

Summary and Conclusions

Summary

The Research Question of this study is whether text mining of media messages can yield interesting insights into citizens' trust in the *rule of law*. Since the media deliver images of political, social and economic developments that may affect trust in the rule of law, the following sub-questions are raised:

- 1,2. whether text mining enables the assessment of the images that citizens would hold if they were to rely completely on the news (1) brought by traditional media such as newspapers and (2) by Internet-based media like Twitter;
3. whether the images of developments that are provided by the media match the data provided by statistical agencies;
4. whether text mining yields valid and reliable results;
5. which media are important if one wants to apply text mining to learn about images of developments that could affect trust in the rule of law;
6. whether media coverage does indeed affect citizens' trust in the rule of law.

The *rule of law* is the best translation for the Dutch word *rechtsstaat*. Whereas the Anglo-Saxon concept *rule of law* is usually straightforwardly defined as government by law – thus as adherence to the due process of law – the Dutch concept *rechtsstaat* appears to be rooted more strongly in De Montesquieu's *trias politica*, and in the German philosophy of law. Dutch treatises on the *rechtsstaat* show a fair amount of agreement that the *rule of law* should entail a *constitution* that guarantees *civil rights* to citizens and that moreover guarantees the *separation of powers* between the legislature, the executive, and the judiciary, with the guarantee of an independent role for the judiciary (Voermans & Emmerik, 2011). *Trust in the rule of law* based on its Dutch definition rests on citizens' belief that civil rights are guaranteed by the constitution, and on citizens' belief that there is a separation of powers, with an independent role for the courts. Trust in the rule of law is greatest indeed among citizens who believe that the most important event in the history of the Netherlands was one of the following events that shaped the current Dutch constitution (i.e., constitutional monarchy 1813-1815, parliamentary democracy 1848, general suffrage also for women 1917-1919) (Survey Research VU/IntomartGfK, 2012).

Important developments in the context of this study are therefore the developments in media images with respect to the separation of powers between Parliament, the government and the judiciary, and with respect to civil rights. Media images can be broken down into four different aspects, which successively were raised in chapters three to six of this study.

- a. How much *attention* is paid by the media to the institutions that guarantee the rule of law and to the issues that are of paramount importance in the context of the rule of law?
- b. How are these institutions (and the actors who represent them) and these issues *framed* in the news, i.e., which of them are mentioned in media messages at every turn in the same breath?

- c. To which institutions and actors do the media attribute *authority*, literally judged based on who appears to be the author of the news?
- d. Which *sentiment*, which tenor or tone, does the news express about these institutions, about these issues, and about the relationships between these institutions and topics?

To show how text mining can contribute to new insights into the development of trust in the rule of law, the news from newspapers, television, news websites and Twitter will be compared, as well as the news from three periods (1993-1994, 2004-2005, 2009-2012).

The choice for specific text mining approaches is motivated by the imperative to build on the legal and political science literature on the rule of law when comparing different periods, different media types, and different media titles, i.e., in measuring them using the same yardstick. A top-down approach was chosen, in which everything in the examined texts will be related to an *ontology* that reflects the concepts from the literature on the rule of law (see Appendix B for the ontology, and chapter 2 for its explanation). The analysis of media authority and the tenor of the coverage (chapters 5 and 6) depends on the choice for automatic parsing of all texts by means of the Dutch grammar parser ALPINO (Bouma & North, 2005). Parsing is necessary to determine to whom media authority belongs, or to find out who is against, or precisely in favor, of what, or of whom. However, errors are inevitable in the automatic parsing of complex sentences: "The accusation of Wilders' lawyer Bram Moszkowicz entailed that Schalken had attempted to influence Jansen by trying to persuade him of the fairness of the decision to commend the OM (=Dutch Prosecutor) to prosecute Wilders for his statements." Based on the law of large numbers, reliable statements can still be made at the level of entire periods or entire media. The rules for the construction of Dutch tweets deviated so much from the rules of grammar for spoken Dutch, that the analysis of social media in terms of media authority and sentiment analysis could solely be based on websites.

The sub-questions 1 to 6 of this text mining study will now be answered, after which the overarching Research Question will be answered.

1.2: Images of developments: attention, frames, media authority and sentiments

A media attention analysis (see Chapter 3) shows that the role of political parties gradually became stronger in the news. The attention for the judiciary as compared with the attention for the legislative and the executive, which are the two other powers in the *trias politica*, did not increase. In terms of media authority (chapter 4), the authority of the judge has increased somewhat, but nevertheless the advantage of the OM (the Prosecutor) held over the judge, and especially the advantage of the advocacy over the judge in terms of media authority did increase.

New issues evolved during the twenty-year research period of 1993-2012, some of which also practically disappeared. The research period ends as it began: the leader of the largest anti-immigration party is acquitted in court. Janmaat was acquitted in 1993, Wilders in 2011. In the period in between, the news changed dramatically, especially between 1993-1994 and 2004-2005. After 9/11 and after the murder of Pim Fortuyn on 6/5/12, a rougher news climate prevailed, with much attention for terror and violence. Thus, the rule of law became a much-debated topic in the news.

In the first place, civic rights as the basis for the rule of law came under discussion in the news after 9/11. In judicial treatises on the rule of law (e.g. WRR, 2002) the implicit assumption was that civic

rights such as the freedom of religion, freedom of expression and the right to non-discrimination reinforced each other in the tolerant Netherlands. It came as a surprise that the discussion in the news about (varying combinations of) terrorists, Islamic fundamentalists, immigrants and criminals culminated in the consensus that civic rights collided with each other and, moreover, were abused. The question raised by Pim Fortuyn as to whether Article 1 of the Dutch constitution on anti-discrimination should be abolished in favor of the freedom of expression for the Dutch gave rise to the position of the popular website GeenStijl.nl and the popular news magazine PowNews that it is acceptable for political statements to be “oversimplified and unnecessarily hurtful”. Ayaan Hirsi Ali came to discuss the freedom of education and the freedom of religion, because Muslim fundamentalists could abuse them. The Court of Arnhem let trafficker Saban B. on weekend leave. In the years 2009-2012 journalists directed their arrows on the implementing agencies who released psychiatric criminals quite early (prisons, TBS institutions (institutions for the detention of psychiatric criminals at Her Majesty’s pleasure))(cf. Chapter 6).

It is only in recent years that other topics are ‘trending’ in the news, such as the sexual offences against children and adolescents in the Catholic Church. Gradually privacy issues pop up in the news more often, as well as various types of cybercrime that may endanger privacy.

In the second place, the separation of powers as the basis for the rule of law came under discussion in the news. An analysis in terms of media authority (chapter 5) shows that the authority of the executive (the government) declined severely during the twenty-year research period, while the authority of parliamentarians increased. Media authority attributed to the judge increased somewhat, but nevertheless the gap between the authority of the Prosecutor and the judge, and especially the gap with the authority of the advocacy increased to the detriment of the authority of the judge. If an *independent* judiciary is perceived as a powerful law-enforcer, then it must be concluded that the news suggests that the court is hardly independent because of the judge’s soft decisions as compared to the tough demands of the Prosecutor. An alternative conception of judicial independence is that the judge is allowed to relax or pass over the sentences demanded by the Prosecutor. In the news after 9/11, the media protested often against this alternative interpretation of judicial independence (e.g. no lifelong sentence for the assassinator of Pim Fortuyn, acquittal for terrorist Samir A.). A semantic network analysis (chapter 6) shows that in 2004 and 2005 the government was seriously engaged in tracing criminals. The pattern in the news from this period shows that the Prosecutor became the paramount hunter for terrorists and criminals, but the Prosecutor also has to deal with the opposition of judges. The confrontation between the Prosecutor and the judge remains prominent in the research period 2009-2012. An analysis of Associative Framing (chapter 4) shows that after 9/11 the judge gradually drowned in political waters: the probability that in an article about the court no politician is mentioned decreases over the years. In the news a raging debate about the independence of the judge takes place, quite unlike the reassuring legal reflections in judicial treatises on the rule of law. Does the judge's independence mean that attempts to prove that politicians should comply with their 'rule of law', or does this concept refer to the judge’s duty to refrain from political interference and from patronizing politicians? Only the full acquittal of Wilders on June 23, 2011 put an end to this lingering debate (pretence 2007, process 2010-2011).

All in all, text mining offers an accurate and differentiated picture of messages in traditional media such as newspapers and internet-based media such as Twitter, although some (usually subjective)

historical knowledge is indispensable for interpreting the abundance of quantitative trends and patterns produced by text mining.

3: Does media coverage match with time series of statistical agencies?

Most developments that are outlined in the media have no counterpart in time series gathered by statistical agencies (CBS, SCP, Eurostat). The trend of the number of violent crimes is an exception. The number of crimes shows a gradual decrease (Wittebrood & Nieuwbeerta, 2006). The media coverage in the years 2004-2005, however, in many ways shows a peak relative to the media coverage in the earlier (1993-1994) and later research period (2009-2012). This peak in the news, which consists of increased attention for terrorists, Muslim fundamentalists, immigrants and criminals, derives from a strong role in the news for the Prosecutor, and for the Minister of Justice. In 2004 and in 2005, the Minister of Justice becomes the second most newsworthy political actor, after the Prime Minister. In the years before and the years after 2004-2005, the Minister of Justice often did not even enter the political top 10 of media publicity (Kleinnijenhuis, et al, 2009). Time series from statistical agencies are clearly less important for news selection than the political agenda and the social agenda, which were dominated after 9/11 by terrorists, Islamic fundamentalists, immigrants and criminals. Whatever the case, the media coverage is not consistent with time series of statistical offices.

4: Validity and reliability of obtained results with text mining

Chapter 2 and the appendices elucidate which types of validity tests and reliability tests were performed. The tests indicate that text mining often leads to erroneous conclusions on the level of specific news items and specific sentences. Results per sentence and even per news item are fairly unreliable in some aspects. If a specific text mining approach (i.e., chapter 5, media authority; chapter six, sentiment analysis in the context of semantic network analysis) presupposes the outcomes of various preprocessing tools (e.g. sentence boundary allocation, part-of-speech tagging, automated grammar parsing, sentiment detection) then possible sources of errors easily accumulate.

Even at a higher aggregation level (medium type per period, or even complete periods, as described in chapter 6) it is almost certain that not every single sentiment is measured appropriately, but we do not know what precisely is due to erratic preprocessing (e.g. sentiment in the relationship of the judge towards suspects / criminals, or precisely in the relationship between the Prosecutor and the advocacy?). Nevertheless, the law of large numbers suggests that at a higher level of aggregation, the *pattern* of text properties that are found with text mining reflect the reality of news reports. This forms the basis for a cautious comparison of the pattern in reporting in one period with the pattern of reporting in a different period.

Incremental improvements of applied research techniques in text mining is possible. For example, the content analysis in order to identify key phrases (see chapter 2) can be improved further, but this requires new research efforts.

The text mining approaches presented in this study enables continued use in the future (based on the usual GNU license). Links to data from continuous survey research into public opinion remain necessary, because it is not guaranteed that everything that will be reported in the media in the years to come will influence the audience of readers, viewers and Internet users along the same lines of influence as in the research period 1993-2011.

It may be added that an alternative for text mining to achieve similar statements about the media coverage regarding trust in the rule of law is difficult to imagine. It is possible to envisage a different way to process a quarter of a million newspaper reports, transcripts of over five hundred television programs, a number of websites and ¾ million tweets, for example with random samples of news items that could be coded by a team of coders. But even in this case, it would remain very difficult to arrive at systematic, detailed and differentiated maps of the news, that would make possible the news effects study presented in chapter 7.

5: Media and news items to be included in text mining studies

In this study, newspapers, television news, websites and tweets are included. Digital availability is a prerequisite for the inclusion in text mining studies. This excluded, for example, political websites and tweets as sources of political information in 1993-1994 and even in 2005-2006. Newspapers could be included from the moment that they became digitally available. This explains the choice of the twenty years research period 1993-2012.

If the sciences behind the many text mining techniques continue to develop newer and more advanced approaches for text mining, then new research techniques for text mining will be developed in the foreseeable future that enable a reliable analysis of news items that pose problems for the current text mining techniques (e.g., short Twitter messages, messages that require advanced anaphora resolution, messages with humor, posts in which a lot of prior knowledge is assumed).

Digital availability does not imply digital availability for others. Messages that are digitally locked down (SMS, email, MSN, private Facebook, WhatsApp, etc.) must be disregarded in public text mining studies. For future text mining studies, gradually more media sources will become available. Old newspaper volumes are digitized. More and better applications are developed to scrape messages from a greater variety of social media. Applications that analyze responses to messages in conjunction with the messages themselves have already been developed. The development of applications for automated face recognition in motion pictures is still in its infancy. It will still take years before images and texts can be combined easily.

6: Is trust in the rule of law affected by the news in one's media?

It is far from obvious that media images of developments with respect to the rule of law affect cognitions and attitudes of audience members. A news effects study was performed to test whether the news that was described here did actually have an impact on its audience. The news effects study shows that the news media coverage of issues related to rule of law that were covered in the media of one's personal choice did indeed affect one's personal trust in the rule of law in the next quarter. It was known already that trust is strongly correlated with socio-demographic background characteristics such as education and wealth. It was known also that trust is often a manifestation of generalized trust in any person or institution whatsoever. This study shows that focus on the person and the institution of the court, and on the judicial institutions in general, and especially a focus on the content of legal processes, has a positive effect on trust in the rule of law. These research results show the viability of the new policy of the judiciary as proclaimed in a new Press Directive (2008) to strive for greater openness and transparency.

The study shows that the rule of law is decreased by crime news, discussions on civic rights, the rule of law, news about the release or escape of prisoners, and by news about the PVV (the anti-immigrant party).

Based on the distinction of Van den Berg, Hoekendijk and Niemeijer (2012) between policy options to increase trust in the rule of law at the relational level, the organizational level, and the substantive level, it can be said that the media apparently increase trust at the relational level (bridging, management of expectations) and at the organizational level (information, accountability, transparency, professionalism) with their news provision about the institutions of law about the judicial process. In contrast, the large amount of news about crime, escapes of psychiatric criminals and the release from prison of dangerous criminals (e.g. Willem Holleeder) on the substantive level may give rise to thoughts that the law fails (no austerity, no rigor, no morality), which in turn undermines trust in the judiciary.

Why news about the PVV, the rule of law, fundamental rights would undermine trust in the rule of law depends of course on the specific context of the news about these subjects during the research period (cf. previous section). The preceding chapters showed that in the last part of the research period, and particularly after the final acquittal of PVV leader Geert Wilders, these topics in the news fade into the background. In the field of justice other issues come to the fore (pedophilia, cybercrime, and privacy), whereas economic issues (credit crisis, euro crisis) decreased the news value of the immigration issue as a whole. Based on the results of the news effects analysis, it is therefore not unsurprising that trust in the rule of law since the third quarter of 2011 is higher than ever before since measurements were taken (in 2009) (Dekker et al, 2012b, p.10).

All in all, the conclusion is that the media do indeed exert an influence on citizens' trust in the rule of law. The media coverage can demonstrably contribute to, or undermine, trust in the judiciary.

Conclusion

The answers to the sub-questions enable an answer to the overarching Research Question.

Text mining may indeed contribute to new insights into the formation of trust in the rule of law. Indeed, text mining can unravel the image that traditional media such as newspapers and internet-based media like Twitter present of developments on which citizens base their trust in the rule of law. The reliability and validity of text mining is sufficient at the aggregate level (coverage per period, medium, etc.), except