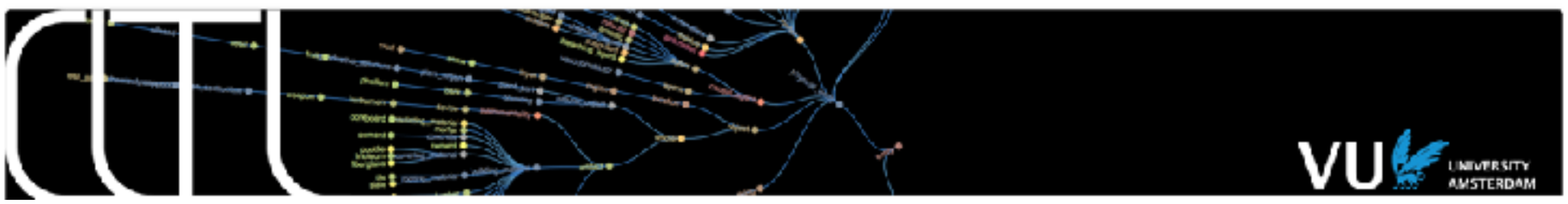
Storylines:
A Computational Framework

Tommaso Caselli & Piek Vossen
t.caselli@{gmail.com;rug.nl}
@tommaso_caselli

Storyline Coding and Annotation Workshop @Narrative Matters 2018
Enschede, July 02, 2018



But I'm not the only one…

- Other storytellers: Antske Fokkens & Roser Morante,
- SPINOZA-NWO Understanding Language by Machines

http://www.understandinglanguagebymachines.org

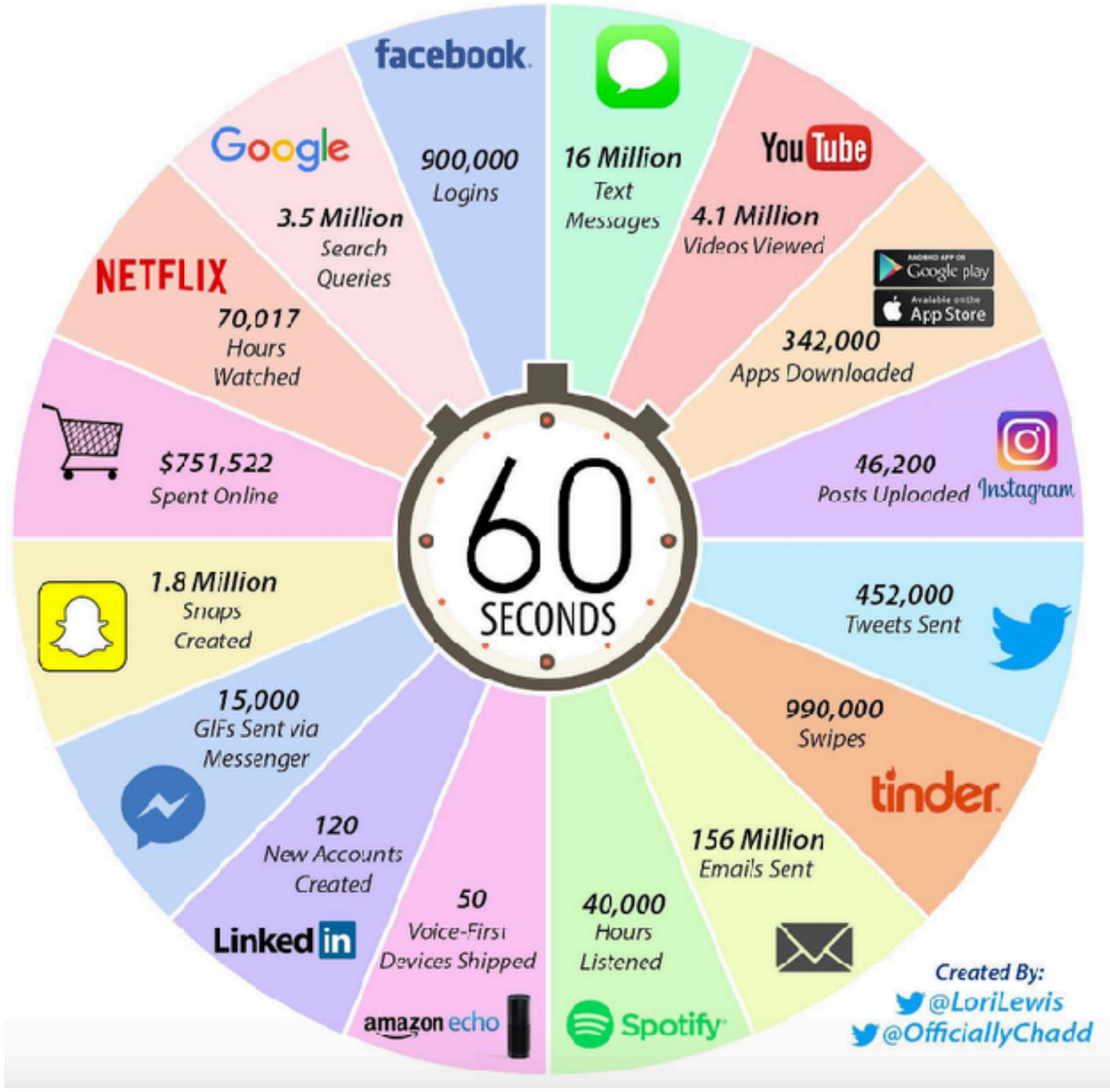
Netherlands Organisation for Scientific Research

VU UNIVERSITY AMSTERDAM

NewsReader

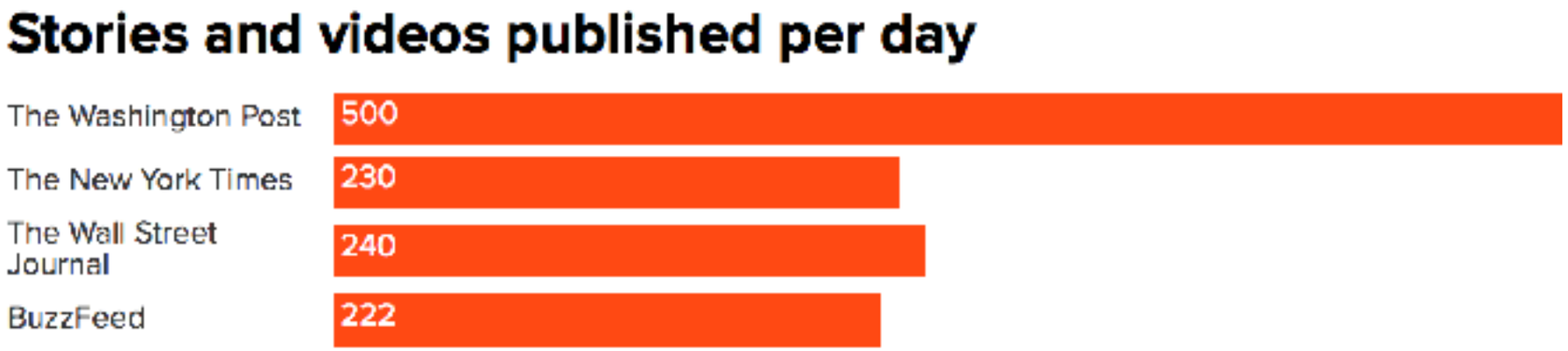
POST HOC ERGO PROPTER HOC

An endless stream of *data*



http://www.visualcapitalist.com/happens-internet-minute-2017/

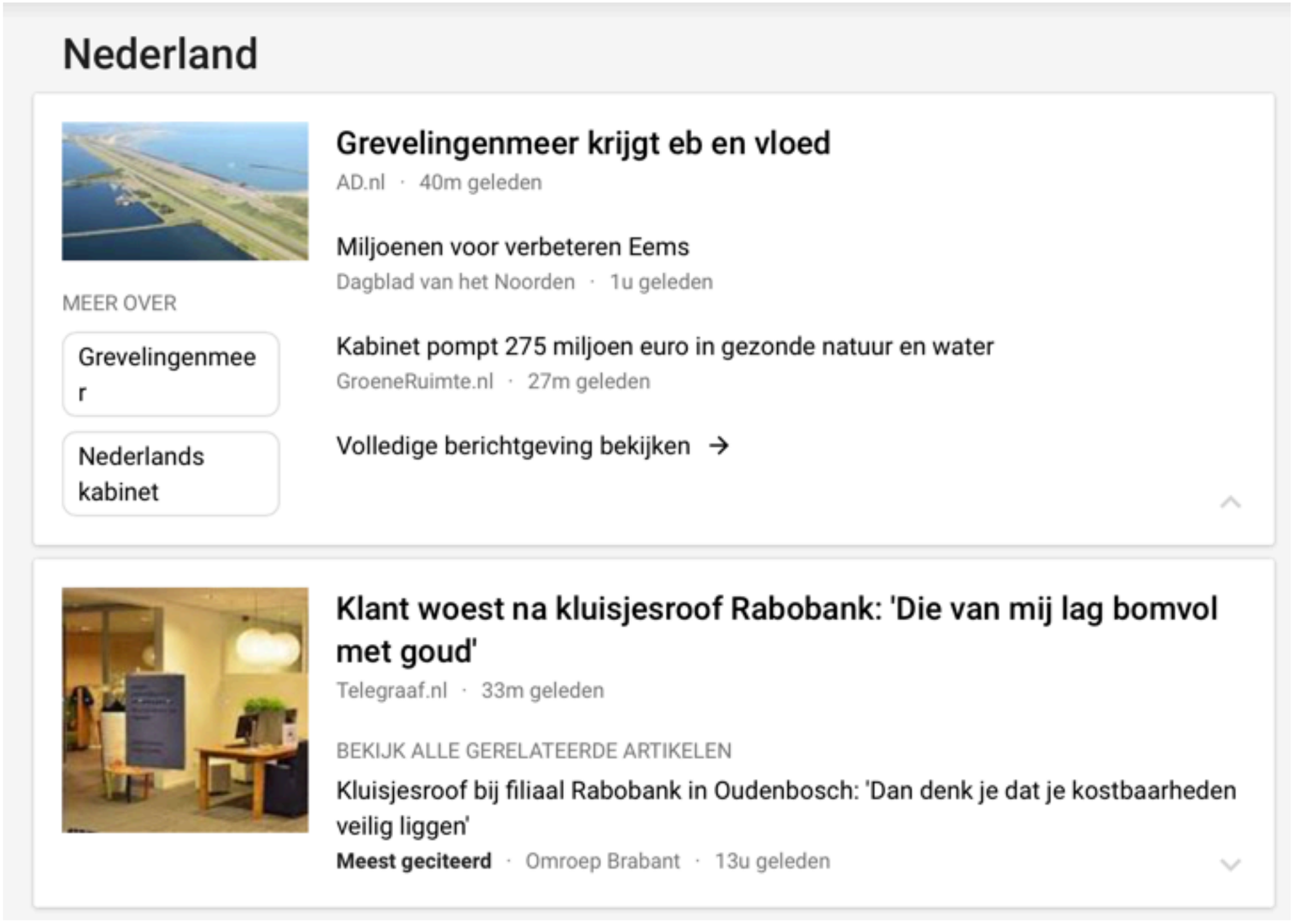
An endless stream of *data*



https://www.theatlantic.com/technology/archive/2016/05/how-many-stories-do-newspapers-publish-per-day/483845/

- Printing the whole WWW will require 305.5 billion pages (*The Washington Post* - https://goo.gl/nPY7db*)*

Google News



https://news.google.com/news/headlines/section/topic/NATION.nl_nl/Nederland?ned=nl_nl&hl=nl&gl=NL

An endless stream of *data*

- Nevertheless, we (*most of us*) make sense of these data:
  - integrate old and new information
  - remove redundant/duplicated information
  - resolve contradictory statements

An endless stream of *data*

- Nevertheless, we (*most of us*) can make sense of these data:
  - integrate/remove/resolve AND
  - develop ***explanatory models***

Explanatory models? Wat?

- *Stories* are the explanatory models we used everyday to make sense of the world
    - everyday we tell each other **stories** to interpret and explain the world

- *Storylines* are a machine-based counterpart of human explanatory models of the world

Humans, “the storytelling animal” (Gottschall, 2012)

- Humans: “*We are such stuff as dreams are made on, and our little life is rounded with a sleep*.” (Shakespeare, *The Tempest*, Act IV, Scene I)

- Stories: What are they made of?

Narratological Framework (Bal, 1997)

- A story:
  - is an instance of a *fabula*, the way an event is narrated
    - *fabula*: chronologically and logically ordered sequence of events, involving one or more actors
    - *actor*: it is a agent of the story, performing or being the target of actions
    - *action*: a change from a state to another
- it has a *focaliser*: the perspective with which the story is told

Narratological Framework (Bal, 1997)

- *Fabula* is a complex device/notion:
  - it has a tripartite structure composed by 3 elements:
    - *exposition and rising action*
    - *climax*
    - *falling action and resolution*

fabula/plot


Climax
Falling Action
Rising Action
Resolution
Exposition

Wait…

Something to agree upon

- Events = ?
- Actors = ?
- Chronologically order = ?
- Logically order = ?
- Perspectives = ?

Something to agree upon

- Events = *things that happen(ed)*
- Participants = *actors of the events*
- Chronologically order = *time*
- Logically order = *causal relations; rising actions; falling actions*
- Perspectives = *opinions; emotions; attribution ("who is saying/thinking what")*

Something to agree upon

- **Events = *things that happen(ed)***
- Participants = *actors of the events*
- **Chronologically order = *time***
- **Logically order = *causal relations; rising actions; falling actions***
- Perspectives = *opinions; emotions; attribution ("who is saying/thinking what")*

Something to agree upon

- What is the *relationship* between these concepts and language?
- How can we (*automatically)* identify them?
- How to integrate all this into a *computational framework* for storyline?

…and… Action!

Events

- "*things that happens in the world*"
- Eventualities (Bach, 1986): events and states
  - ontological primitives
  - dynamic vs static

Events

- Different proposals on how to classify events:
  - their internal temporal properties (Vendler, 1967; *Aktionsart*): *I ran* vs. *I ran to the shop*
  - syntactic and semantic alternations (Levin, 1993): *I loaded my truck with hay* vs. *I loaded hay on the truck*
  - semantic frames, or scenario evoked (Fillmore, 1968; *Frame semantics*)
  - inferential properties (Pustejovsky et al., 2003; *TimeML classes*)

Events

- Identifying eventualities in a textual document is not a trivial task:
  - what is the text span to keep as valid?

"*The police prevented the manifestation*"

"*The police {**prevented the manifestation**}$_{e2}$*"

"*The police {**prevented**}$_{e1}$ the {**manifestation**$_{e2}$}*"

Events

- Identifying eventualities in a textual document is not a trivial task:
  - what is the text span to keep as valid?

    "*The police prevented the manifestation*"

    "*The police {**prevented the manifestation**}$_{e2}$*"

    "*The police {**prevented**}$_{e1}$ the {**manifestation**$_{e2}$}*"

  - which event classes best fit the storyline framework and are useful for their extraction?

Time & Timelines

- The *fabula:*
  - trivially: *what comes first and when*
    - chronological order of sequences of events = *timeline*
- Time and (natural) language have a complicated relationship
  - language is a symbolic, vague, and economical system

Time & Timelines

- Time and temporal relations can be expressed with different means:

- lexical items (words and phrases)

- structural devices

Time & Timelines

- Time & Language:
  - Direct reference: temporal expressions
  - Indirect reference: Tense, Aspect, (Mood), Lexical Aspect (*Aktionsaart*), Connectives (signals),

Time & Timelines

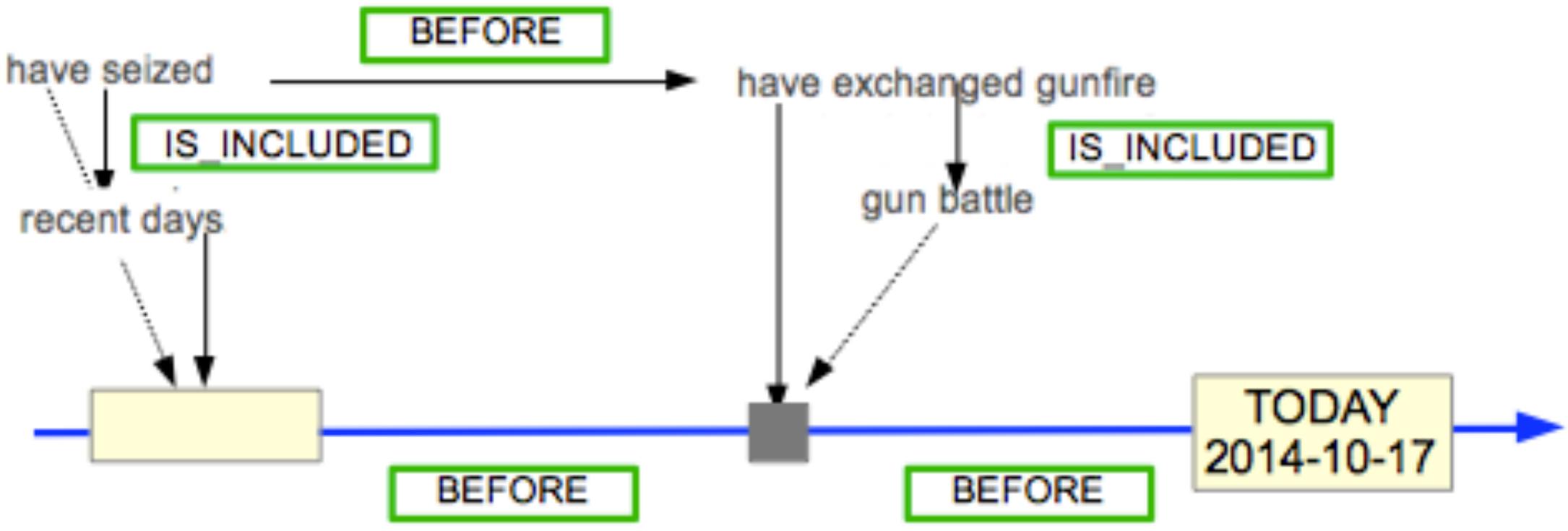
- Time & Language:
  - Direct reference: **temporal expressions**
  - Indirect reference: **Tense**, Aspect, (Mood), Lexical Aspect (*Aktionsaart*), **Connectives** (signals)

Time & Timelines

- Temporal Relations:
  - are stable (time is ordered) - *what is in the past stays in the past*
  - are inferential processes
  - are a by-product of discourse structure (Asher and Lascarides, 1993) - no discourse/text, no temporal relations

Time & Timelines

Ukrainian forces have exchanged gunfire with unknown armed men in the first reported gun battle in the east of the country, where pro-Russian protesters have seized a number of government buildings in recent days.



Time & Timelines

- What granularity of temporal relations?
  - TimeML: 13 values
  - RED: 10 values (3 are temporal & causal relations e.g. BEFORE-CAUSES)
  - CAEVO: 6 values (including VAGUE)
  - ECBStar: 8 values
    - ESC v1.0: 4 values (only anchoring)
  - CaTeRs: 3 values

Logical Connections & Plot Links

- Causality: main ingredient of a narrative
- logically connected events
- loose definition: *a connection between two processes, or states, whereby the first is responsible for the occurrence, or holding as true, of the second, and the second is dependent on the first*

Logical Connections & Plot Links

- Explicit relations:
  - causal connectives (*because, etc*); causation verbs (*trigger, etc*), conditionals.
- Implicit relations:
  - complex nominals (*malaria mosquitos*); discourse structure.

Logical Connections & Plot Links

- Linguistically grounded approach:
  - annotation only when in presence of a connective/signal
    - *Causelines*
- CaTeNa Corpus
- BeCauSe v2.0 Corpus

Logical Connections & Plot Links

- *In a letter, prosecutors **told** Mr. Antar's layers that <u>because</u> of the recent Supreme Court **rulings**, they could **expect** that any fees collected by Mr. Antar may be **seized*** [wsj_788; CaTeNa]

- *expect - [because]- rulings*

Logical Connections & Plot Links

- Commonsense reasoning approach:
  - casualty is interpreted as a subclass of contingency relation
    - remember our definition of causality!!
- Includes more subclasses:
  - enablement
  - prevention

Logical Connections & Plot Links

- Enablement:
    - *John studied hard and finally pass the exam*. [studied]-[pass]
- Prevention
    - *The rain prevented the demonstration* [rain]-[demonstration]

Logical Connections & Plot Links

- To generalise:
    - *rising actions*: relations that express a cause, an enablement, a prevention, or circumstantiality
    - *falling actions*: relations that express an effect, a consequence, a speculation
    - Relations that expresses *Storylines*
        - *make explicit WHY events happened*
        - *contribute to the reconstruction of the plot structure*

Logical Connections & Plot Links

- *The **earthquake killed** 14 and **left** hundred **trapped** in **collapsed** buildings.*

- *earthquake [rising action] killed*

- *earthquake [rising action] trapped*

- *earthquake [rising action] collapsed*

- *collapsed [rising action] trapped*