

A Study in Contradiction: Data and Annotation for AIDA Focusing on Informational Conflict in Russia-Ukraine Relations

Jennifer Tracey, Ann Bies, Jeremy Getman, Kira Griffitt, Stephanie Strassel

Overview

- This poster describes data resources created for **Phase 1** of the DARPA Active Interpretation of Disparate Alternatives (**AIDA**) program
- Program aims to develop language technology that can help humans manage **large volumes of sometimes conflicting information** about world events
- Technology must be capable of building **multiple hypotheses** to account for **alternative interpretations** of what happened
- Informational conflict across a set of data including **multiple languages, genres, and media types** including formal and informal text, video and image
- The corpus described here is designed to support these goals for the **Russia-Ukraine relations scenario** selected for Phase 1 of the program
- Corpus expected to start appearing in **LDC catalog** starting in **2022**, divided into background, practice and eval sets

Scenario/Topic Development

- Political & military relations between Russia and Ukraine in 2014-2015
- Rich in opposing narratives and portrayals of the facts, across a variety of media types and languages
- Several subtopics within the broader domain were selected for data collection and annotation

Set	Topic
Practice	Who Started the Shooting at Maidan? (February 2014)
Practice	Ukrainian War Ceasefire Violations in Battle of Debaltseve (January-February 2015)
Practice	Donetsk and Luhansk Referendum, aka Donbas Status Referendum (May 2014)
Eval	Suspicious Deaths and Murders in Ukraine (January-April 2015)
Eval	Odessa Tragedy (May 2, 2014)
Eval	Siege of Sloviansk and Battle of Kramatorsk (April-July 2014)

Queries

- Who fired on protesters at the Maidan protests?
- Who fired on police at the Maidan protests?
- What groups or organizations were the Maidan shooters affiliated with?
- How many people were shot during the Maidan protests?
- Who owned the weapons used during the Maidan protests?

Facet-level Hypotheses

Query: Who fired on protesters at the Maidan protests?

- H1: Members of the Berkut, a police force loyal to the Yanukovich government, fired on protesters.
- H2: Snipers loosely affiliated with the Ukrainian government fired on protesters.

Topic-level Prevailing Theory: Snipers affiliated with Ukraine's Berkut riot police, under the direction of Aleksandr Yakimenko (who was at the time head of Ukraine's SSU intelligence service) and in collaboration with Russia's intelligence service the FSB, killed at least 53 anti-government activists protesting in Kiev's Independence Square (aka Maidan) on February 20, 2014, using AK-47s and sniper rifles.

Required Events

life.die.deathbyviolentevents
Victim: at least 53 protesters

conflict.attack.firearmattack
Victim: protesters
Attacker: snipers
Weapon: AK-47s and sniper rifles
Place: Maidan
Temporal: on 2014-02-20

contact.collaborate.meet
Participant 1: snipers
Participant 2: FSB
Place: Maidan

conflict.demonstrate.marchprotest
Place: Maidan
Temporal: in February 2014

Required Relations

- snipers **affiliated with** Berkut
- Berkut **affiliated with** Ukraine
- Yakimenko **affiliated with** snipers
- Yakimenko **leader of** SSU
- SSU **affiliated with** Ukraine
- FSB **affiliated with** Russia
- protesters **affiliated with** pro-Maidan

Assessment

Class assessment: Is the system response a **valid entity** of the **type** indicated?

Zero-hop assessment: Is the system response a **valid mention** of the **query entity**?

Graph assessment: Does the system **predicate justification** provided adequate **evidence** for the assigned **type**, argument **role**, and **entity** filling the argument role?

Hypothesis assessment: Judge **events & relations** within system hypothesis for:

- **Relevance:** responsive to some aspect of the statement of information need
- **Coherence:** free of illogical or contradictory information
- **Coverage:** provides enough information to fully express a topic-level hypothesis

Source Data

Text, image & video in English, Russian, and Ukrainian

Background: data of unspecified topic content from the approximate time period of the topics of interest

Practice: docs collected specifically for relevance to practice topics, for use in system development

Evaluation: docs used in system evaluation - some specifically relevant to eval topics plus background data from roughly the same time period and sources which may or may not have general relevance to the wider Phase 1 scenario

	Eng	Rus	Ukr	Mixed	Total
Background	783	1753	1101	6031	9668
Practice	245	518	269	639	1671
Evaluation	752	511	670	63	1996
Total	1780	2782	2040	6733	13335

	Text	Image	Video	Total
Background	8189	13770	943	22902
Practice	1608	10537	479	12624
Evaluation	1996	6194	322	8512
Total	11793	30501	1744	44038

Annotation

- Topic-relevant events, relations with arguments & attributes
- Cross-document coreference and KB linking
- Pilot only: explicit annotation of query-level hypotheses
- After pilot, prevailing theories incorporated into system assessment rather than pre-hoc manual annotation

1 attack event base mention

conflict.attack.firearmattack
Victim: 17 protesters
Attacker: members of the riot police
Place: Kiev

According to Ukraine's security chiefs today *members of the riot police shot 17 protesters in Kiev, gunned down while agitating for the ouster of then-President Viktor Yanukovich.* Ukraine's new authorities have pointed the finger at members of the country's disbanded "Berkut" riot police.

2 additional mentions of riot police entity in text and image



3 coreference and KB linking for all mentions of riot police entity

Berkut riot police
KB ID: 80000093

4 pilot only: judge whether annotation supports hypothesis

Hypothesis: Members of the Berkut, a police force loyal to the Yanukovich government, fired on protesters.
Annotation: Fully supported by the labeled conflict.attack.firearmattack event

Phase 1 Corpus Results

	Docs	Events	Relations	Entities
Background	704	1967	569	4529
Practice	204	1955	1848	8930
Evaluation	248	2456	2029	10508
Total	1156	6378	4446	23967