

MetaNet: Deep semantic automatic metaphor analysis

George Lakoff, Ellen Dodge, and Luca Gilardi

Center for the Neural Mind and Society
CNMS.berkeley.edu

Research IT Reading Group
November 17, 2016

Road map

- Background:
 - MetaNet project
 - What are metaphors? Why are they important?
- Overview of current MetaNet system
- Ongoing and future work: objectives and challenges

The Metanet Project

Started in 2012 as part of IARPA's Metaphor Project

Initial objectives:

- Build a multi-lingual metaphor repository
- Automatically extract metaphors from text
- Compare conceptual metaphors used by different groups and cultures in order to better understand their different beliefs and worldviews
- Support multilingual analysis

Why are metaphors important?

- Metaphors provide a way to reason about one domain of experience (the **target** domain) in terms of some other domain (the **source** domain).

For instance, **social problems** are often discussed in terms of **disease** :

- **Poverty** is the world's deadliest **disease**.
- This city is experiencing an **epidemic** of **gun violence**.

Why are metaphors important?

Metaphoric phrases are expressions of more general systems of conceptual metaphors.

- *Poverty and gun violence are plagues/epidemics/infections*
- *We need to cure/alleviate/treat/eradicate these problems.*

Metaphors affect the kinds of inferences we make about a given target domain.

- Infections and diseases harm people (the 'patients')
- Left unchecked, they are likely to spread and affect more people
- Society needs to act to prevent negative effects (it is not up to the patient to cure themselves)

Why are metaphors important?

The same concept can potentially be understood via different, alternative metaphor systems, each of which support different sets of inferences.

Another metaphoric conceptualization of poverty:

- Poverty is a location that people get 'stuck' in
*the **abyss** of poverty, **mired** in poverty*
- People are responsible for their own economic state
- We can address poverty by helping impoverished people leave this location:

*lift people from poverty, a **ladder out of** poverty*

Metaphor analysis can thus provide valuable insights into the different ways people may think about a particular domain

Why are metaphors important?

Metaphors are pervasive in language use, even though we are not necessarily conscious of their presence

- A MetaNet-based corpus study¹ indicates that over a third of the expressions involving poverty are highly metaphoric.
- Other studies indicate that metaphor explains 20% of all word meanings (Shutova and Teufel 2010; Steen et al. 2010)
- Google Translate: as many as 44% of metaphorical expressions translated incorrectly (Shutova, Teufel, and Korhonen 2013)

If we 'ignore' metaphor, we are losing a lot of valuable information about how people are conceptualizing and reasoning about a given domain

Challenges of analyzing metaphor

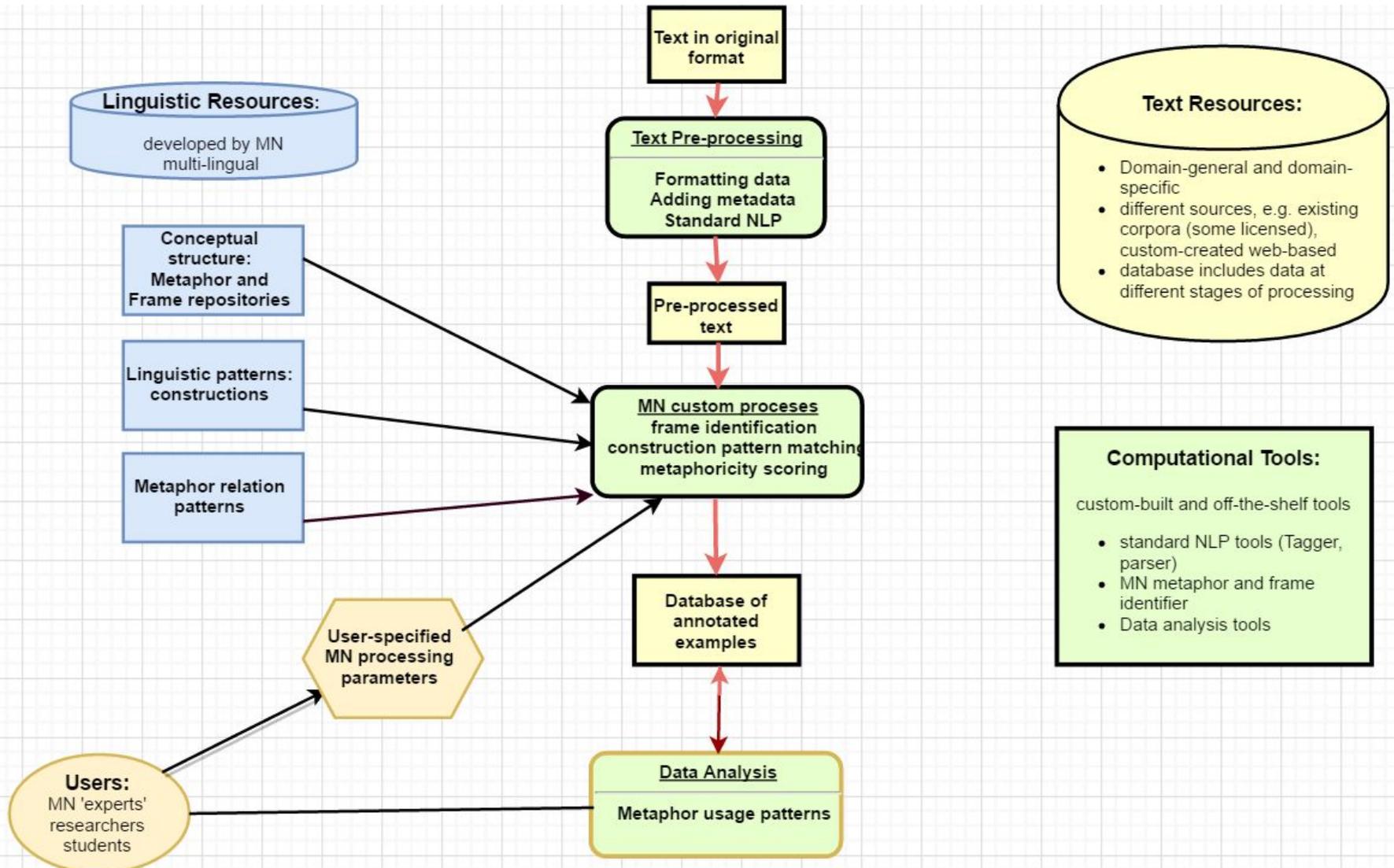
- Comprehensive metaphor analysis needs to do more than just identify whether specific words are being used metaphorically (or not)
- The methodologies employed in most types of current metaphor research cannot adequately perform metaphor analyses that are both **large-scale** and **in-depth**.

MetaNet system

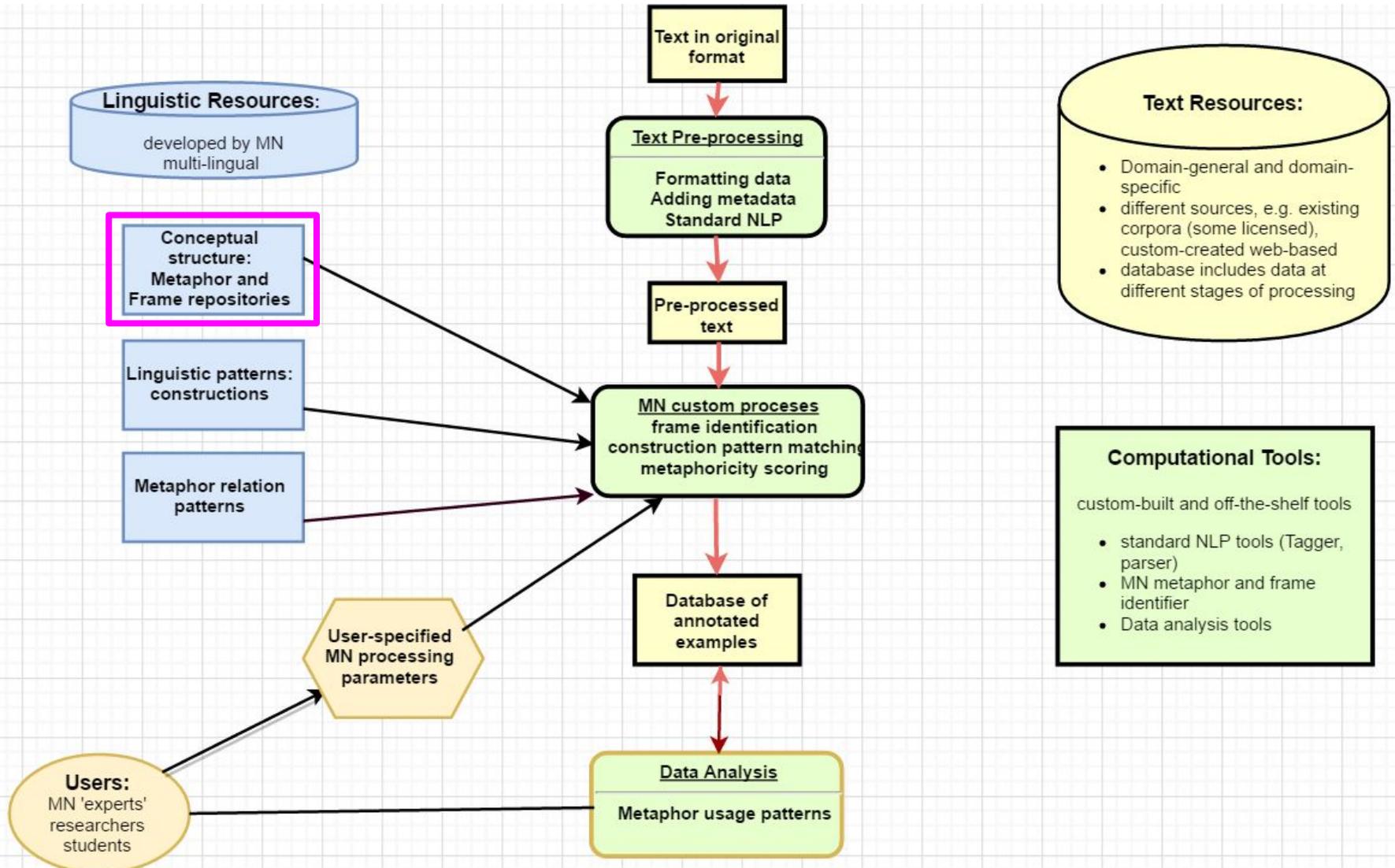
Road map:

1. Overview of current system:
 - A working system that automatically identifies metaphor in large-scale corpora, producing a database of annotated metaphor examples
 - Data visualization and analysis tools that have been used to analyze metaphor use in many domains, using different corpora
2. Future directions and challenges:
 - System improvements
 - Increased accessibility

System overview

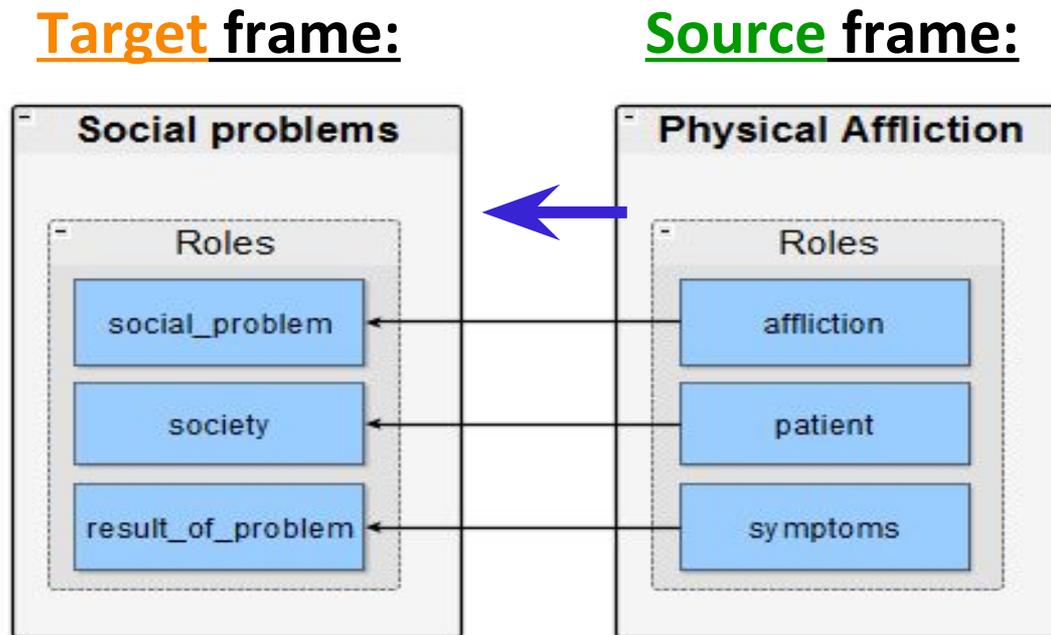


System overview



Metaphor and Frame Repository

- Formalization of conceptual metaphor theory (CMT) (Lakoff & Johnson, 1980)
- Conceptual metaphors represented as **mappings** from **Source** frames to **Target** frames

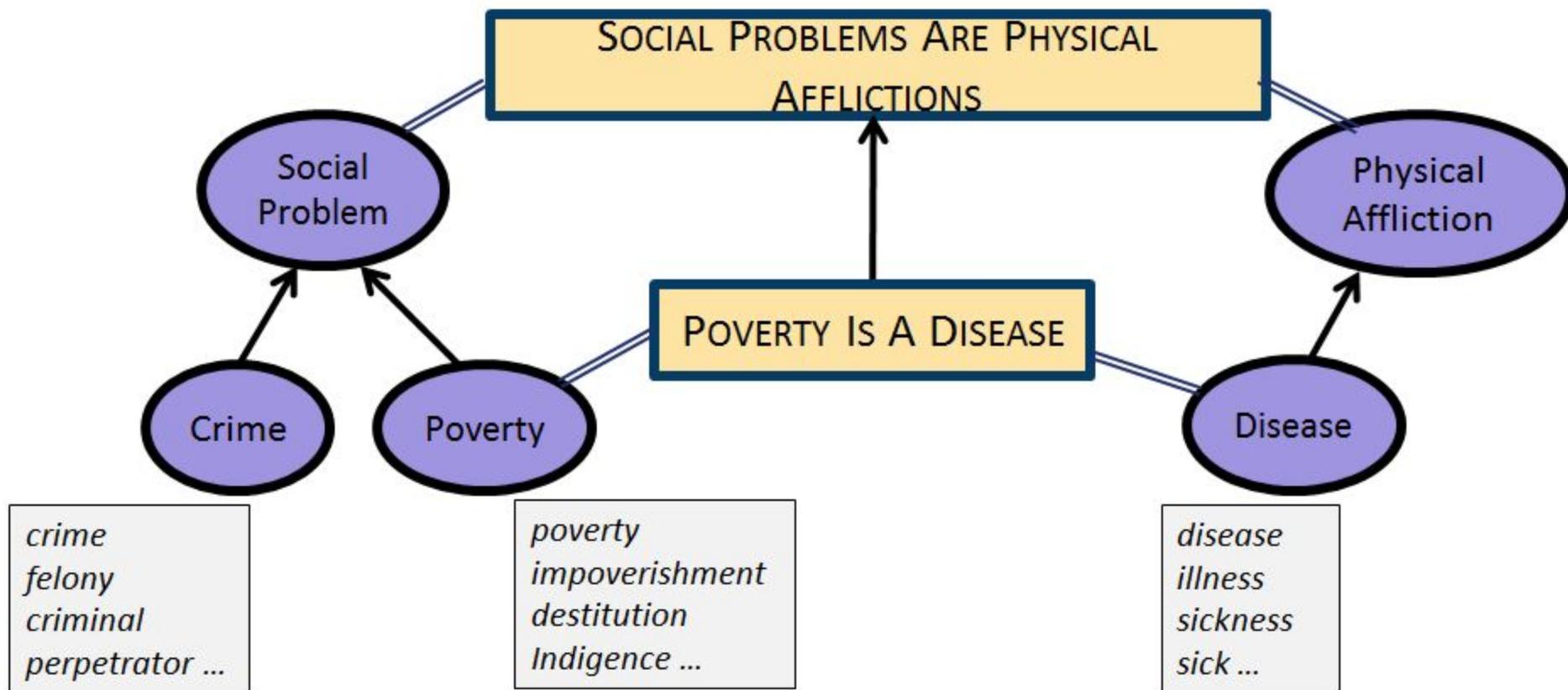


Metaphor: Social Problems Are Physical Afflictions

Metaphor and Frame Networks

Interconnected structures:

- Lexical items and expressions evoke frames
- Frames represent source and target domains of metaphors
- Individual frames and metaphors are parts of larger networks

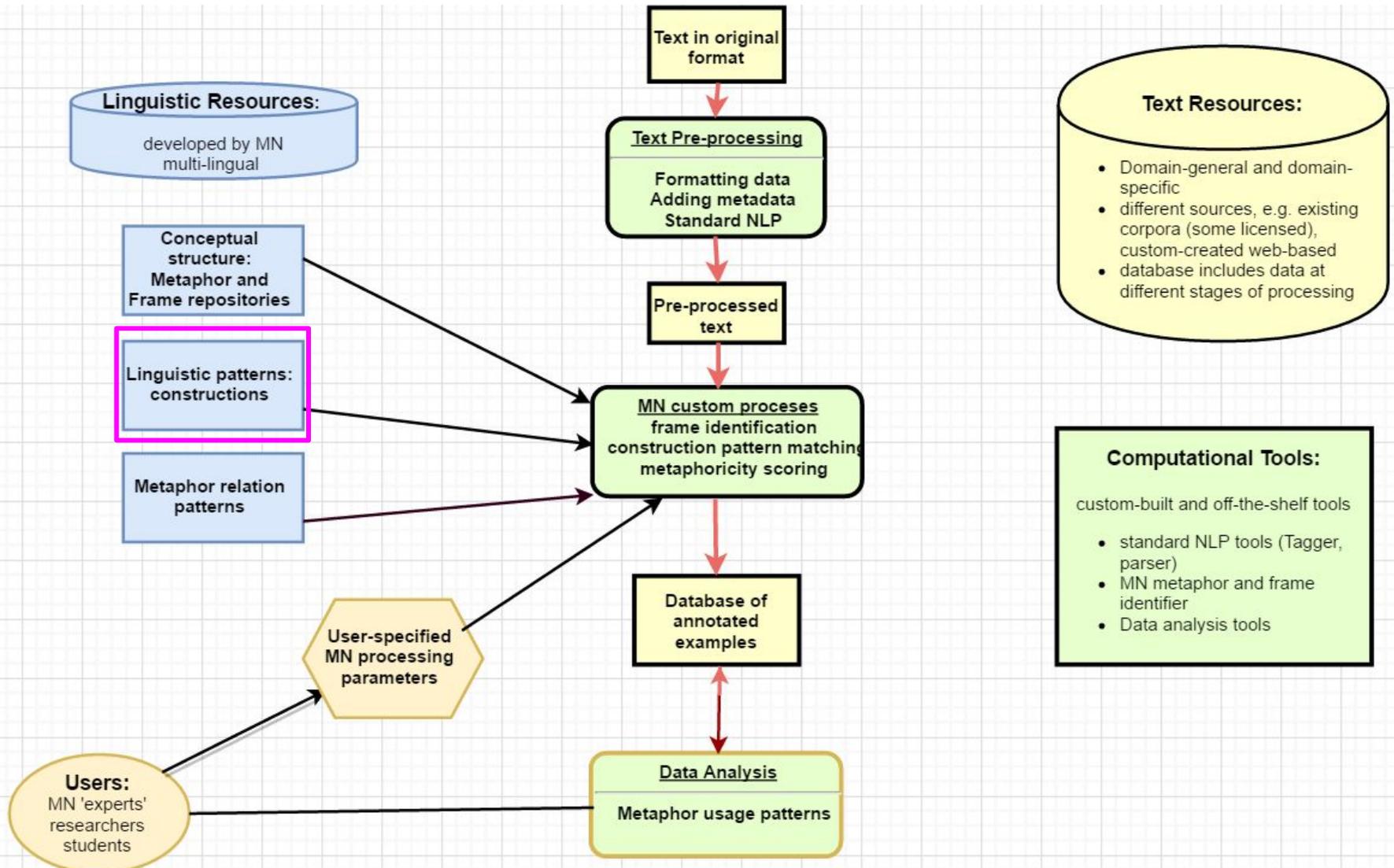


Metaphor and Frame Repository

- Implemented in Semantic Media Wiki (Krötzsch et al. 2007)
- Provides a collaborative tool for knowledge-based construction.
- Multilingual wikis: American English, Mexican Spanish, Russian, and Persian
- English developmental wiki currently contains:
 - Over 650 frames
 - 800 conceptual metaphors
- Publically-accessible wiki:

<https://metaphor.icsi.berkeley.edu/pub/en/>

System overview



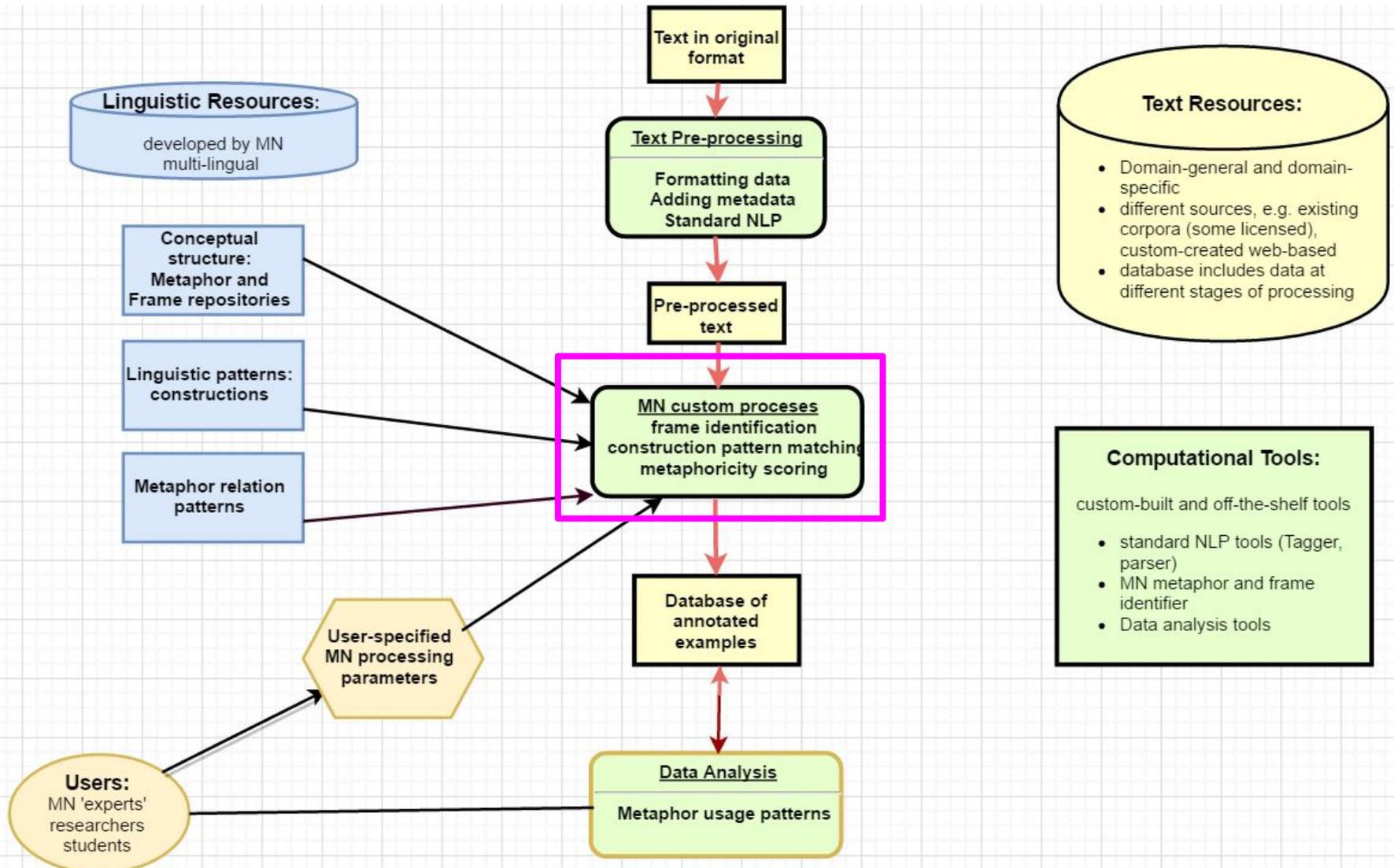
Metaphoric constructional patterns

- Conceptual metaphors are typically expressed in particular syntactic patterns
 - MetaNet constructional patterns are defined with respect to part of speech and dependency relations (of the kind produced by standard dependency parsers)
 - System identifies instances of source and target words occurring in these patterns
 - Words have links to frames, and thus provide access to the MN conceptual networks

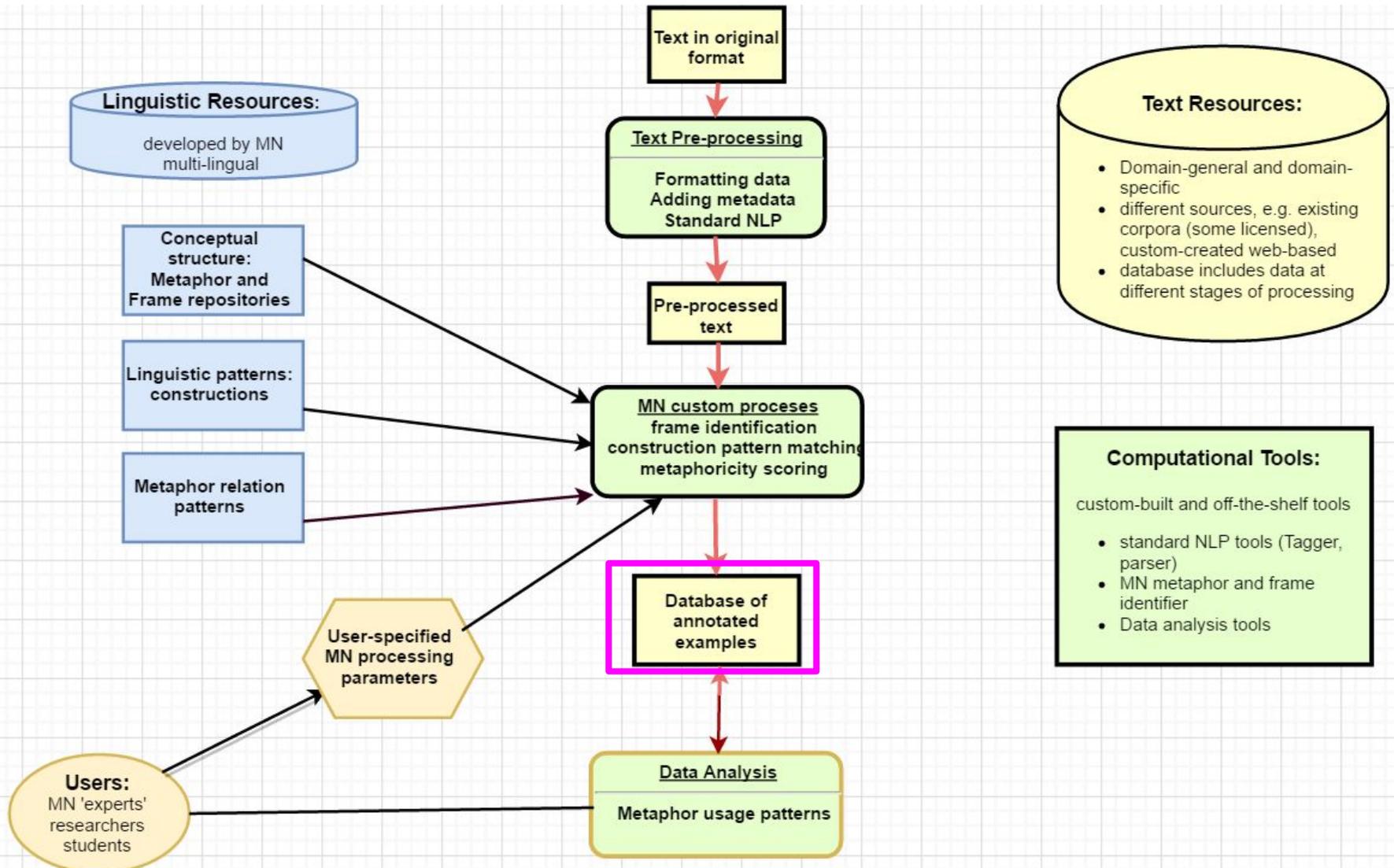
Metaphoric constructional patterns

Constructional pattern	Examples
T-subj_S-verb	<i>poverty infects</i>
T-subj_S-verb-conj	<i>poverty infects and maims</i>
T-subj-conj_S-verb	<i>homelessness and poverty infect</i>
S-verb_T-dobj	<i>escape poverty</i>
S-verb_T-dobj-conj	<i>escape despair and poverty</i>
S-verb_Prep_T-noun	<i>slide into poverty / pull up out of poverty</i>
S-noun_of_T-noun	<i>trap of poverty</i>
T-noun_poss_S-noun	<i>poverty's undertow</i>
S-noun_prep_T-noun	<i>path to poverty</i>
T-noun_mod_S-noun	<i>poverty trap</i>
S-adj_mod_T-noun	<i>burdensome poverty</i>
T-noun_cop_S-noun-adj	<i>poverty is a disease / poverty is burdensome</i>

System overview



System overview



Metaphor Viewer

We should look at fixing the mental health system in the United States in conjunction with implementing gun control, but we should not place the blame for the **gun death epidemic** in our country on those who were unfortunate enough to be born with mental problems.

In fact, the reduction of gun availability in our society would help alleviate the **epidemic** of **gun violence** that we are living in and would save many lives.

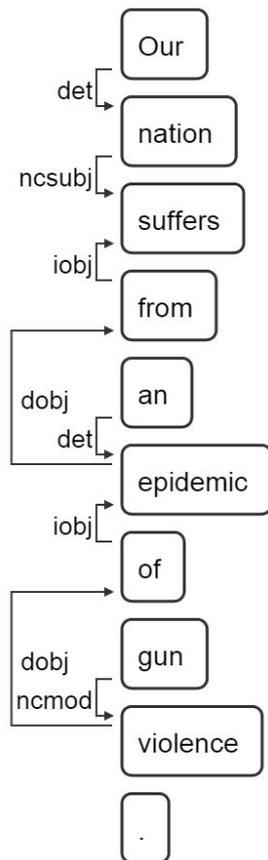
Or how about discussing the potential merits of a national no-sell list that would give law enforcement and mental health personnel the opportunity to flag potentially dangerous people to licensed gun-merchants? Would these kinds of modest regulations of gun ownership end the **scourge** of **gun violence** in America? Of course not people will always snap.

In the wake of Sandy Hook, US President Barack Obama announced the most aggressive federal gun-control plan in decades to combat what he termed an "**epidemic** of **gun violence**".

This lobby diverts attention from the fact that although revisiting gun control will not solve Americas **cancerous gun violence**, well-considered changes in legislation and enforcement will make a difference, and probably a big difference.

Our nation suffers from an epidemic of gun violence .

Dependency Tree



Linguistic Metaphors

gun violence epidemic

Source LU

Lemma: epidemic
form: epidemic
POS: epidemic.n
Schema(s): Disease spread
Schema families: Disease schema family
Concept(s): DISEASE
Concept score: DISEASE:1:0.13043
Map method: CNMS

Target LU

Lemma: gun violence
Form: gun violence
Schema: Injurious gun use
Schema family: (none)
Concept: GUN_VIOLENCE
Concept group: GUN OVERSIGHT

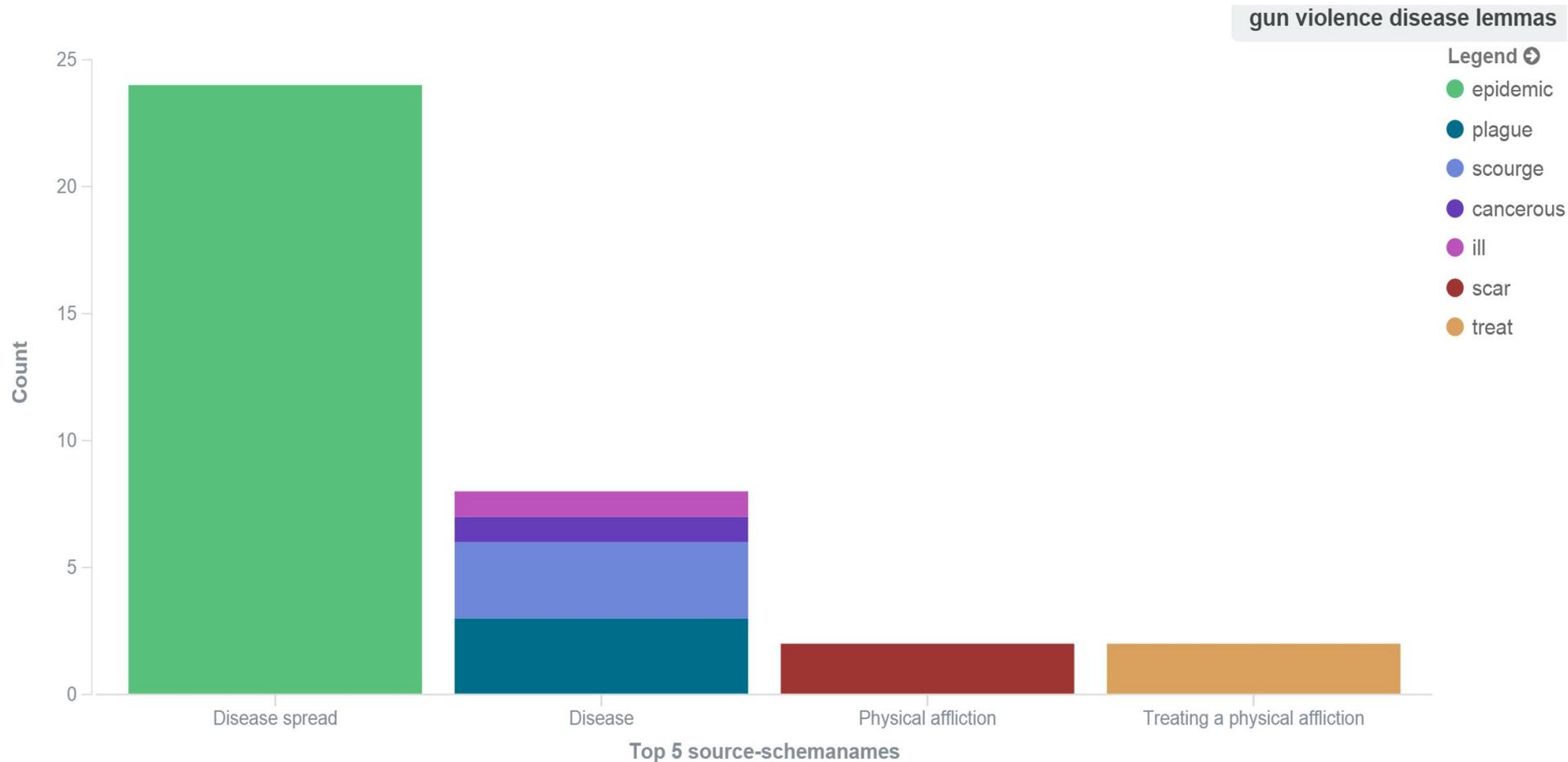
Extraction

Construction: S-noun_prep_T-noun
Score: 0.72
Scoring metric: Iposinwiki, vetsurvival
Extractor: CMS:CNMS
Seed: None

Distribution visualization

Target: **Gun violence**

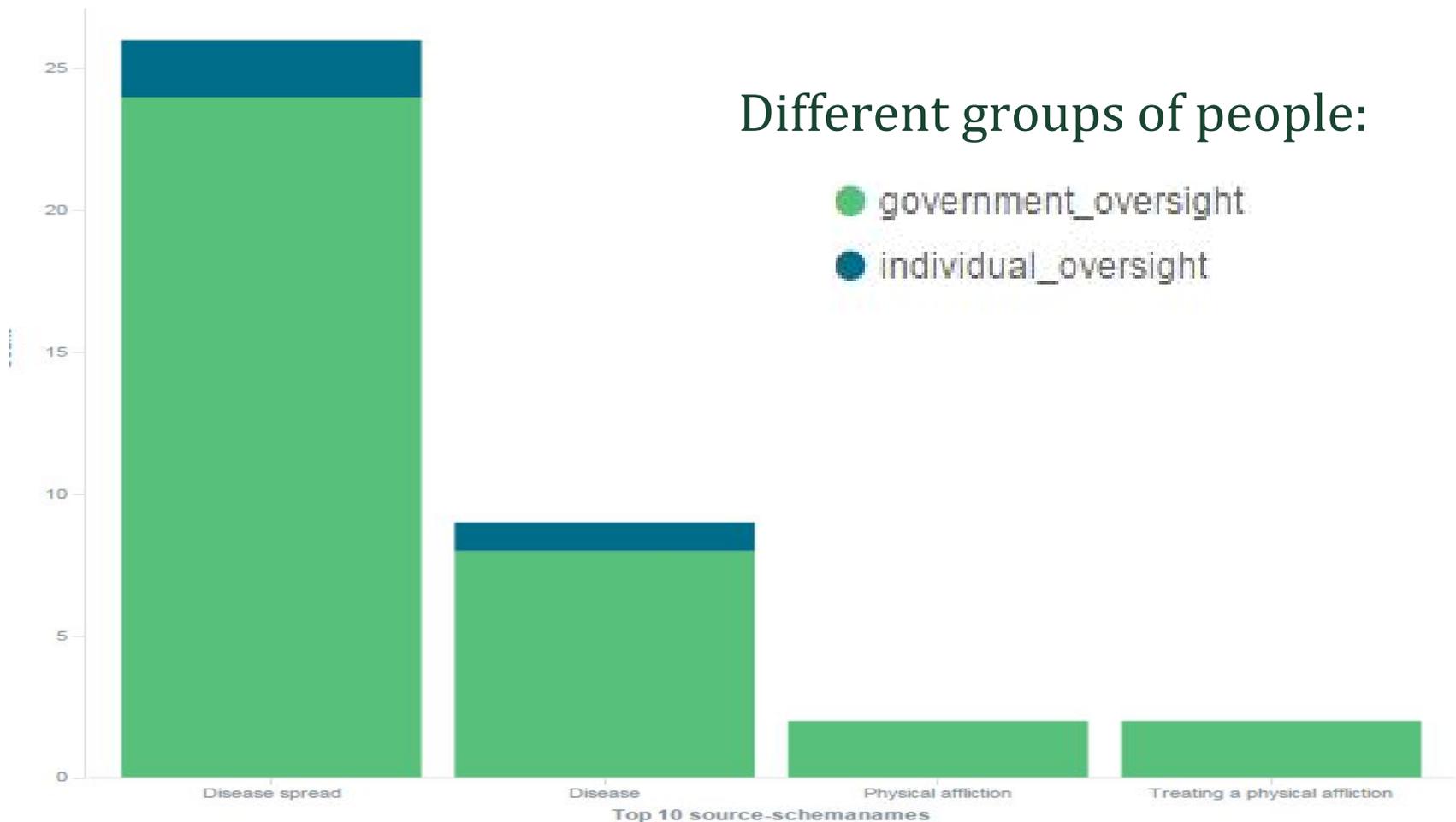
Source: **Disease** schemas and lemmas



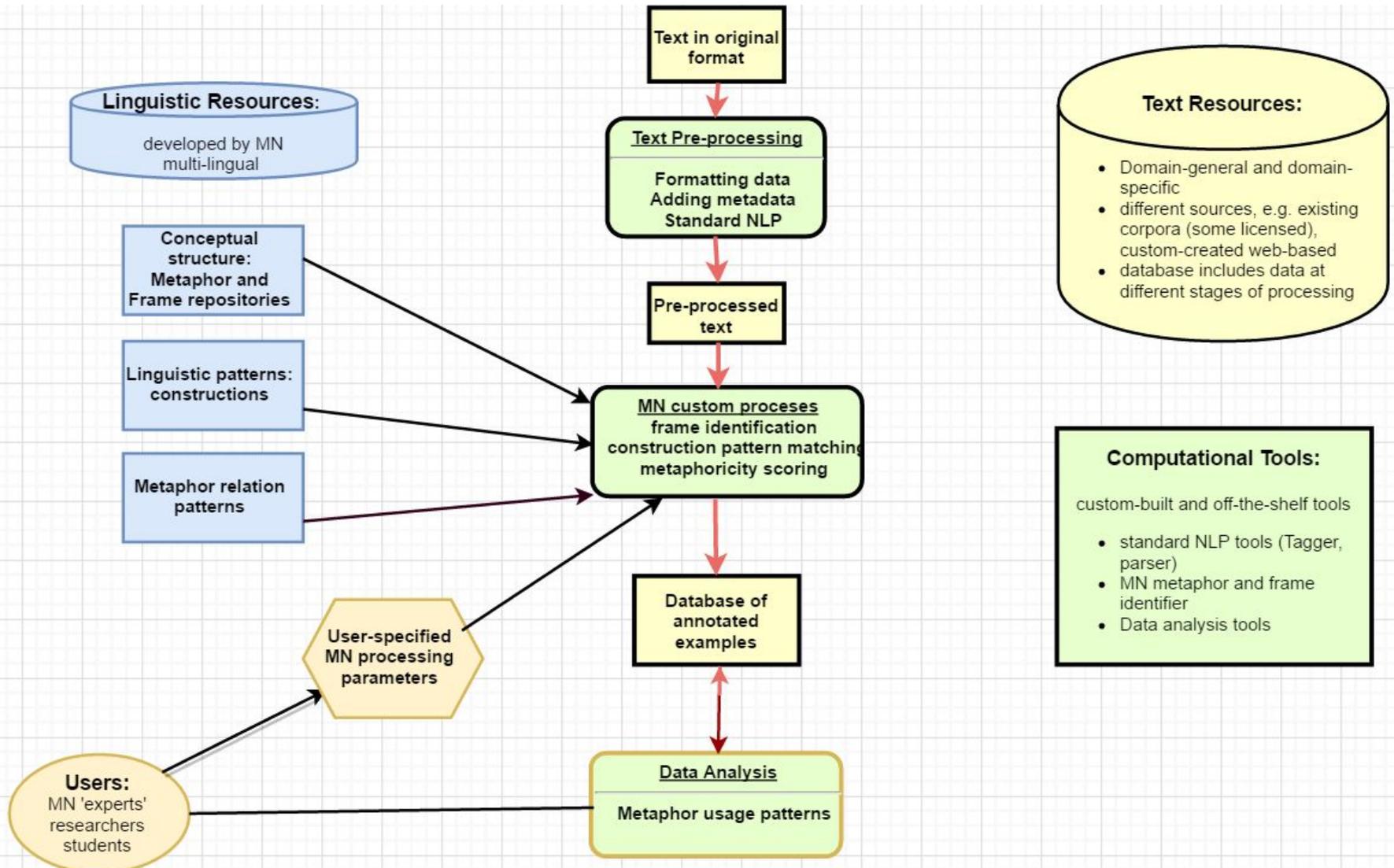
Distribution visualization

Target: **Gun violence**

Source: **Disease**



System overview



Current State of System

Domain-general nature of system makes it possible to readily analyze new domains.

- Prior work includes analysis of several different domains, including poverty, gun rights and gun control, democracy, taxation, and cancer.

High-speed processing makes it possible to analyze different domain-general and domain-specific large corpora

- Thus far, have used various corpora, including large pre-built corpora (Gigaword, British National Corpora) as well as custom-built web-based corpora

Current Objectives

Improvements: structural

- Improve capabilities and robustness of the system (especially the “extraction” system)
- Recast (and streamline) the system in terms of one of the “big data” framework, such as Apache Spark or Storm
- Better modularity to favor faster development and easier experimentation of new features

Improvements: evaluation

- Better methods for evaluation and iterative improvements
- Facilities for creation of gold standard annotations

Current Objectives

Improvements: coverage

- Expand conceptual networks
 - Better UI for collaborative work (now in Semantic MediaWiki)
- More corpora, in multiple languages, with more flexible metadata

Improvements: analytics

- More visualization features, better integrated both with each other and the conceptual network UI
- Ability to export result for further/offline processing

Current Objectives

Improvements: extensions to different uses

- *Separation of concerns*: processing vs. data
 - I.e., let users process their own data (which also avoids some copyright/licensing/privacy issues)
- Increase accessibility, make more user-friendly:
 - MetaNet 'specialists'
 - Researchers in different domains
 - Education in Data Science, Linguistics, etc.

Methods and Challenges

“Cloudification”

- Phase 1: preliminary relocation of current system nearing completion
- Phase 2: start rewriting a subset of the current system

Exploration of different use cases:

- Personal use:
 - Downloadable Docker instance
 - Possibly, publication of source code on online repo
- Use as an online service
 - Batch mode for large text corpora
 - Interactive analysis of single sentences

Methods and Challenges

- Corpora, data sharing
 - Explore ways to connect online resources (esp. social media)
 - Infrastructure to allow storage of users' data for processing and successive collection (temporary or not, depending on availability)
- Incorporating more resources
 - FrameNet
 - VerbNet, WordNet, WikiData, Freebase...
- Use ECG:
 - As a way to express a richer set of constructions
 - For parsing, thus producing deeper semantics and more specific features for the extractor

Thank you!

MetaNet:

- Wiki: <https://metaphor.icsi.berkeley.edu/pub/en/>
- General web page (with publication list):
<https://metanet.icsi.berkeley.edu/>

Center for the Neural Mind and Society: CNMS.berkeley.edu