A Parameters

- **Sampling parameters**: for the multiple summaries generation, we varied the sampling parameters as follows :

  - Temperature from 0.5 to 1 with a step of 0.1 $\implies T = [0.5, 0.6, 0.7, 0.8, 0.9]$;
  - Top-p from 0.75 to 0.96 with a step of 0.05 $\implies Top - p = [0.75, 0.80, 0.85, 0.90, 0.95]$;
  - Top-k from 40 to 70 with a step of 10 $\implies Top - k = [40, 50, 60]$

  We generated 77 summaries in total for each example (75 using sampling methods + greedy generation + beam of size 4 generation).

- **Generation parameters**: for each generation, we took 1024 tokens as the size of the input document and 128 tokens as the size of the output summary.

- **Entities Tags**: the different entities tags extracted with the FLERT system are : cardinal, date, event, facility, geo-political entity, language, law, location, money, groups, ordinal, organisation, percent, person, product, quantity, time, and work of art.

B Cohen kappa agreement between annotators

We computed the Cohen Kappa between the annotator for the annotation of Rouge max summaries selected and report the result in the table below:

	A1 - A2	A1 - A3	A2 - A3
kappa	0.6833	0.6468	0.6302

C Variation detection heuristics

We have introduced several variation detection heuristics:

- **Text processing**: this consists of *cleaning* the input text from all non-alphabetic characters.

- **Stemming**: to replace words by their radical or root.

- **Levenshtein distance** [1]: for a given word and text, the aim is to find words in the text which are close to the input word when a character deletion, replacement, transposition or insertion is applied. We consider that two word strings can be considered as equivalent if their Levenshtein distance is below a threshold.

- **Alpha2Digit**[2]: this uses a python package which converts numbers written as letters into text written as digits.

- **SplitWord**: the idea is to process multi-word entities by looking to each word individually to check if they belong to the source document.

- **BERTScoreCheck**: As BERTScore is based on contextual representations, the idea is to find an entity in the source document that is close to the entity we want to test.
  As BERTScore is a measure between sentences, we have developed a method for using it for variation detection. Figure 1 gives an example. Let's suppose that the sentence $S1$ comes from the document, and that "*Netherlands*" is a candidate variation of an entity in this sentence. Here, we replace "*Netherlands*" with each entity to create new sentences and then calculate the BERTScore of the sentences obtained in relation to the starting sentence $S1$. The highest score between $S1$ and $S2$ means that "*Netherlands*" has more chance to be a variation of the entity "*Holland*" than "*Isabelle*".

---

[1] https://norvig.com/spell-correct.html
[2] https://pypi.org/project/text2num/

$$S_1 = \textit{\textbf{Isabelle}} \textit{ come from } \textit{\textbf{Holland}}$$

$$S_2 = \textit{\textbf{Netherlands}} \textit{ come from } \textit{\textbf{Holland}}$$

$$S_3 = \textit{\textbf{Isabelle}} \textit{ comme from } \textit{\textbf{Netherlands}}$$

$$BERTScore(S1, S2) = 92.30 < BERTScore(S1, S3) = 98.13$$
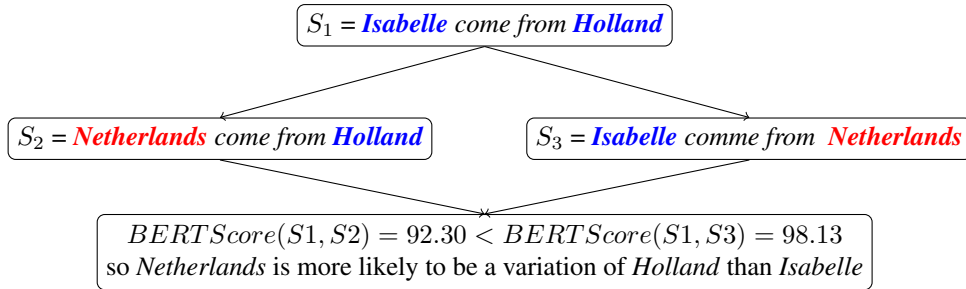so *Netherlands* is more likely to be a variation of *Holland* than *Isabelle*

Figure 1: Variation detection using BERTScore

D   Human evaluation of NEHR selected summaries

The adjudication results of each document-summary pair for 5 examples out of the 10 annotated are shown below (in green the entities used in the correct context and in red the incorrect ones):

Document	Media playback is not supported on this device Rangers, who have already clinched the Championship title and promotion, have now gone three games without a victory. Halkett's first goal of the season - the final piece of action in the first half - won the points. And it will give second-bottom Livi renewed hope that they can avoid relegation via the play-offs. This was only the fifth defeat of Rangers' league season, but Mark Warburton's side have now not won since beating Celtic in the Scottish Cup semi-final. Livingston had gone into the game with only 17 fit players and gave a first start to 16-year-old Matthew Knox. The highly talented midfielder has actually been training with Rangers, who had a bid in the region of £100,000 rejected for his services. Knox is also interesting Chelsea, Everton, Manchester United and Tottenham Hotspur. The fact that virtually the length of the division lies between the teams was underlined by the early possession and pressure. It was all Rangers from the first whistle and Andy Halliday, Gedion Zelalem and Michael O'Halloran all had opportunities to score. Incredibly, though, it should have been Livingston taking the lead. Zelalem lost the ball in his own half and Danny Mullen suddenly found himself clear, but the midfielder shot wide. The warning bells were ringing for the division's champions when Sam Stanton also had a dancing run and a shot past the goalkeeper's left port. The half did, indeed, come to a remarkable finish. In its dying seconds, young Knox whipped in a fabulous free-kick from the left and Halkett met it with the perfect header to flash the ball beyond diving goalkeeper Wes Foderingham. Rangers supporters in the crowd of 5,021 looked like they could hardly believe the scoreline. The second half picked up on the first half script. The Ibrox side continued to huff and puff and fluff chances. Livingston grew in confidence as the league's champions produced one of their poorest showings of the term. The West Lothian side should have had a second goal, but Liam Buchanan put the simplest of headers over the bar from inside the six-yard box. The same player, one on one with Foderingham, was only denied by a block from the the goalkeeper. Livingston were worthy winners. Rangers looked a shadow of the team that grabbed the Old Firm headlines and are preparing for the cup final against Hibernian.
Reference	Champions Rangers' slump in form continued thanks to Livingston defender Craig Halkett's first-half header against his former club.
Beam 4	Matt Halkett 's last-gasp header gave Livingston victory over Scottish Premiership leaders Rangers at Ibrox
Min NEHR	Livingston beat Rangers to move into the top six of the Scottish Premiership and inflict the first league defeat of the season on Ibrox side.
Document	Their 196-run first-innings lead came mainly thanks to Tom Kohler-Cadmore (102) and Ed Barnard (59). Further contributions down the order allowed the visitors to reach 403 all out for five batting bonus points. Glamorgan started well with Jacques Rudolph and Nick Selman putting on 74 for the first wicket, but they ended 141-3 at the close, 55 runs behind. Worcestershire claimed three wickets for 34 runs to reduce Glamorgan to 108-3, but Colin Ingram (41 not out) and Aneurin Donald saw out the rest of the day. Glamorgan bowler Michael Hogan told BBC Wales Sport: "It seemed to go through different stages at different times. It was a little trickier at times and then got easier as ball got softer. The boys toiled hard but it could have been a bit better. "It was good to see Lukas Carey doing well. He has done a lot of work in pre-season and he has worked really hard on his fitness and his technique, so it is good to see him being given a chance and taking it with both hands. "I'm fitter after a winter off, probably lacked a bit of bowling conditioning. I've had a few weeks off with injury so probably I got bit tired towards the end. "We thought with the sun out after about 35 or 40 overs it got a bit easier, but they have got a good attack so we will have to absorb a bit of pressure. I think the boys can do the job and if the boys can build a good total we can hopefully bowl them out."
Reference	Worcestershire batted themselves to a good position on day two against Glamorgan in Cardiff.
Beam 4	Worcestershire made a solid start to their reply to Glamorgan 's first innings total on day one of the County Championship match at Chelmsford
Min NEHR	Worcestershire 's batsmen continued to dominate on day two against Glamorgan at the County Ground

Document	Ronald Dickens, 47, followed a woman in her 20s into Newbold Road in Rugby in November 1985 before dragging her into a lane and raping her. The incident completely changed her life leaving her wondering for 30 years who assaulted her, police said. A cold case review using new DNA techniques led to the breakthrough. Warwickshire Police said its major crime review team identified Dickens, with help from forensic investigators, using techniques not available at the time of the crime. More updates on this and other stories in Coventry and Warwickshire Dickens, of Phipps Avenue, Rugby, admitted rape at at Coventry Crown Court on Monday. Det Sgt Paul Thompson said: "Ronald Dickens subjected the victim to a horrific attack which has altered the course of her life. "This incident completely changed her life and she has been looking over her shoulder for the past 30 years not knowing who assaulted her. He also urged anyone else who has been a victim of a sexual offence to come forward to police to report it.
Reference	A man has been jailed for eight years for the brutal rape of a woman in 1985, thanks to advances in DNA and forensic investigations, police say.
Beam 4	A man has admitted raping a woman in Warwickshire 30 years ago
Min NEHR	A man has been jailed for a minimum of 30 years for a "horrific" rape in 1985
Document	The footage shows gunmen beating and shooting a group of prisoners who were cowering on the floor. It has been alleged that Islamist militants carried out the attack after seizing army checkpoints on Thursday. Unconfirmed reports say troops have now quit all bases near the strategic northern town of Saraqeb. The town lies near the main roads linking Aleppo to Damascus and the coastal city of Latakia. The army, meanwhile, continued its air strikes across Syria on Thursday. In all, more than 150 people reportedly died in fighting, said the Syrian Observatory for Human Rights (SOHR), a UK-based activist group. The SOHR said that among the victims were more than 70 government soldiers, 43 civilians and 38 rebels. The claim has not been independently verified. In other developments: The alleged shootings took place after the rebels overran the strategic army checkpoints between on Thursday. The video appears to show agitated rebels kicking and pushing the soldiers or pro-government militiamen, known locally as "shabiha", to the ground inside one of the seized buildings. Shots are then fired into the cowering mass of bodies. Rupert Colville, of the UN High Commission for Human Rights, said it appeared that the victims "were no longer combatants and therefore, at this point, it looks very like a war crime". He added that the video, if proved to be genuine, would almost certainly form part of a future prosecution. Rights group Amnesty said in a statement: "This shocking footage depicts a potential war crime in progress, and demonstrates an utter disregard for international humanitarian law by the armed group in question." No group has so far admitted carrying out the alleged killings. However, a rebel fighter from Idlib province, Abu Abdul Rahim, told the Guardian that a Salafi-jihadist group was behind the killings, which he said had occurred in al-Nayrab, to the west of Saraqeb. He said Salafists of the Dawood brigade and Suqur al-Sham did not answer to any military council affiliated to the rebel Free Syrian Army. For months, activists have reported similar summary killings by government forces virtually every day. But there has been mounting evidence of similar tactics being used by some rebel groups too, although many have signed a code of practice banning such abuses, says the BBC's Jim Muir in neighbouring Lebanon. US Secretary of State Hillary Clinton earlier warned that radical Islamist fighters were trying to hijack the Syrian revolution. The comments have drawn an angry response from some opposition leaders, who say that it is the failure of the outside world to support the uprising with practical help that has left the field open to the radicals. The SOHR says more than 36,000 people have been killed since protests against President Bashar al-Assad erupted in March 2011. The SOHR is one of the most prominent organisations documenting and reporting incidents and casualties in the Syrian conflict. It says its reports are impartial, though its information cannot be independently verified.
Reference	A video appearing to show Syrian rebels murdering soldiers or pro-government militiamen could be evidence of a war crime, the UN has said.
Beam 4	The UN 's human rights chief has called on Syria 's government to investigate a video purporting to show the killing of dozens of pro-government soldiers.
Min NEHR	A video showing what appears to be the summary killing of more than 50 Syrian government soldiers in a rebel-held town is being investigated by the UN
Document	AB de Villiers smashed 71 off 29 balls and Hashim Amla an unbeaten 69 off 38 as the hosts chased down 172 with 5.2 overs to spare in Johannesburg. England lost their last seven wickets for 14 runs as they were bowled out for 171 despite Jos Buttler's 28-ball 54. Eoin Morgan made 38 and Joe Root 34, while Chris Morris claimed 3-26. Defeat was England's fifth in succession in limited-overs cricket, and means they ended a tour on which they won the Test series with losses in both the one-day and T20 series. It was their final competitive game before the World Twenty20 in India starts on 8 March, while South Africa face Australia in a three-match T20 campaign on home soil from 4 March. Bayliss: 'It was men against boys' The manner in which England were dismantled by De Villiers and Amla dented the optimism built up during a successful 12 months in the limited-overs format. Captain Morgan said: "It's just a blip - we've had such a good run of things in white-ball cricket. "I honestly think we are on the right track playing a brand of cricket that can win an ICC event, whether it is this one coming up or one in 2017." If England had any hopes of defending an apparently testing total at the Wanderers, they were ended by a riotous 125-run opening stand between De Villiers and Amla spanning 8.2 overs. De Villiers will rightly take the headlines for a remarkable display of hitting, marrying power with placement against an increasingly ill-disciplined attack. He flicked the first ball of the innings from Reece Topley for four and plundered 22 off the next over from Chris Jordan on his way to a 21-ball fifty, the fastest by a South African in T20s. "I really enjoy it up the order," said De Villiers. "You can play normal shots and get value for your strokes. That's where I want to bat." By the time he sliced Adil Rashid to wide long-off, he had scored 60 of his 71 runs in boundaries, and victory was as good as sealed. Such was his dominance that Amla's elegant 27-ball half-century went largely unnoticed, while Faf du Plessis had the luxury of sealing victory with a cover-driven four off Jordan. Jordan's 2.4 overs cost 48 runs, while Topley, whose mistake allowed South Africa to win the first T20 in Cape Town on Friday, leaked 32 from two wicketless overs. As well as De Villiers and Amla batted, England's defeat was arguably founded on their remarkable batting collapse. A slide from 157-3 to 171 all out in 18 deliveries was largely of their own making, hastened by a rash of wild strokes on a reliable surface offering pace and bounce. Only Joe Root, with 34 off 17 balls, Eoin Morgan (38 off 23) and Buttler, who smashed four fours and four sixes, emerged with any credit. This is how the collapse unfolded:
Reference	South Africa swept to a comprehensive nine-wicket victory over England in the second Twenty20 to seal a 2-0 series win.
Beam 4	South Africa completed a 3-0 Twenty20 series whitewash over England with an eight-wicket victory in the final match of the tour.
Min NEHR	England 's hopes of winning the World Twenty20 suffered a severe blow as South Africa comfortably beat the visitors by eight wickets in their final one-day international.