



Center for the Advancement of
Data and Research in Economics

CLINT:

Introducing the **C**ategorically **L**INKed Timeline database

October 7, 2016

San Cannon

Center for the Advancement of Data and Research in Economics
Federal Reserve Bank of Kansas City

The views expressed in this presentation are those of the author and don't necessarily reflect the position of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

OUTLINE

1. A story
2. The content
3. The challenges
4. The current state
5. The plan

A STORY



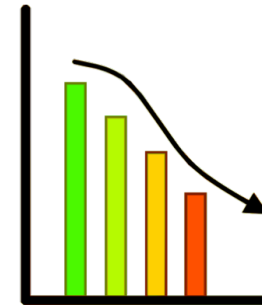
ONCE UPON A TIME...

There was an research assistant at the Federal Reserve Board who worked on “the Monthly Model”



She was in charge of running the forecast model on her own for the first time.

And her forecast for near term GDP was tanking.

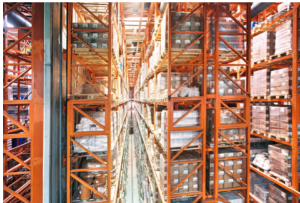


And she couldn't figure out why.

WHAT TO DO?

She worked diligently to find the root of the problem:

Manufacturing output was projected to slide into the abyss because....



Durable goods production was forecast to plummet because...

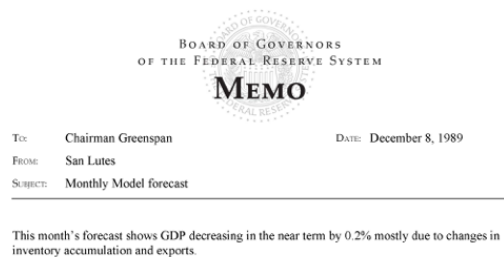
Aircraft production was projected to go to zero because....



Boeing had been on strike during the baseline period for her forecast.

A HAPPY ENDING

It took our intrepid heroine more than three days to pick through the connections to find the problem.

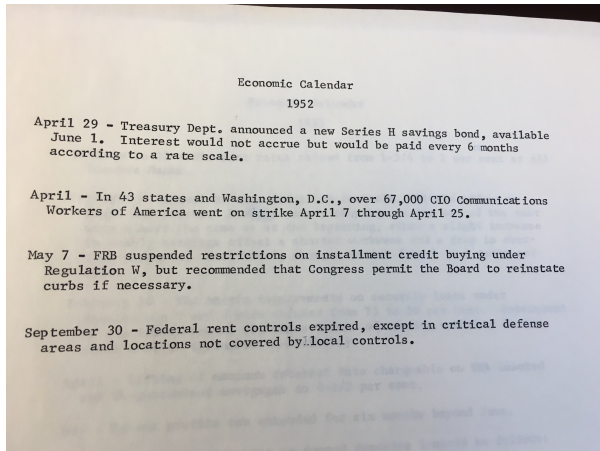


One small “adjustment factor” later and the forecast was done and the memo to Chairman Greenspan was written.

MEANWHILE ...

In an office in an another part of the building...

There was a green notebook. It was an unassuming notebook but one that contained lots of information on “events” that happened from the 1940’s through the 1970’s



The kind of notebook where a Boeing strike would have been recorded.

And her life would have been so much easier. But she found it too late.

YEARS PASS....

And the former RA finds herself in a position to be able to do something with the Green notebook and possibly save future RA's some grief.

Two summers, two interns

Lots of existential angst

Lots of classification challenges

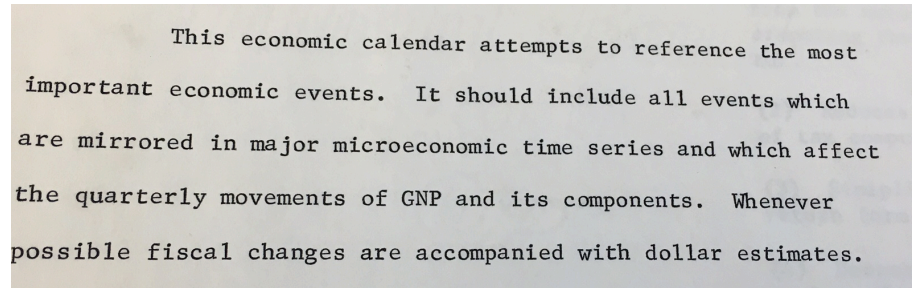
And the CLINT database was born.

THE CONTENT



WHAT'S THE POINT?

The intention for the Green notebook was clear:



This economic calendar attempts to reference the most important economic events. It should include all events which are mirrored in major microeconomic time series and which affect the quarterly movements of GNP and its components. Whenever possible fiscal changes are accompanied with dollar estimates.

And we wanted to continue in that vein:

Not only do we want to capture the event but also the potential effect of that event.

FIRST: CAPTURE THE EVENT

The Green notebook was a great starting point.

We added similar events from the post 1970s period.

We added different events that we thought met the Green notebook criteria.

We consulted other institutions and websites that collected and presented “events”. ¹

¹And are working to secure permission to include them in the database. My copyright expert would want me to say that.

NEXT: IDENTIFY METADATA

It wouldn't be fun without metadata and we had to identify elements we wanted to use to describe our event.

We started with trying to decide a few basic elements: event name, event type, date, location.

But to connect these events to “quarterly movements of GNP and its components”, we need to link the event to some part of the economy.

So we created industry and sector metadata fields.

FOLLOWED BY: DEFINING METADATA ELEMENTS

Once we identified the relevant metadata elements (or thought we did), we needed to define them and the permissible values for those fields.

- ▶ **Name** and **description** are allowed to be free text. Allows for copious copying and pasting.
- ▶ **Dates** are W3C three part dates. Very precise.
- ▶ **Location** can be pulled from standard code lists (e.g. ISO country codes, MSA identifiers, FIPS codes, etc.)
- ▶ **Industry** can also use standard designations like NAICS (or SIC for longevity purposes).

And some items we had to make up. Is there a standard code list for “economic sector” or “event type”?

CREATING CONTROLLED VOCABULARIES: SECTORS

Let's start with the concept of "sector" - if we have industries, why do we need sector?

Because some concepts fit more readily with sector. "Business sector" and "labor markets" can be affected by events regardless of industry.

SECTOR LIST

We currently define 8 sectors: ¹

- ▶ Business
- ▶ Consumers
- ▶ Finance
- ▶ Government
- ▶ Housing
- ▶ Labor Markets
- ▶ Macroeconomic
- ▶ Natural Resources

¹Note: Everything is a work in progress. Selected items will be rationally defended and suggestions are cheerfully accepted after the presentation.

CREATING CONTROLLED VOCABULARIES: EVENT TYPE

Green notebook lists mainly fiscal or monetary policy items, legislative changes, or major industrial actions.

But what about Hurricane Katrina? The LA riots? Deepwater Horizon? Failure of Lehman Brothers?

Arguably, all have had economic effects. None are policy or legislation.

EVENT TYPE LIST

Working through our base compendium of events, we define 16 types into which they can fall.¹

- ▶ Legislative
- ▶ Monetary Policy
- ▶ Fiscal Policy
- ▶ Regulation
- ▶ Court ruling
- ▶ Executive order
- ▶ Conflict
- ▶ Statement
- ▶ Change in leadership
- ▶ Firm outcome
- ▶ Agreement
- ▶ Business conflict
- ▶ Casus Fortuitous
- ▶ Eventum

¹See footnote on previous slide. Applies here too.

CURRENT METADATA ITEMS

Besides the text and date fields, we have the following categorizations:

- ▶ Event type
- ▶ Industry
- ▶ Sector
- ▶ Agent
- ▶ Foreign/domestic
- ▶ Country
- ▶ Region

Some have multiple entries to allow for multiple classifications.

THE CHALLENGES



WHAT IS AN EVENT?

This should be easy, right?

Events happen.

Except when they don't and that is what has the effect.

And some “events” are words but those have effects.

Then there are the events that are nested inside other events.

SOME EXAMPLES

Which of these are events?

- ▶ Flash crash
- ▶ “Irrational exuberance”
- ▶ Unemployment report has unexpected number.
- ▶ Official statistical release changes methodology.
- ▶ The Dust Bowl
- ▶ World War II
- ▶ Invasion of Normandy

WHEN DID IT HAPPEN?

Even more fun than defining events? Assigning a date to them.

Some events are brief and can be captured with a single date.

Some events last more than a day so they need a start and an end date.

Some events have multiple dates for starting or ending.

SOME EXAMPLES

When did these “happen”?

- ▶ “The financial crisis”
- ▶ The Patient Protection and Affordable Care Act
- ▶ The Dust Bowl
- ▶ Brexit
- ▶ World War II

OTHER CLASSIFICATION ISSUES

There are lots of “sticky” questions:

- ▶ What kind of event is a municipal bankruptcy?
- ▶ Is Puerto Rico a country? What about Taiwan?
- ▶ Who is part of “Europe”?
- ▶ Should we try to define technology “eras”?

AND THE HARDEST OF ALL

Classifying the economic effect.

We're pretty sure that strikes affect labor markets. Depending on the strike, we might be able to identify industry and/or geography.

We might have an idea about monetary policy effects.

But how do we define the scope? Hurricane Katrina was a natural disaster (Casus Fortuitous in Clint lingo) that directly affected the Gulf Coast region. Which industries? Which sectors? For what dates?

And we need to consider which economy is affected.

THE CURRENT STATE



SOME STATISTICS

We have about 3000 events identified.

Event Type	Code	Number of Instances in Database
Legislative	L	580
Monetary Policy	MP	473
Regulation	R	179
Casus Fortuitous	CF	102
Conflict	C	142
Statement	ST	83
Marker	MR	634
Court Ruling	CR	21
Change in leadership	CL	72
Eventum	ET	39
Firm Outcome	FO	156
Report	RT	21
Fiscal Policy	FP	178
Agreement	A	69
Executive Order	EO	12
Business Conflict	BC	167
Total		2928

MORE STATISTICS

Industry	Code	Number
ALL	ALL	339
Agriculture	AG	56
Mining	M	8
Construction	CS	18
Manufacturing	MA	83
Wholesale and retail trade	WR	51
Transportation and Utilities	TU	76
Information	IT	28
Financial Activities	FA	1104
Professional and Business Services	BS	3
Educational and Health Services	EH	37
Leisure and Hospitality	LH	4
Other Services	OS	7
Public Administration	PA	514
Armed Forces	AF	113
Technology	TC	21
Energy	E	160
None	N	351
Total		2973

Sector	Code	Number
ALL	ALL	201
Business	B	207
Finance	F	985
Labor Markets	LM	158
Consumers	CN	18
Government	G	753
Housing	H	93
Macroeconomics	ME	167
Natural Resources	E	0
None	N	227
Total		2809

DATA POINT 1

Event: 616

- Description: Northridge Earthquake. A destructive earthquake registering 6.8 on the Richter scale, centered in Northridge, struck Southern California at 4:31 a.m.
- url: http://www.dof.ca.gov/Forecasting/Economics/Chrono_Sig_Events/
- Event Type: Casus Fortuitous
- Country: United States
- State: California
- City: Northridge, Los Angeles
- Date Start: January 17,1994
- Date End: January 17,1994

DATA POINT 2

Event: 999

- Description: US Abandonds gold standard. Nixon ends the convertability of USD to gold through Executive Order 11615. This brings an end to the Bretton Woods Era of exchange-rate fixity.
- url: <http://www.federalreservehistory.org/Events/DetailView/33>,
<http://www.presidency.ucsb.edu/ws/?pid=3115>
- Event Type: Monetary Policy, Executive Order
- Sector: Business, Finance, Government, Macroeconomics
- City: Washington DC
- Country: USA
- Agent: Richard Nixon Administration
- Date Start: August 15, 1971
- Date End: August 15, 1971

DATA POINT 3

The one that started it all:

Event: 4562

- Description: Boeing machinists strike. Boeing machinists go on strike, halting production for the maker of 60% of the world commercial jets. This event halted production on \$80 billion of planes.
- url: <http://www.nytimes.com/1989/10/05/us/boeing-strike-halts-work-on-80-billion-in-planes.html>
- Event Type: Business Conflict
- Sector: labor market, business
- Industry: manufacturing
- Agents: Boeing machinist union (IAM 751), Boeing
- Country: USA
- Date start: October 4, 1989
- Date End: Nov 21, 1989

THE PLAN



COMPILING CONTENT

- ▶ Add more events.
- ▶ Clean up and classify events.
- ▶ Document sources and secure permissions.
- ▶ Include research on effects (e.g. paper citations)
- ▶ Create workflow to make population self-sustaining.

DESIGN DATABASE

- ▶ Reformat from flat storage to (possibly) graph database.
- ▶ Add linkages and relationships.
- ▶ Add perspective structure.
- ▶ Build usable search interface.
- ▶ Build data capture interface.

SHARE!

- ▶ Get feedback from interested parties (like you!)
- ▶ Launch beta version for public use.
- ▶ Solicit input from the public for relevant events and classifications.
- ▶ Create verification and formatting infrastructure.
- ▶ Develop classification algorithm.
- ▶ ~~Save the world.~~

IMAGE SOURCES

Slide Number	Citation
4	Isaiah658. (2015). Progress Chart Loss. Available at https://openclipart.org/detail/227737/progress-chart-loss Copyright released into the public domain
5	Unknown. (1913). 1913 Rochester Garment Workers Strike. Available at https://commons.wikimedia.org/wiki/File:1913_Rochester_Garment_Workers_Strike.jpg [Public domain]
5	Niklem. (2010, May 27). Hemofarm manufacturing warehouse, Vrsac.jpg. Available at https://commons.wikimedia.org/wiki/File:Hemofarm_manufacturing_warehouse,_Vrsac.jpg Reprinted courtesy of the Rights Holder under a Creative Commons license CC BY-SA 3.0
5	Manolo Gómez. (2007). 747. <i>Flickr.com</i> . Retrieved from https://flic.kr/p/3eBUAH Reprinted courtesy of the Rights Holder under a Creative Commons license CC BY 2.0
5	Dinesh Wijekoon. (2015, April 8). Factory. <i>Flickr.com</i> . Retrieved from https://flic.kr/p/8GsyeE Reprinted courtesy of the Rights Holder under a Creative Commons license CC BY 2.0

LAST SLIDE

Thanks for listening!
sandra.cannon@kc.frb.org
(816) 881-2596

