

From Nodes to Narratives: A Knowledge Graph-based Storytelling Approach*

Mike de Kok^{1,2,*}, Youssra Rebboud^{2,*}, Pasquale Lisena², Raphael Troncy² and Ilaria Tiddi¹

¹Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

²EURECOM, Sophia Antipolis, France

Keywords

Narratives, Knowledge Graphs, Information Extraction, Event-centric-knowledge-graphs

1. Appendix

1.1. Results REBEL Model

Relation	Precision	Recall	F1	Support	Micro F1	Macro F1
Cause	80.00	96.00	87.27	25	94.03	93.13
Enable	94.55	94.55	94.55	55		
Prevent	98.21	91.67	94.83	60		
Intend	96.67	95.08	95.87	61		

Table 1

Detailed result for the validation dataset using the REBEL model

Relation	Precision	Recall	F1	Support	Micro F1	Macro F1
Cause	79.31	100.00	88.46	46	84.65	81.53
Enable	91.67	61.11	73.33	18		
Prevent	84.62	68.75	75.86	16		
Intend	95.83	82.14	88.46	28		

Table 2

Detailed result for the test dataset using the REBEL model

Woodstock'22: Symposium on the irreproducible science, June 07–11, 2022, Woodstock, NY

*Corresponding author.

✉ m.a.h.de.kok@student.vu.nl (M. d. Kok); youssra.rebboud@eurecom.fr (Y. Rebboud); pasquale.lisena@eurecom.fr (P. Lisena); raphael.troncy@eurecom.fr (R. Troncy); i.tiddi@vu.nl (I. Tiddi)

🌐 <https://kmitd.github.io/ilaria/> (I. Tiddi)

🆔 0009-0009-9843-4707 (M. d. Kok); 0000-0003-3507-5646 (Y. Rebboud); 0000-0003-3094-5585 (P. Lisena); 0000-0003-0457-1436 (R. Troncy); 0000-0001-7116-9338 (I. Tiddi)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

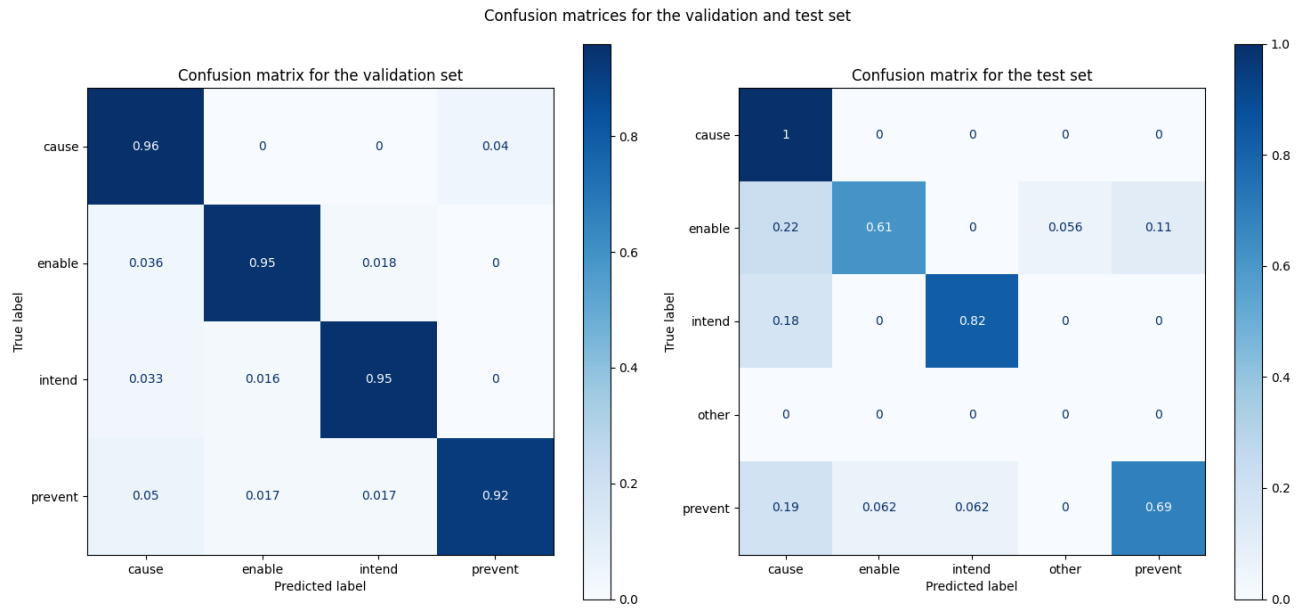


Figure 1: Confusion matrix of relationships in the validation- and testset

Relation	Frequency
has cause	1584
cause of destruction	55
has immediate cause	14
immediate cause of	9
end cause	3
may prevent	2

Table 3
Relationships in the REBEL train dataset

1.2. Event coreference resolution

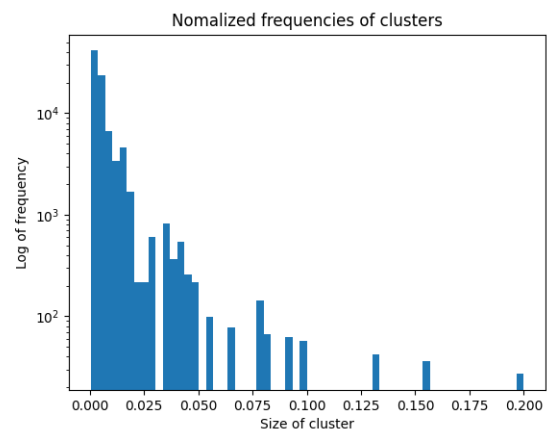


Figure 2: Normalized distribution of the cluster sizes

1.3. JointGT: Qualitative analysis

Triples	Label	Generated
<p>(109 Felicitas, apoapsis, 523329000.0 (kilometres), (109 Felicitas, temperature, 170.0 (kelvins))</p>	<p>(109 Felicitas , with a temperature of 170 kelvins , has an apoapsis of 523329000.0 (kilometres) .), (The temperature of the asteroid called 109 Felicitas is 170.0 kelvins and it has an apoapsis of 523329000.0 km .), (The temperature of the asteroid called 109 Felicitas is 170.0 kelvins and it has an apoapsis of 523329000 kilometres .)</p>	<p>109 Felicitas has an apoapsis of 523329000.0 kilometres and a temperature of 170.0 kelvins .</p>
<p>(3Arena, owner, Live Nation Entertainment), (Dublin, is part of, Republic of Ireland), (3Arena, location, Dublin), (Dublin, is part of, Leinster)</p>	<p>(The owner of 3Arena , Dublin , Leinster , Republic of Ireland is Live Nation Entertainment .), (Dublin is part of Leinster and a city in the Republic of Ireland .Dublin is also home to the 3Arena which is currently owned by Live Nation Entertainment .)</p>	<p>3Arena is located in Dublin , Leinster , Republic of Ireland and is owned by Live Nation Entertainment .</p>
<p>(Barry Cakes, serving size, 30.0 g)</p>	<p>(Barry cakes can be served in 30 gram sizes .), (Serving size for the Barry cakes is 30.0g .), (The serving size of Barry cakes is 30.0g .)</p>	<p>Barry cakes have a serving size of 30.0 g .</p>
<p>(107 Camilla, discoverer, N.R. Pogson), (N. R. Pogson, death place, Chennai), (107 Camilla, periapsis, 479343000.0 (kilometres), (N. R. Pogson, birth place, Nottingham)</p>	<p>(Nottingham born , N.R. Pogson (who died in Chennai) , discovered 107 Camilla, which has a periapsis of 479343000.0kilometres .), (N.R. Pogson discovered 107 Camilla which has a periapsis of 479,343,000 kilometres . Pogson was born in Nottingham and died in Chennai .), (107 Camilla was discovered by N.R. Pogson who was born in Nottingham . He died in Chennai . The periapsis of 107 Camilla is 479343000.0 (kilometres) .)</p>	<p>N.R. Pogson was born in Nottingham and died in Chennai . He discovered 107 Camilla which has a periapsis of479343000.0 kilometres .</p>
<p>(Fulton County, Georgia, country, United States)</p>	<p>(Fulton County is in Georgia , in the United States .), (Fulton County , in Georgia is part of the U.S .)</p>	<p>Fulton County , Georgia is in the United States .</p>

Table 4: Sample of the WebNLG testset and the generated output of the base model.

Triples	Labels	Generated
(demand, cause, benefited)	(The company benefited from continued strong demand and higher selling prices for titanium dioxide, a white pigment used in paints, paper and plastics.)	benefited ” is the cause of the demand .
(center, enable, innovation)	(We recently moved our headquarters and research and development center to Abu Dhabi to benefit from its rapidly growing innovation hub, and with the new AWS Region, we are excited to see it thrive at an accelerated rate.)	Innovation is the enabler of the center .
(made, cause, eaten)	(”Cheerios and Honey Nut Cheerios have eaten away sales normally going to Kellogg’s corn-based lines simply because they are made of oats,” says Merrill Lynch food analyst William Maguire.)	The cause of eating is eaten .
(offer, cause, reimburse)	(The directors said if Messrs. Drabinsky and Gottlieb mail an offer to shareholders by Nov. 22, it will reimburse them a maximum of C\$8.5 million for expenses related to a bid.)	The cause of the offer is to reimburse .
(proclaimed, prevent, bloodshed)	(DR Congo President Felix Tshisekedi on May 6 proclaimed a 30-day ”state of siege” – effectively martial law – in North Kivu and neighbouring Ituri province in a bid to curb bloodshed by the ADF.)	Bloodshed can be prevented .

Table 5: Sample of the FARO testset and the generated output of the base model

Triples	Labels	Generated
(demand, cause, benefited)	(The company benefited from continued strong demand and higher selling prices for titanium dioxide, a white pigment used in paints, paper and plastics.)	The company said it benefited from the strong demand for its products and services from a growing number of customers.
(center, enable, innovation)	(We recently moved our headquarters and research and development center to Abu Dhabi to benefit from its rapidly growing innovation hub, and with the new AWS Region, we are excited to see it thrive at an accelerated rate.)	The United States has created a new innovation center that will help stimulate economic growth and job creation.
(made, cause, eaten)	(“Cheerios and Honey Nut Cheerios have eaten away sales normally going to Kellogg’s corn-based lines simply because they are made of oats,” says Merrill Lynch food analyst William Maguire.)	The company said it made a profit from the sale of the shares because it made enough money to make a profit from the sale.
(offer, cause, reimburse)	(The directors said if Messrs. Drabinsky and Gottlieb mail an offer to shareholders by Nov. 22, it will reimburse them a maximum of C\$8.5 million for expenses related to a bid.)	The company has also announced that it will offer a new credit facility to small businesses, in an effort to reimburse them for the cost of capital expenditures.
(proclaimed, prevent, bloodshed)	(DR Congo President Felix Tshisekedi on May 6 proclaimed a 30-day “state of siege” – effectively martial law – in North Kivu and neighbouring Ituri province in a bid to curb bloodshed by the ADF.)	The United Nations has proclaimed a new resolution to prevent the spread of bloodshed.

Table 6: Sample of the FARO testset and the generated output of the combined model

1.4. Generating text from selected events

Operation Breaking Dawn

- begin date: 2022-08-05
- end date: 2021-08-07
- location: Israel
- participant: Israel Defense Forces
- triples: < *bombardment*, *cause*, *firing* >, < *bombardment*, *cause*, *retaliated* >, < *fired*, *cause*, *killed* >

Base model: *"Israel Defense Forces participated in Operation Breaking Dawn which began on August 5th, 2022 and ended on July 7th, 2022. The bombardment was retaliated by firing."*

Combined model: *"Israel Defense Forces participated in Operation Breaking Dawn which began on August 5th, 2022 and ended on July 7th, 2022. The Israelis were involved in Operation Breaking Dawn and were killed in the retaliation of the Israelis."*

2021 storming of the United States Capitol

- date: 2021-01-06
- location: United States of America
- participant: Proud Boys
- triples: < *measures*, *enable*, *removing* >, < *died*, *cause*, *shot* >, < *addressed*, *cause*, *shot* >

Base model: *"Save America March is located in the United States of America and was launched on January 6th, 2021. Its cause was stormed and it was removed."*

Combined model: *"Save America March is located in the United States of America and was launched on January 6th, 2021. The US government has implemented a series of measures to prevent the spread of disasters."*

2021 Fukushima earthquake

- : date: 2021-02-13
- location: Japan
- triples: < *earthquake*, *cause*, *collapsing* >, < *earthquake*, *cause*, *injured* >, < *broken*, *cause*, *damage* >

Base model: *"2021 Fukushima earthquake, which was caused by collapsing, is located in Japan and was broken."*

Combined model: *"The 2021 Fukushima earthquake, which hit Japan on February 13th,*

2021, injured many people and caused extensive damage and collapsing.”

2021 Sundance Film Festival

- begin date: 2021-01-28
- end date: 2021-02-03
- location United States of America
- triples: < whittled, cause, unveiling >, < awards, enable, stunt >, < coronaviruspandemic, cause, delayed >

Base model: *”2021 Sundance Film Festival, located in the United States of America, began on January 28th, 2021 and ended on March 3rd, 2021-02-03. The event was delayed by the coronavirus pandemic, and awards were whittled.”*

Combined model: *”2021 Sundance Film Festival, which is located in the United States, began on January 28th, 2021 and ended on March 3rd, 2021. It was delayed because of the coronavirus pandemic.”*

Giza church fire

- date: 2022-08-14
- location: Egypt
- triples: < mobilized, cause, taken >, < went, cause, deaths >, < shortcircuit, cause, fire >

Base model: *”The cause of the Giza church fire, which was located in Egypt, was a short circuit. The church was mobilised on the 14th of August, 2022.”*

Combined model: *”The Giza church fire, which occurred on 2022-08-14, was located in Egypt. It was a short circuit that killed several people.”*

Nationwide COVID-19 memorial

- date: 2021-01-19
- location: United States of America
- triples: < pandemic, prevent, swearing – in >, < confirmation, cause, run >, < sworn, cause, saluted >

Base model: *”The nationwide COVID-19 Memorial, located in the United States of America, was sworn in on January 19th, 2021. It was saluted and run to prevent a pandemic.”*

Combined model: *”The National COVID-19 Memorial is located in the United States of*

America and was inaugurated on January 19th, 2021. It was saluted and sworn in to prevent a pandemic."

Save America March

- date: 2021-01-06
- location: United States of America
- triples: < measures, enable, removing >, < emergencies, cause, died >, < stormed, cause, drew >

Base model: *"Save America March is located in the United States of America and was launched on January 6th, 2021. Its cause was stormed and it was removed."*

Combined model: *"Save America March is located in the United States of America and was launched on January 6th, 2021. The US government has implemented a series of measures to prevent the spread of disasters."*

1.5. Annotation of article: Russia launches Iranian satellite amid Ukraine war concerns

An Iranian satellite launched by Russia blasted off from Kazakhstan Tuesday and reached orbit amid controversy that Moscow might use it to boost its surveillance of military targets in Ukraine

Triples: ("Russia", "launch", "satellite"), ("launch", "location", "Kazakhstan"), ("launch", "time", "Tuesday"), ("satellite", "enable", "military surveillance")

Base model: *"The launch of a satellite in Kazakhstan , which has a military surveillance capability , took place on Tuesday ."*

Combined model: *"The launch of a satellite from Russia was on Tuesday , in Kazakhstan . The satellite will enable military surveillance of the country ."*

As Russia's international isolation grows following Western sanctions over its invasion of Ukraine, the Kremlin is seeking to pivot towards the Middle East, Asia and Africa and find new clients for its embattled space programme.

Triples: ("Russia", "cause", "invasion of Ukraine"), ("invasion of Ukraine", "cause", "western sanctions"), ("western sanctions", "cause", "Russia's international isolation"), ("Russia", "intend", "find new clients")

Base model: *"Russia 's invasion of Ukraine is caused by western sanctions and Russia 's*

international isolation . Russia 's intention is to find new clients ."

Combined model: *"The invasion of Ukraine is a cause of Russia 's isolation and western sanctions are a reason for Russia's invasion of Ukraine . Russia also seeks to find new clients in the region ."*

Speaking at the Moscow-controlled Baikonur Cosmodrome in the Kazakh steppe, Russian space chief Yury Borisov hailed "an important milestone in Russian-Iranian bilateral cooperation, opening the way to the implementation of new and even larger projects".

Triples: ("Russian-Iranian bilateral cooperation", "enable", "implementation of new and even larger projects")

Base model: *"The Russian-Iranian bilateral cooperation is enabled by the implementation of new and even larger projects ."*

Combined model: *"The Russian-Iranian bilateral cooperation will enable the implementation of new and even larger projects ."*

Iran's Telecommunications Minister Issa Zarepour, who also attended the launch of the Khayyam satellite, called the event "historic" and "a turning point for the start of a new interaction in the field of space between our two countries".

Triples: ("Iran Telecommunication", "minister", "Issa Zarepour"), ("Issa Zarepour", "attended", "launch of the Khayyam satellite"), ("satellite", "enable", "a turning point for the start of a new interaction")

Base model: *"Iran Telecommunications Minister Issa Zarepour attended the launch of the Khayyam satellite , which was a turning point for the start of a new interaction ."*

Combined model: *"Iran Telecommunication Minister Issa Zarepour attended the launch of the Khayyam satellite , which was a turning point for the start of a new interaction ."*

Nasser Kanani, the Iranian foreign ministry spokesman, said on Twitter that "the brilliant path of scientific and technological progress of the Islamic republic of Iran continues despite sanctions and the enemies' maximum pressure".

Triples: ("western sanctions", "not prevent", "scientific and technological progress of Iran"), ("enemies' maximum pressure", "not prevent", "scientific and technological progress of Iran")

Base model: *"The enemies ' maximum pressure is the scientific and technological progress of Iran ."*

Combined model: *"The United Nations has imposed a maximum pressure on Iran in an effort to prevent the country from achieving scientific and technological progress."*

Iran, which has maintained ties with Moscow and refrained from criticism of the Ukraine invasion, has sought to deflect suspicions that Moscow could use Khayyam to spy on Ukraine.

Triples: ("Iran", "maintained ties", "Moscow", 5), ("Iran", "not criticized", "Ukraine invasion", 5), ("Khayyam", "enable", "spy", 5)

Base model: *"Khayyam is a spy in Iran , which has maintained ties with Moscow . The Ukraine invasion is not criticized in Iran ."*

Combined model: *"Khayyam was allowed to spy on Moscow and Iran , which has not been criticized for its invasion of Ukraine ."*

Responding to the launch, Washington said Russia's growing cooperation with Iran should be viewed as a "profound threat".

Triples: ("Washington", "respond", "launch"), ("Russia's growing cooperation with Iran", "cause", "profound threat")

Base model: *"Russia 's growing cooperation with Iran is the cause of Washington 's response , which is a profound threat ."*

Combined model: *"Russia's growing cooperation with Iran has created a profound threat to the region, and the United States has responded with a missile launch."*

"No third country is able to access the information" sent by the satellite due to its "encrypted algorithm", it said.

Triples: ("satellite", "sends", "information"), ("encrypted algorithm", "prevent", "third countries accessing information"), ("satellite", "has", "encrypted algorithm")

Base model: *"The satellite has an encrypted algorithm which prevents third countries accessing information ."*

Combined model: *"The satellite sends information via an encrypted algorithm , which*

prevents third countries from accessing information .”

Iran is currently negotiating with world powers, including Moscow, to salvage a 2015 deal aimed at reining in Tehran’s nuclear ambitions.

Triples: (“Iran”, “intend”, “salvage 2015 deal”), (“2015 deal”, “prevent”, “Iran’s nuclear ambitions”)

Base model: *“Iran ’s nuclear ambitions are intended to prevent the 2015 deal being salvaged .”*

Combined model: *“The United Nations has imposed a new nuclear deal on Iran in an effort to prevent the country from developing nuclear weapons.”*

Western governments worry that satellite launch systems incorporate technologies interchangeable with those used in ballistic missiles capable of delivering a nuclear warhead, something Iran has always denied wanting to build.

Triples: (“satellite”, “contains”, “ballistic missile technologies”, 9), (“ballistic missile technologies”, “enable”, “delivery of nuclear warhead”, 9)

Base model: *“The satellite contains ballistic missile technologies which enable the delivery of nuclear warheads .”*

Combined model: *“The satellite contains ballistic missile technologies which enable the delivery of nuclear warheads .”*

Iran successfully put its first military satellite into orbit in April 2020, drawing a sharp rebuke from the United States.

Triples: (“Iran”, “launch”, “first military satellite”, 10), (“first military satellite”, “orbit”, “April 2020”, 10), (“first military satellite”, “cause”, “sharp rebuke from the United States”, 10)

Base model: *“The first military satellite , launched in April 2020 , was launched in Iran . A sharp rebuke from the United States was the cause of the first military satellite .”*

Combined model: *“Iran launched its first military satellite in April 2020 , causing a sharp rebuke from the United States .”*

Borisov, who last month replaced bombastic nationalist Dmitry Rogozin as head of the Russian space agency, had acknowledged that the national space industry

is in a "difficult situation" amid tensions with the West. Triples: ("Borisov", "replaced", "Dmitry Rogozin", 11), ("Borisov", "acknowledged", "difficult situation", 11), ("tensions with the West", "cause", "difficult situation ", 11)

Base model: *"Borisov , who was replaced by Dmitry Rogozin , is in a difficult situation which has caused tensions with the West ."*

Combined model: *"Borisov , who was replaced by Dmitry Rogozin , acknowledged a difficult situation with the West because of tensions with the West ."*

1.6. Example Narrative Graph of article: Russia launches Iranian satellite amid Ukraine war concerns

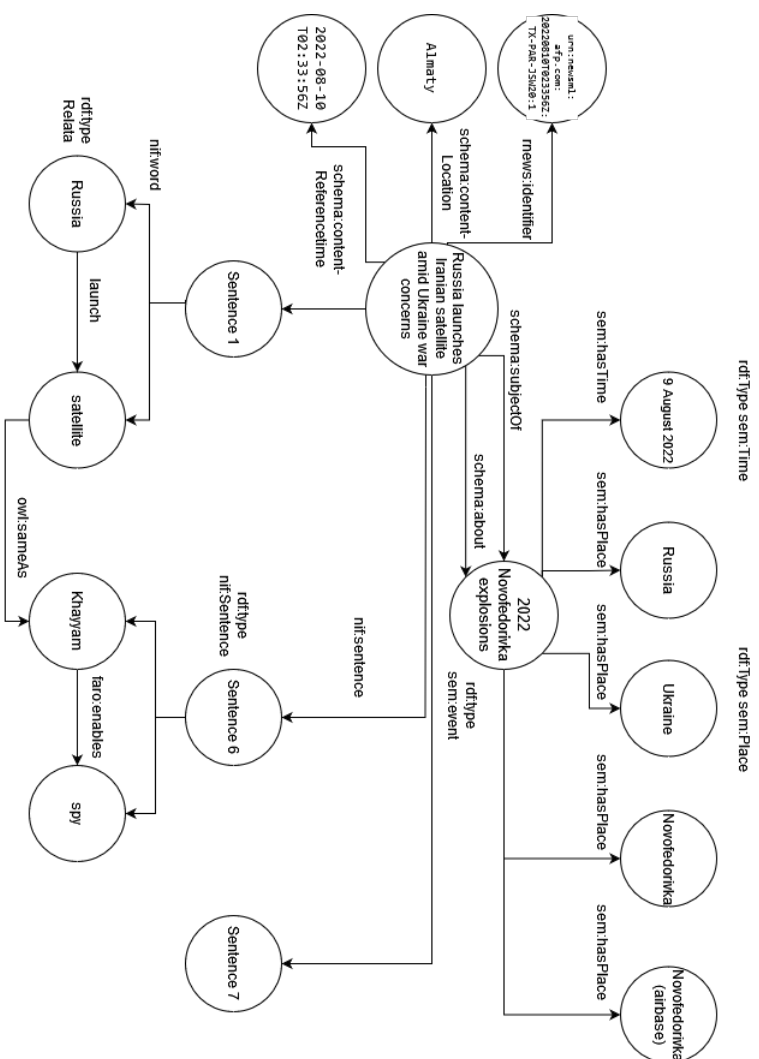


Figure 3: Part of the constructed narrative graph of the article "Russia launches Iranian satellite amid Ukraine war concerns"

1.7. Overview Narrative Graph

