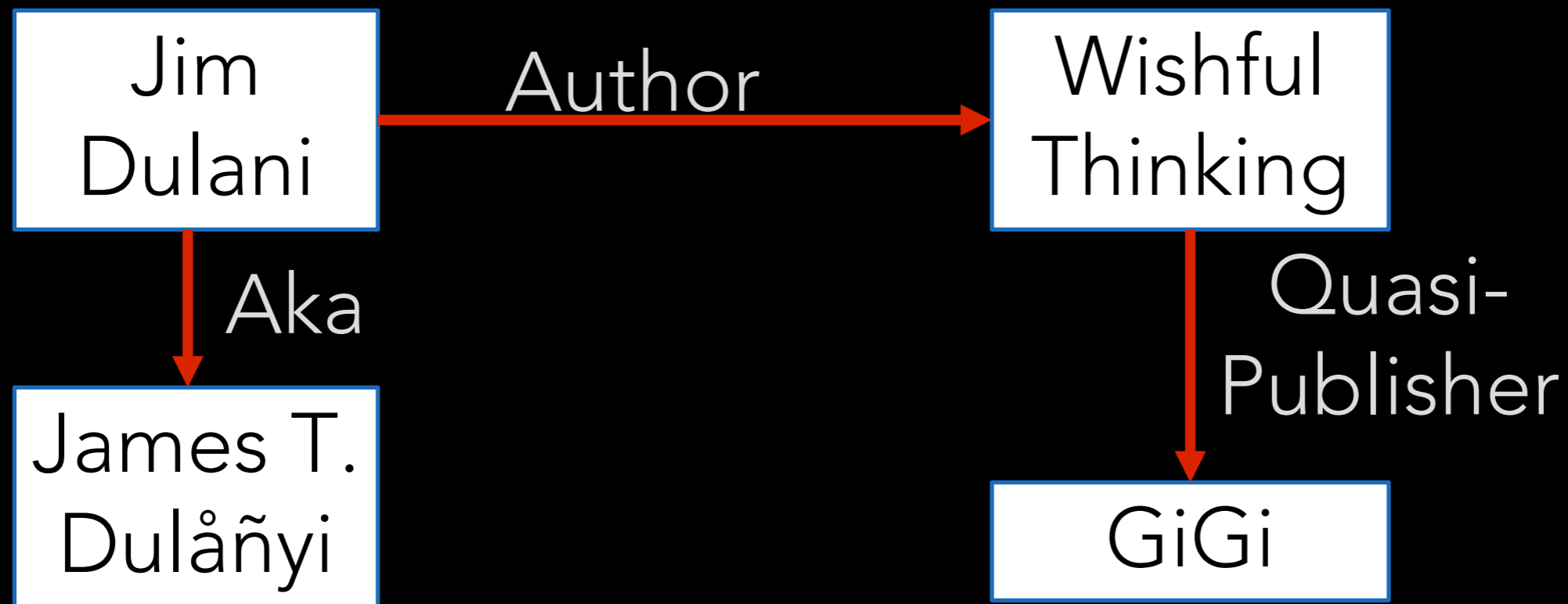
# GRAPH

Title	Author	Publisher
Wishful Thinking	Jim Dulani	GiGi



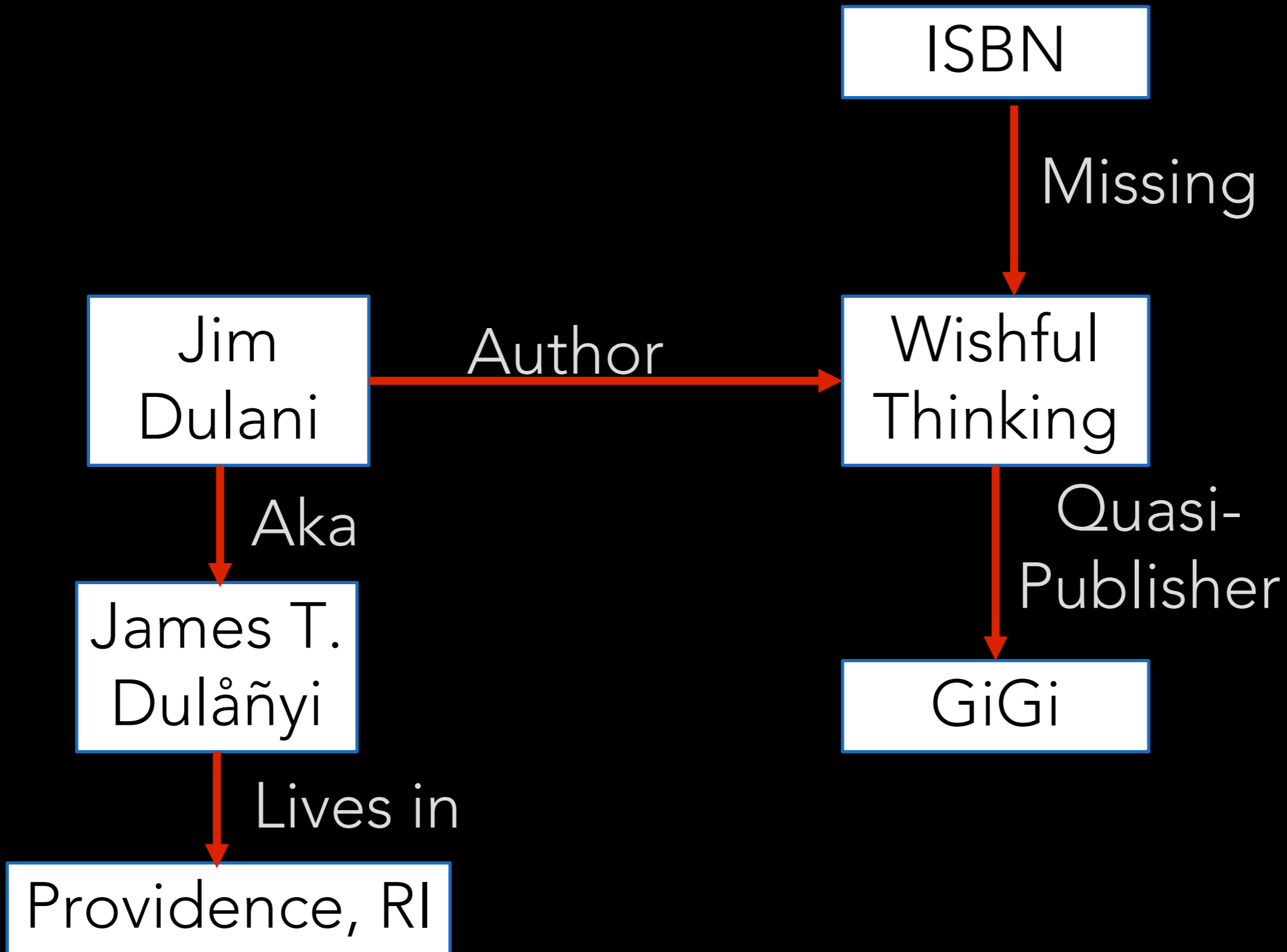


# GRAPH MODIFIED



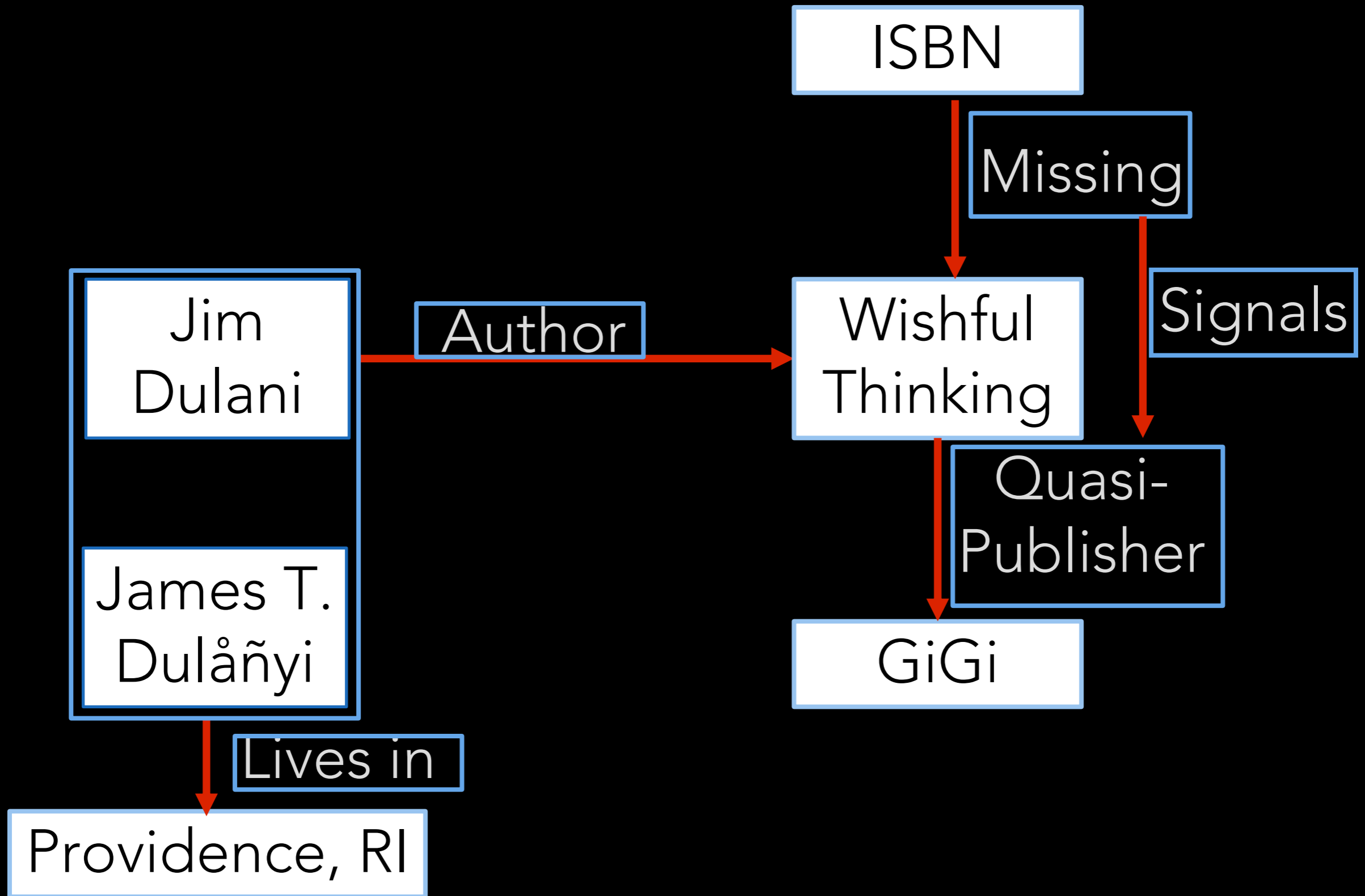


# GRAPH MODIFIED, MORE





# TOPIC MAP VIEW





# WHY SWITCH TO GRAPHS?

- Make it hospitable for multiple data repositories to co-exist
- Make it possible to invent new inference rules
- Make it possible to enable human curation.
- Maintainable over the long term, because a graph can have areas that are messy without breaking the rest.



# GRAPH VS. GRAPHIC

- Visually speaking, a graph structure can be purely textual.
- The Web is a graph.
- A hierarchical tree is a graph.
- A structured document is a graph.
- Relational databases are graphs.
- Bottom line: graphs can ingest legacy information, and do more.





# GRAPH TECHNOLOGIES

- Graph Databases
- Relational databases configured to handle graphs
- Graph-based standards:



# HUMANS STILL NEEDED!

- Good information needs to be curated.
- Information needs to be auditable.
- Companies need to be held accountable for what they do and say.
- For computers, mess is a bug.
- For humans, mess is a feature.



# SO, MOVING TO A NEW HOME WHEN IT GETS MESSY?

- No, of course not. Instead:
  - Don't undervalue the mess. Enable it instead.
  - Mess doesn't go away just because things are swept under the rug.
  - Consider cleaning when things go out of control.
  - Use cleaning tools. Automate what you can, do the rest yourself.
  - You don't have to clean everything, certainly not in one shot.
  - Once you're done cleaning, consider you may have to clean again later.



• Questions ?

Contact: [mb@infloom.com](mailto:mb@infloom.com)