

Trustworthy AI: Towards the Golden Age of RE?

Matthieu Vergne

Meritis PACA
Les Algorithmes – Aristote B
2000 Route des Lucioles
06901 Sophia Antipolis Cedex
FRANCE



RE4AI Workshop – June 23rd, 2020

Introduction

Who am I?

2016 Ph.D. in Requirements Engineering (RE) in Trento, Italy

Who am I?

2016 Ph.D. in Requirements Engineering (RE) in Trento, Italy

2017 Postdoc in Artificial Intelligence (AI) in Ikoma, Japan

Who am I?

2016 Ph.D. in Requirements Engineering (RE) in Trento, Italy

2017 Postdoc in Artificial Intelligence (AI) in Ikoma, Japan

Now Software Engineer in Sophia Antipolis, France

- *1st European & French technopole (1969)*
- *Cluster of AI companies and research institutes*
- *1 of the 4 areas chosen to support national strategy for AI (3IA)*

Context

History:

2018-04 European Commission [2018] presents its vision of AI

- *“one of the most strategic technologies of the 21st century”*
- *“a solid European framework is needed”*

Context

History:

2018-04 European Commission [2018] presents its vision of AI

- *“one of the most strategic technologies of the 21st century”*
- *“a solid European framework is needed”*

2018-06 Creation of AI high-level expert group (AI HLEG)

Context

History:

2018-04 European Commission [2018] presents its vision of AI

- *“one of the most strategic technologies of the 21st century”*
- *“a solid European framework is needed”*

2018-06 Creation of AI high-level expert group (AI HLEG)

2019-04 AI HLEG [2019] publishes Ethics Guidelines for Trustworthy AI

Context

History:

2018-04 European Commission [2018] presents its vision of AI

- *“one of the most strategic technologies of the 21st century”*
- *“a solid European framework is needed”*

2018-06 Creation of AI high-level expert group (AI HLEG)

2019-04 AI HLEG [2019] publishes Ethics Guidelines for Trustworthy AI

Goals:

- Show these guidelines offer many opportunities for RE
- Show RE is a core element in achieving these guidelines

Trustworthy AI Requirements

Key requirement 1/7

R1. Human agency and oversight

Allow humans to assess or challenge the AI system, as well as governance mechanisms of humans over the system.

Key requirement 1/7

R1. Human agency and oversight

Allow humans to assess or challenge the AI system, as well as governance mechanisms of humans over the system.

RE example

Vierhauser et al. [2017] monitor and diagnose the requirements violations through a framework with monitoring capabilities (violation info, system access, heterogeneity support, etc.).

Key requirement 2/7

R2. Technical robustness and safety

General safety, accuracy, reliability and reproducibility, but also resilience of the AI system to attack and security.

Key requirement 2/7

R2. Technical robustness and safety

General safety, accuracy, reliability and reproducibility, but also resilience of the AI system to attack and security.

RE example

Provenzano and Hänninen [2017] present a method to comply to an official standard bringing best practices and safety requirements.

Key requirement 3/7

R3. Privacy and data governance

Especially privacy of acquired and generated user data, the quality and integrity of the data used in AI learning processes, and a controlled access to user data.

Key requirement 3/7

R3. Privacy and data governance

Especially privacy of acquired and generated user data, the quality and integrity of the data used in AI learning processes, and a controlled access to user data.

RE example

Bokaei Hosseini et al. [2018] mitigate misinterpretations of privacy policies through shared ontologies.

Key requirement 4/7

R4. Transparency

Support the traceability and explainability of the processes and decisions of the AI system, as well as communicate about its abilities and limitations.

Key requirement 4/7

R4. Transparency

Support the traceability and explainability of the processes and decisions of the AI system, as well as communicate about its abilities and limitations.

RE example

Hübner and Paech [2018] to have a performant traceability by showing how the components of the produced AI system relate to its requirements.

Key requirement 5/7

R5. Diversity, non-discrimination and fairness

Especially the avoidance of unfair bias, allowing all people to access and use the AI system, and the participation of stakeholders who might be directly or indirectly affected.

Key requirement 5/7

R5. Diversity, non-discrimination and fairness

Especially the avoidance of unfair bias, allowing all people to access and use the AI system, and the participation of stakeholders who might be directly or indirectly affected.

RE example

Nunes Rodrigues et al. [2018] provide a comprehensive list of personas and persona-based modelling to ensure that the relevant stakeholder profiles are explicitly considered.

Key requirement 6/7

R6. Environmental and societal well-being

Monitor the impact on society and democracy and consider the broader society, other sentient beings and the environment, as stakeholders for sustainability and environmental friendliness.

Key requirement 6/7

R6. Environmental and societal well-being

Monitor the impact on society and democracy and consider the broader society, other sentient beings and the environment, as stakeholders for sustainability and environmental friendliness.

RE example

Condori Fernandez and Lago [2018] integrate green strategies in quality requirements for optimizing energy and other resources consumption.

Key requirement 7/7

R7. Accountability

Support auditability and reporting of negative impacts (e.g. whistle-blowers and NGOs), identify trade-offs and redress actions when conflicts arise between the key requirements.

Key requirement 7/7

R7. Accountability

Support auditability and reporting of negative impacts (e.g. whistle-blowers and NGOs), identify trade-offs and redress actions when conflicts arise between the key requirements.

RE example

Hübner and Paech [2018] help auditors by ensuring a satisfying level of quality for traceability links.

Key point: RE examples are easy to find

- Various RE venues: REFSQ, EmpiRE, CAiSE, iStar, etc.

Key point: RE examples are easy to find

- Various RE venues: REFSQ, EmpiRE, CAiSE, iStar, etc.
- REFSQ 2018 alone allows to illustrate each requirement

Key point: RE examples are easy to find

- Various RE venues: REFSQ, EmpiRE, CAiSE, iStar, etc.
- REFSQ 2018 alone allows to illustrate each requirement
- Quick view of REFSQ 2020 papers:
 - *Assessments (R1), privacy (R3), transparency (R4), feedback (R5)*

Key point: RE examples are easy to find

- Various RE venues: REFSQ, EmpiRE, CAiSE, iStar, etc.
- REFSQ 2018 alone allows to illustrate each requirement
- Quick view of REFSQ 2020 papers:
 - *Assessments (R1), privacy (R3), transparency (R4), feedback (R5)*

AI HLEG guidelines deeply bond with RE

RE is a rich source of ideas to cope with the 7 key requirements, which makes AI a rich source of opportunities for RE practices.

Challenges for RE

Still a lot to do

Space for improvements for RE to support the 7 key requirements:

- Still young on the specificities of Big Data issues.
 - *Arruda and Madhavji [2018]*

Still a lot to do

Space for improvements for RE to support the 7 key requirements:

- Still young on the specificities of Big Data issues.
 - *Arruda and Madhavji [2018]*
- Open issue of ambiguous/incomplete requirements.
 - *Dalpiaz et al. [2018]*

Still a lot to do

Space for improvements for RE to support the 7 key requirements:

- Still young on the specificities of Big Data issues.
 - *Arruda and Madhavji [2018]*
- Open issue of ambiguous/incomplete requirements.
 - *Dalpiaz et al. [2018]*
- Analysts need more support to recover traces.
 - *Chitre et al. [2018]*

Still a lot to do

Space for improvements for RE to support the 7 key requirements:

- Still young on the specificities of Big Data issues.
 - *Arruda and Madhavji [2018]*
- Open issue of ambiguous/incomplete requirements.
 - *Dalpiaz et al. [2018]*
- Analysts need more support to recover traces.
 - *Chitre et al. [2018]*
- Tools to help in RE tasks still far from reliable.
 - *Hübner and Paech [2018]*

Still a lot to do

Space for improvements for RE to support the 7 key requirements:

- Still young on the specificities of Big Data issues.
 - *Arruda and Madhavji [2018]*
- Open issue of ambiguous/incomplete requirements.
 - *Dalpiaz et al. [2018]*
- Analysts need more support to recover traces.
 - *Chitre et al. [2018]*
- Tools to help in RE tasks still far from reliable.
 - *Hübner and Paech [2018]*
- Requirements in complex organizations is by itself challenging.
 - *Alsaqaf et al. [2018], Wohlrab et al. [2018]*

Even more important: where is RE?

AI HLEG guidelines document contains:

- 41 pages
- 59 occurrences of “requirements”, without counting pictures
- 0 occurrences of “requirements engineering”

Even more important: where is RE?

AI HLEG guidelines document contains:

- 41 pages
- 59 occurrences of “requirements”, without counting pictures
- 0 occurrences of “requirements engineering”

Personal experience: there is as much people knowing about AI than people ignoring about RE.

Even more important: where is RE?

AI HLEG guidelines document contains:

- 41 pages
- 59 occurrences of “requirements”, without counting pictures
- 0 occurrences of “requirements engineering”

Personal experience: there is as much people knowing about AI than people ignoring about RE.

Communicate!

Not only with European actors, but also with AI people.

Conclusion

Summary

- AI is a currently hot topic all over the world

Summary

- AI is a currently hot topic all over the world
- European Commission considers AI as key for the future

Summary

- AI is a currently hot topic all over the world
- European Commission considers AI as key for the future
- Guidelines for trustworthy AI systems with 7 key requirements

Summary

- AI is a currently hot topic all over the world
- European Commission considers AI as key for the future
- Guidelines for trustworthy AI systems with 7 key requirements
- Numerous resources to exploit from RE

Summary

- AI is a currently hot topic all over the world
- European Commission considers AI as key for the future
- Guidelines for trustworthy AI systems with 7 key requirements
- Numerous resources to exploit from RE
- Yet challenges and open questions remain

Summary

- AI is a currently hot topic all over the world
- European Commission considers AI as key for the future
- Guidelines for trustworthy AI systems with 7 key requirements
- Numerous resources to exploit from RE
- Yet challenges and open questions remain

The RE community is strong and dynamic enough to tackle the challenges. It should be ambitious enough to ride the AI boat and take its opportunities.


Thanks for your attention.

Remarks?
Questions?




References

- AI HLEG. Ethics Guidelines for Trustworthy AI. Report, European Commission, Apr. 2019. URL <https://ec.europa.eu/futurium/en/ai-alliance-consultation>.
- W. Alsaqaf, M. Daneva, and R. Wieringa. Quality Requirements Challenges in the Context of Large-Scale Distributed Agile: An Empirical Study. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 139–154. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_9. URL http://link.springer.com/10.1007/978-3-319-77243-1_9.
- D. Arruda and N. H. Madhavji. State of Requirements Engineering Research in the Context of Big Data Applications. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 307–323. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_20. URL http://link.springer.com/10.1007/978-3-319-77243-1_20.
- J. Bhatia, M. C. Evans, and T. D. Breaux. Identifying incompleteness in privacy policy goals using semantic frames. Requirements Engineering, 24(3):291–313, Sept. 2019. ISSN 0947-3602, 1432-010X. doi: 10.1007/s00766-019-00315-y. URL <http://link.springer.com/10.1007/s00766-019-00315-y>.
- M. Bokaei Hosseini, T. D. Breaux, and J. Niu. Inferring Ontology Fragments from Semantic Role Typing of Lexical Variants. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 39–56. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_10. URL http://link.springer.com/10.1007/978-3-319-77243-1_10.

- 10.1007/978-3-319-77243-1_3. URL
http://link.springer.com/10.1007/978-3-319-77243-1_3.
- B. Chitre, J. H. Hayes, and A. Dekhtyar. Second-Guessing in Tracing Tasks Considered Harmful? In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 92–98. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_6. URL http://link.springer.com/10.1007/978-3-319-77243-1_6.
- N. Condori Fernandez and P. Lago. The Influence of Green Strategies Design onto Quality Requirements Prioritization. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 189–205. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_12. URL http://link.springer.com/10.1007/978-3-319-77243-1_12.
- F. Dalpiaz, I. van der Schalk, and G. Lucassen. Pinpointing Ambiguity and Incompleteness in Requirements Engineering via Information Visualization and NLP. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 119–135. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_8. URL http://link.springer.com/10.1007/978-3-319-77243-1_8.
- European Commission. Artificial Intelligence for Europe. Communication from the Commission, European Commission, Apr. 2018. URL 

[https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN.COM\(2018\)237final](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN.COM(2018)237final).

- A.-M. Ghiran, R. A. Buchmann, and C.-C. Osman. Security Requirements Elicitation from Engineering Governance, Risk Management and Compliance. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 283–289. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_17. URL http://link.springer.com/10.1007/978-3-319-77243-1_17.
- M. Hosseini, A. Shahri, K. Phalp, and R. Ali. Four reference models for transparency requirements in information systems. Requirements Engineering, 23(2):251–275, June 2018. ISSN 0947-3602, 1432-010X. doi: 10.1007/s00766-017-0265-y. URL <http://link.springer.com/10.1007/s00766-017-0265-y>.
- S. Hotomski, E. Ben Charrada, and M. Glinz. Keeping Evolving Requirements and Acceptance Tests Aligned with Automatically Generated Guidance. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 247–264. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_15. URL http://link.springer.com/10.1007/978-3-319-77243-1_15.
- P. Hübner and B. Paech. Evaluation of Techniques to Detect Wrong Interaction Based Trace Links. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 75–91. Springer International Publishing, Cham, 2018. ISBN 

- 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_5. URL http://link.springer.com/10.1007/978-3-319-77243-1_5.
- T. Li, F. Zhang, and D. Wang. Automatic User Preferences Elicitation: A Data-Driven Approach. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 324–331. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_21. URL http://link.springer.com/10.1007/978-3-319-77243-1_21.
- G. Nunes Rodrigues, C. Joel Tavares, N. Watanabe, C. Alves, and R. Ali. A Persona-Based Modelling for Contextual Requirements. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 352–368. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_23. URL http://link.springer.com/10.1007/978-3-319-77243-1_23.
- T. Olsson and K. Wnuk. QREME – Quality Requirements Management Model for Supporting Decision-Making. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 173–188. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_11. URL http://link.springer.com/10.1007/978-3-319-77243-1_11.
- C. Palomares, X. Franch, and D. Fucci. Personal Recommendations in Requirements Engineering: The OpenReq Approach. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 297–304. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_21. URL http://link.springer.com/10.1007/978-3-319-77243-1_21.

- Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_19. URL http://link.springer.com/10.1007/978-3-319-77243-1_19.
- L. Provenzano and K. Hänninen. Specifying Software Requirements for Safety-Critical Railway Systems: An Experience Report. In P. Grünbacher and A. Perini, editors, REFSQ, volume 10153, pages 363–369. Springer International Publishing, Cham, 2017. ISBN 978-3-319-54044-3 978-3-319-54045-0. doi: 10.1007/978-3-319-54045-0_25. URL http://link.springer.com/10.1007/978-3-319-54045-0_25.
- M. Vierhauser, R. Rabiser, and J. Cleland-Huang. From Requirements Monitoring to Diagnosis Support in System of Systems. In P. Grünbacher and A. Perini, editors, REFSQ, volume 10153, pages 181–187. Springer International Publishing, Cham, 2017. ISBN 978-3-319-54044-3 978-3-319-54045-0. doi: 10.1007/978-3-319-54045-0_14. URL http://link.springer.com/10.1007/978-3-319-54045-0_14.
- J. P. Winkler and A. Vogelsang. Using Tools to Assist Identification of Non-requirements in Requirements Specifications – A Controlled Experiment. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, REFSQ, volume 10753, pages 57–71. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_4. URL http://link.springer.com/10.1007/978-3-319-77243-1_4.

- R. Wohlrab, P. Pelliccione, E. Knauss, and S. C. Gregory. The Problem of Consolidating RE Practices at Scale: An Ethnographic Study. In E. Kamsties, J. Horkoff, and F. Dalpiaz, editors, *REFSQ*, volume 10753, pages 155–170. Springer International Publishing, Cham, 2018. ISBN 978-3-319-77242-4 978-3-319-77243-1. doi: 10.1007/978-3-319-77243-1_10. URL http://link.springer.com/10.1007/978-3-319-77243-1_10.