

Tech talk at **World Summit AI 2019**
9th October 2019, Zaandam, Netherlands



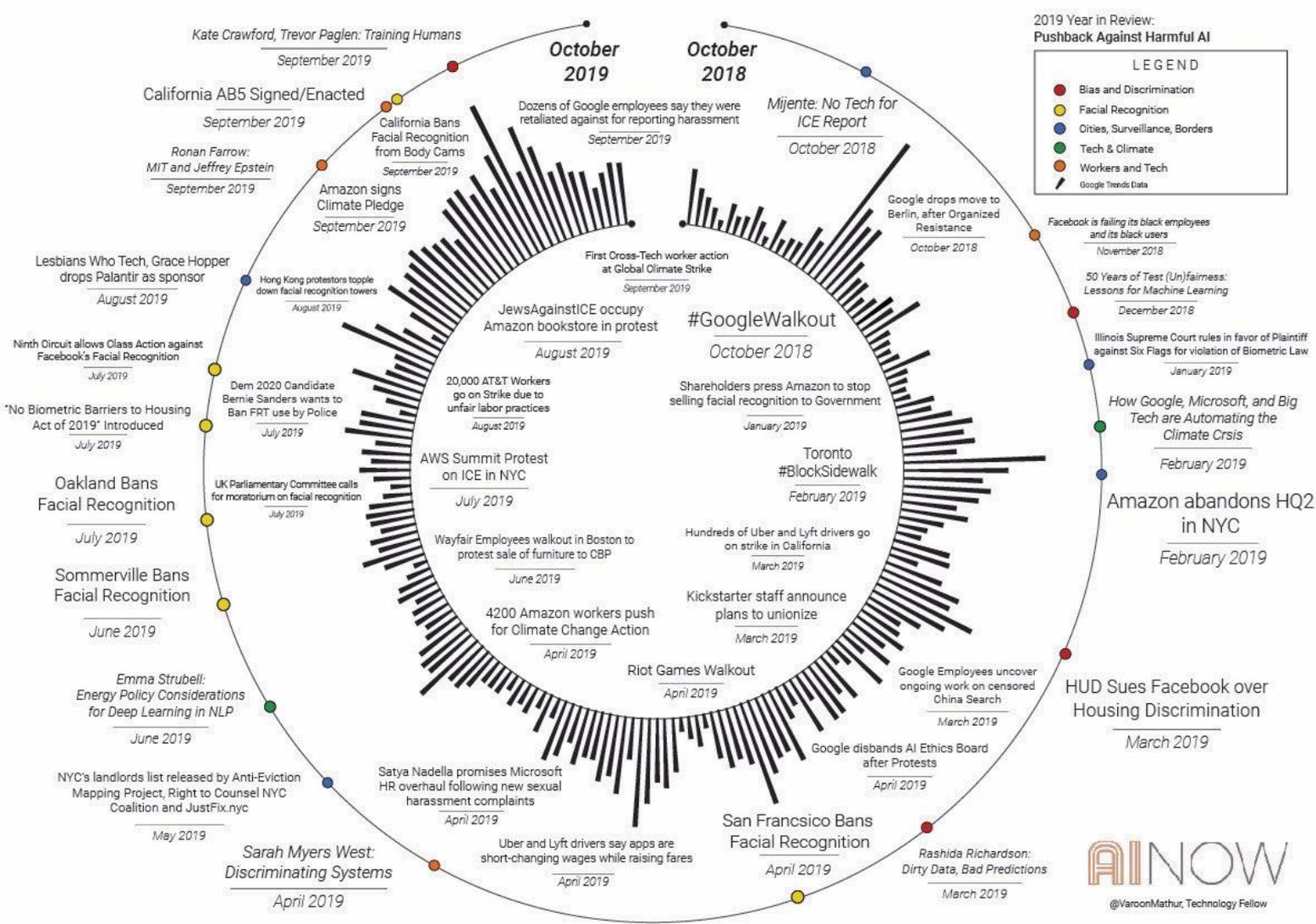
Hochschule für
Wirtschaft und Recht Berlin
Berlin School of Economics and Law



What makes us human makes us also AI (ir)Responsible: When and how to inject your AI with ?

Prof. Dr. Dagmar Monett
Berlin School of Economics and Law, and AGISI.org

@dmonett



Pushback Against Harmful AI

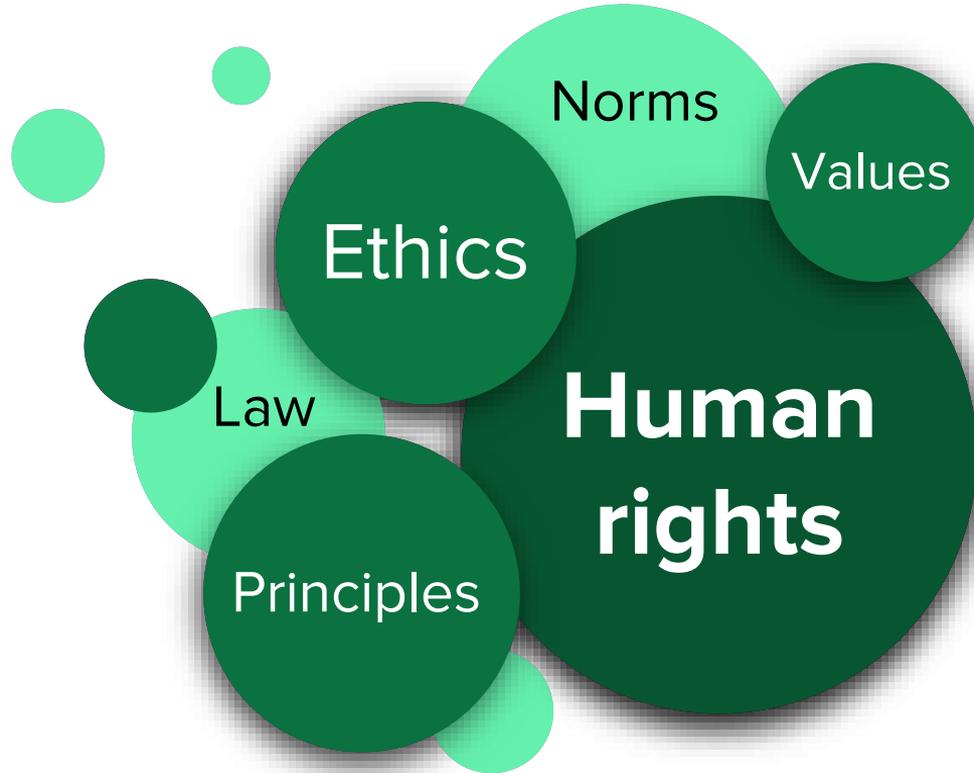
Courtesy of
 @VaroonMathur,
 Tech fellow
 @AINowInstitute



“Ethical judgment corresponds with an ethical system.”

If the system is not ethical, we cannot have ethical outcomes.

It's actually much more than only ethics...



**The Universal Declaration of
Human Rights, standards,
guidelines, policies,
recommendations.**

To name a few.



THE UNIVERSAL DECLARATION OF Human Rights

PREAMBLE Recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world.

Article 1 All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

Article 2 Everyone has the right to life, liberty and security of person. No one shall be subjected to arbitrary arrest, detention or exile. No one shall be deprived of his life, liberty, property, family, health, honor and reputation. No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. No one shall be subjected to arbitrary arrest, detention or exile.

Article 3 Everyone has the right to life, liberty and security of person. No one shall be subjected to arbitrary arrest, detention or exile.

Article 4 No one shall be held in slavery or servitude; slavery and the slave trade shall be prohibited in all their forms.

Article 5 No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment.

Article 6 Everyone has the right to recognition as a person before the law.

Article 7 All are equal before the law and are entitled without distinction to equal protection of the law. All are equal before the law and are entitled without distinction to equal protection of the law. All are equal before the law and are entitled without distinction to equal protection of the law.

Article 8 Everyone has the right to an effective remedy by the competent national tribunals for the redress of any wrong done to him.

Article 9 No one shall be subjected to arbitrary arrest, detention or exile.

Article 10 Everyone has the right to a fair and public hearing by an independent and impartial tribunal in the determination of his rights and obligations and of any criminal charge against him.

Article 11 Everyone charged with a criminal offence has the right to be presumed innocent until proved guilty according to the law. In the course of a criminal trial, he shall be entitled to a full and fair hearing in a public hearing. He shall be entitled to a full and fair hearing in a public hearing.

Article 12 No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. No one shall be subjected to arbitrary arrest, detention or exile.

Article 13 Everyone has the right to free movement and residence within the borders of each state. No one shall be arbitrarily deprived of his nationality nor of the right to change his nationality.

Article 14 Everyone has the right to seek and to enjoy in other countries asylum from persecution. No one shall be subjected to arbitrary arrest, detention or exile.

Article 15 Everyone has the right to a nationality. No one shall be arbitrarily deprived of his nationality nor of the right to change his nationality.

Article 16 Everyone has the right to marry and to found a family, which is the basis of the peace and stability of the world. No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. No one shall be subjected to arbitrary arrest, detention or exile.

Article 17 Everyone has the right to own property alone as well as in association with others. No one shall be arbitrarily deprived of his property. No one shall be arbitrarily deprived of his property.

Article 18 Everyone has the right to freedom of thought, conscience and religion. No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. No one shall be subjected to arbitrary arrest, detention or exile.

Article 19 Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without restriction and to receive and impart information and ideas without restriction as to sources or frontiers.

Article 20 Everyone has the right to peaceful assembly and to associate with other persons, to form and to join trade unions, for the protection of his interests.

Article 21 Everyone has the right to take part in the government of his country, directly or through freely chosen representatives. Everyone has the right of equal access to public service in his country.

Article 22 Everyone has the right to social security. No one shall be subjected to arbitrary arrest, detention or exile.

Article 23 Everyone has the right to work, to free choice of employment, to just and favorable conditions of work, to protection against unemployment, to higher pay for more complex tasks, to a reasonable leisure time, to a fair wage and to a just and favorable system of remuneration. No one shall be subjected to arbitrary arrest, detention or exile.

Article 24 Everyone has the right to rest and leisure, including reasonable limitation of working hours and periodic holidays with pay, and to a fair wage and to a just and favorable system of remuneration.

Article 25 Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing, medical care and social services, and to education. No one shall be subjected to arbitrary arrest, detention or exile.

Article 26 Education shall be made available to all. It shall be free, at least in the elementary and basic stages. Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. It shall promote understanding, tolerance and friendship among all nations, racial and religious groups, and shall foster the activities of the United Nations and its Members.

Article 27 Everyone has the right to own property alone as well as in association with others. No one shall be arbitrarily deprived of his property. No one shall be arbitrarily deprived of his property.

Article 28 Everyone has the right to a social and international order in which the rights and freedoms set forth in this Declaration can be fully realized. No one shall be subjected to arbitrary arrest, detention or exile.

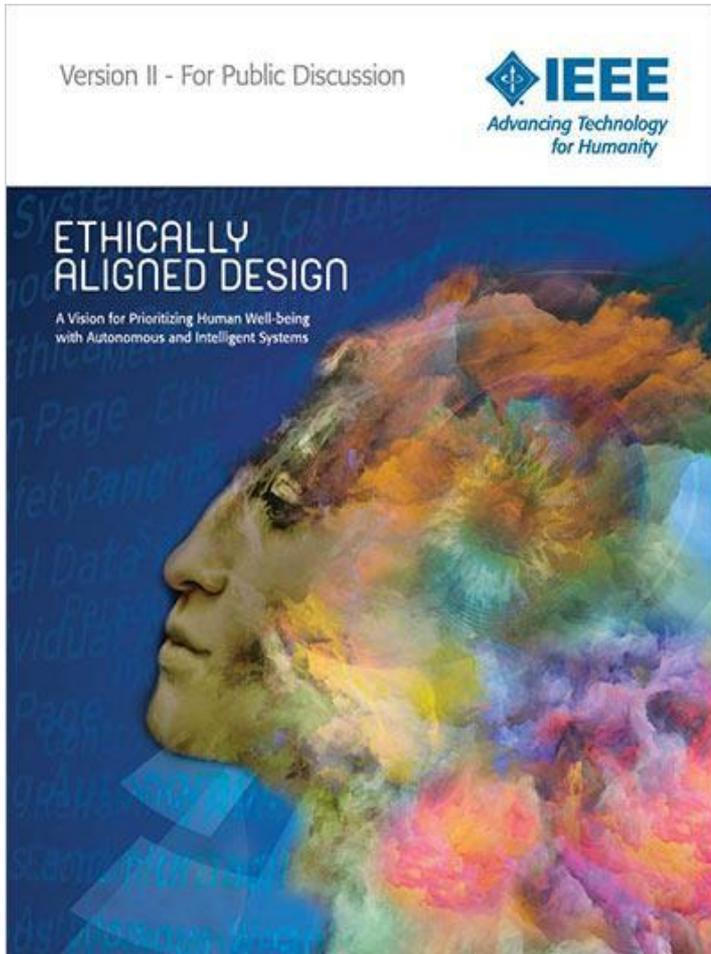
Article 29 Everyone has the right to the peaceful enjoyment of his property. No one shall be arbitrarily deprived of his property. No one shall be arbitrarily deprived of his property.

Article 30 Nothing shall be interpreted as authorizing any State, group or individual to engage in any activity aimed at the destruction of any of the rights and freedoms set forth in this Declaration.

“Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.”

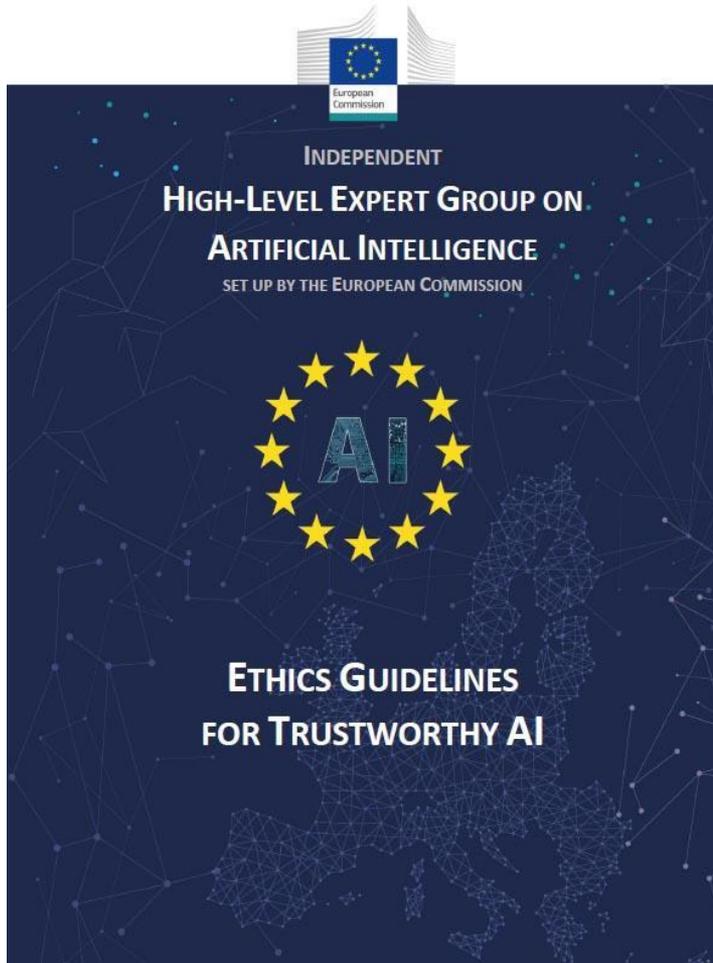
<https://www.un.org/en/universal-declaration-human-rights/>





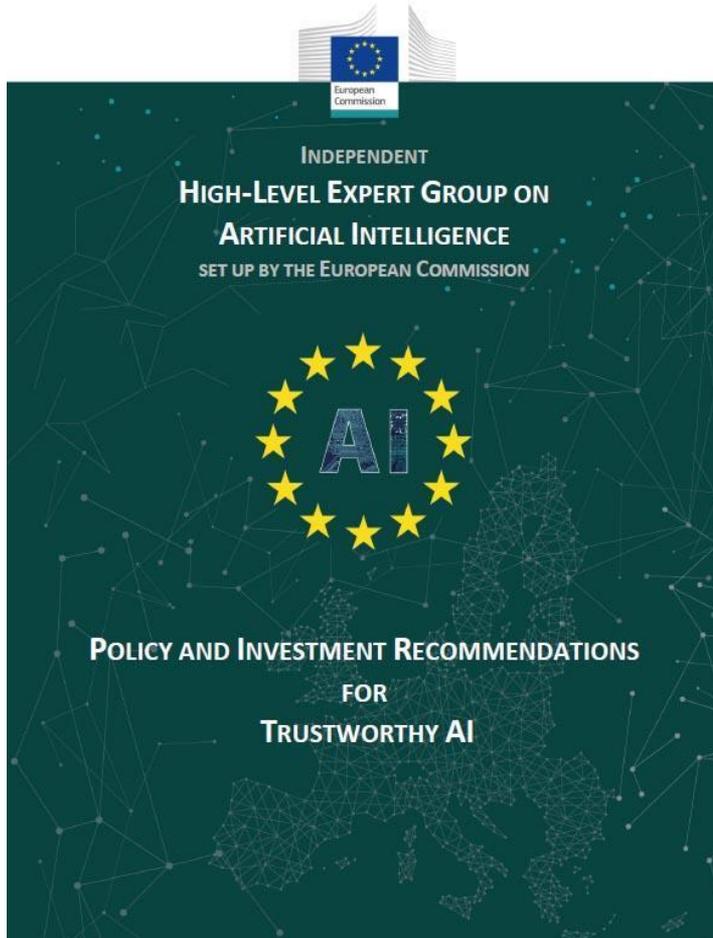
(2017)

*“[IEEE’s] commitment to engage global communities in creating practical guidelines that **will help ensure the development of ethically aligned autonomous and intelligent products, services, and systems.**”*



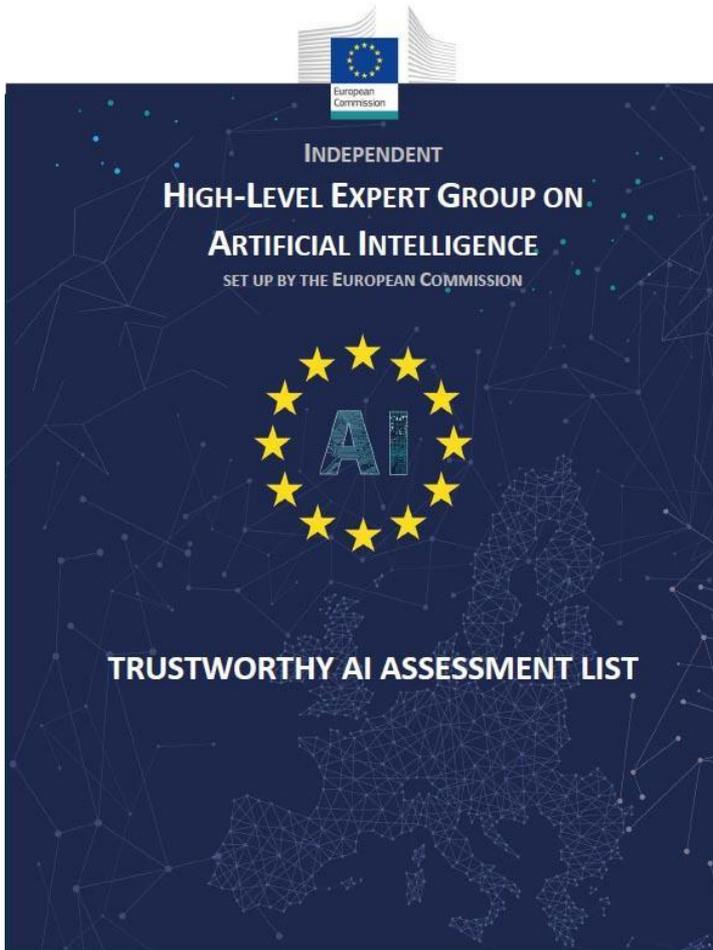
(2019)

*“Ensure that the development, deployment and use of AI systems meets the **seven key requirements for Trustworthy AI**: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being and (7) accountability.”*



(2019)

*“[R]ecommendations that can guide Trustworthy AI towards sustainability, growth and competitiveness, as well as inclusion – while **empowering, benefiting and protecting human beings.**”*



(2019)

“[A] non-exhaustive Trustworthy AI assessment list [with 147 questions] to operationalise Trustworthy AI. It particularly applies to AI systems that directly interact with users, and is primarily addressed to developers and deployers of AI systems.”

**What makes us human
makes us also
AI (ir)Responsible**

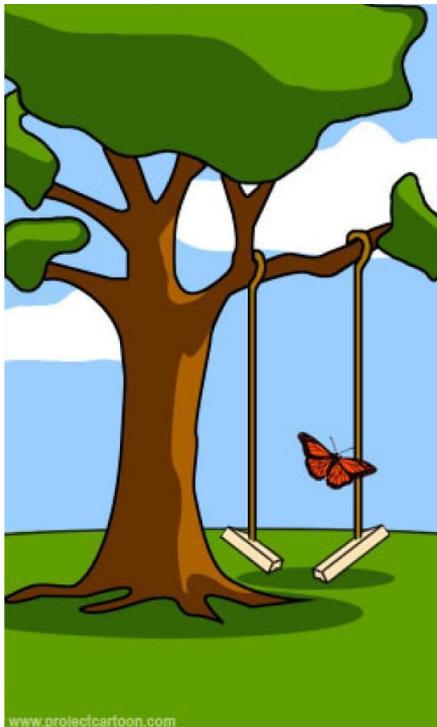


What the company says about its AI product

How robust the product is

“Did you assess how your system behaves in unexpected situations and environments?”

(Trustworthy AI assessment list, 2019)



The first security tests

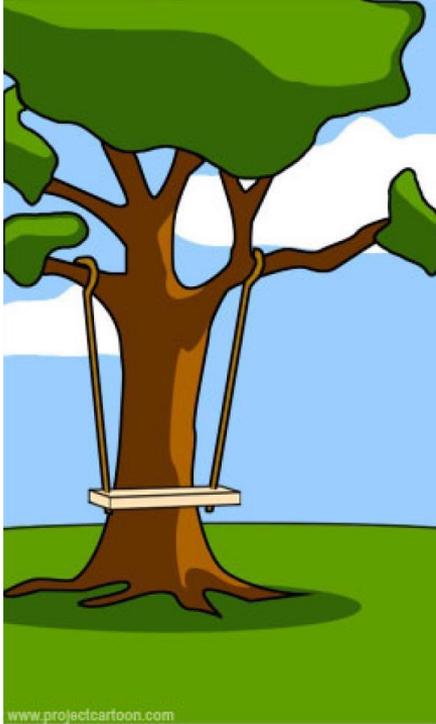


After the first security tests

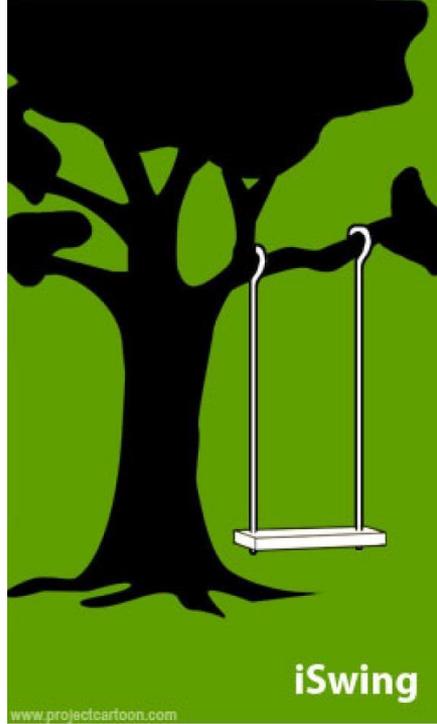
“Did you consider the potential impact or safety risk to the environment or to animals?”

(Trustworthy AI assessment list, 2019)

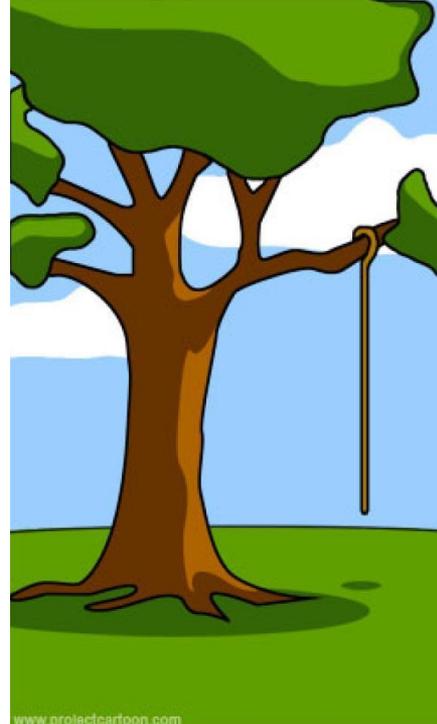
And what about “AI”?



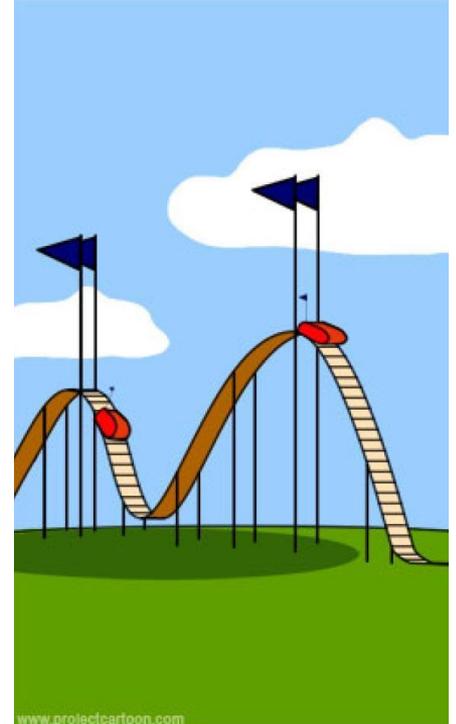
www.projectcartoon.com
Explaining the AI's decision-making process



www.projectcartoon.com
iSwing
What marketing says about how much AI there is inside



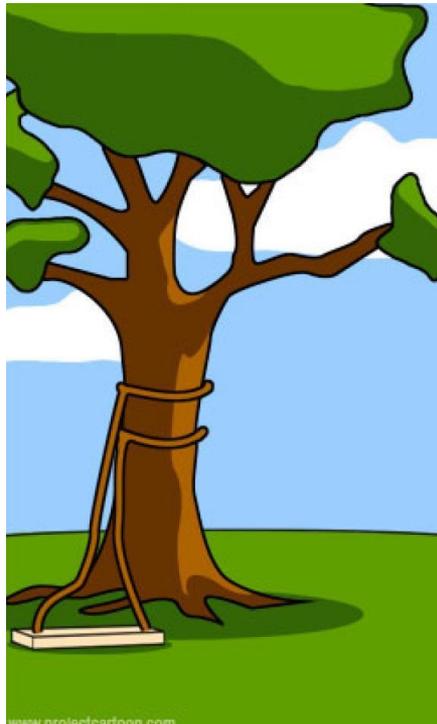
www.projectcartoon.com
How much AI there is inside



www.projectcartoon.com
The company's reaction in a live interview



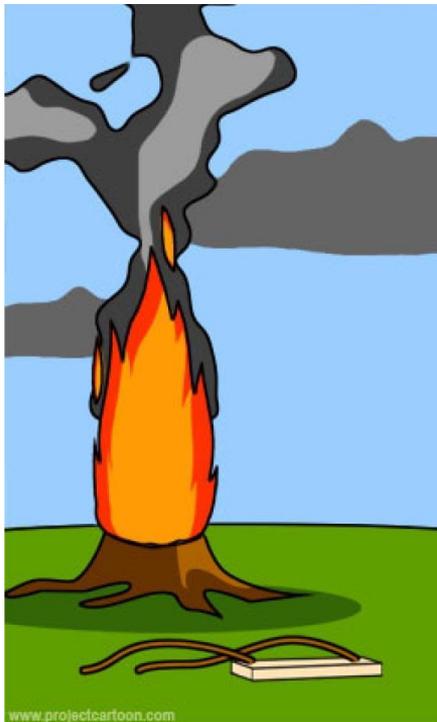
What the end user truly needs



How human-centered the product is

“In case the AI system features a chat bot or conversational system, are the human end users made aware of the fact that they are interacting with a non-human agent?”

(Trustworthy AI assessment list, 2019)



How the user data was processed



The product and the bias

“Did you consider ways to develop the AI system or train the model without or with minimal use of potentially sensitive or personal data?”

(Trustworthy AI assessment list, 2019)

Be aware of human biases

D. Monett, L. Hoge, and C. W. P. Lewis. **Cognitive Biases Undermine Consensus on Definitions of Intelligence and Limit Understanding**. In U. Furbach, S. Hölldobler, M. Ragni, R. Rzepka, C. Schon, J. Vallverdu, and A. Włodarczyk (eds.), Joint Proceedings of the Workshops on Linguistic and Cognitive Approaches to Dialog Agents (LaCATODA 2019) and on Bridging the Gap Between Human and Automated Reasoning (BtG 2019), Vol. 2452, pp. 52-59, CEUR-WS, co-located with IJCAI 2019, 10 - 16 August, 2019, Macao, China.

Position of
definitions

Anchoring Effect (Tversky & Kahneman, 1974)

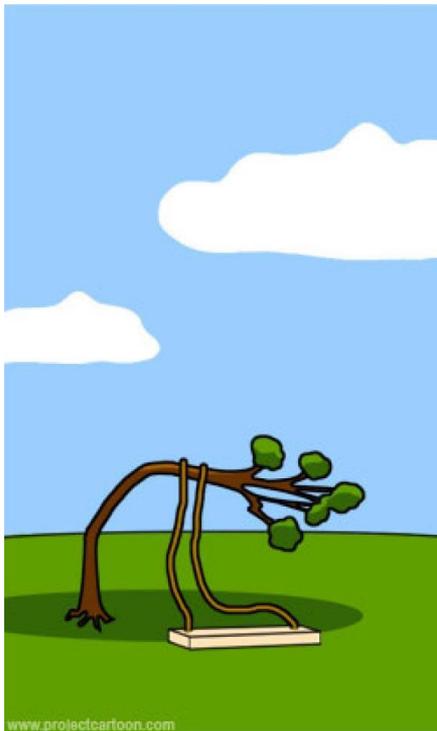
Comments when
argumenting

Confirmation bias (Nickerson, 1998)

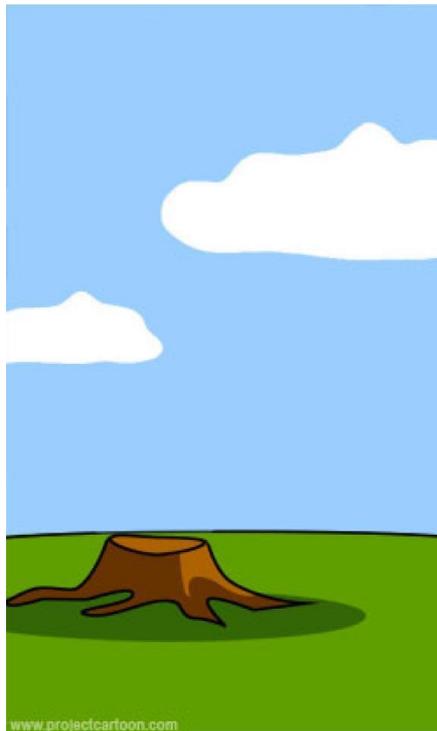
Expressing
disagreement

Focalism (Kahneman et al., 2006)

... among other biases.



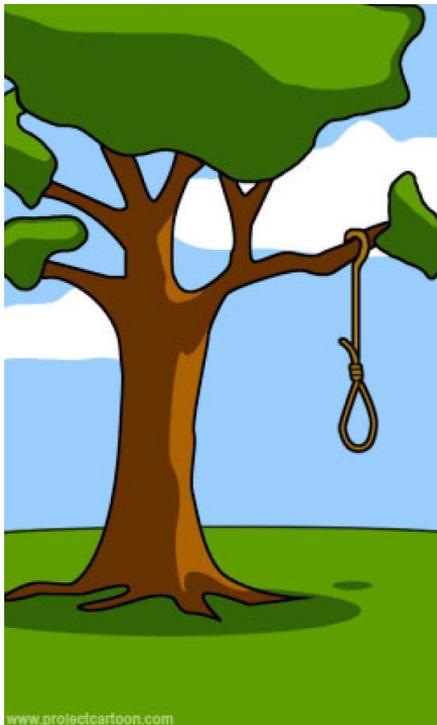
How fairness was considered



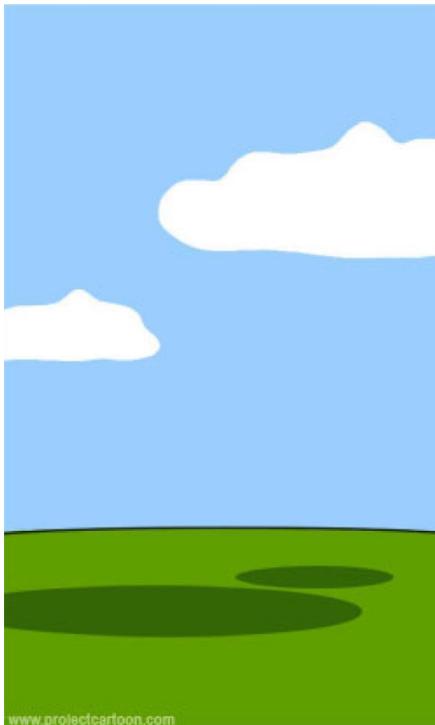
How ethical aspects were considered

“Did you consider diversity and representativeness of users in the data? Did you test for specific populations or problematic use cases?”

(Trustworthy AI assessment list, 2019)



www.projectcartoon.com
Seeking for accountability



www.projectcartoon.com
What the company thinks
about regulation

“Did you put in place mechanisms that facilitate the system’s auditability by internal and/or independent actors?”

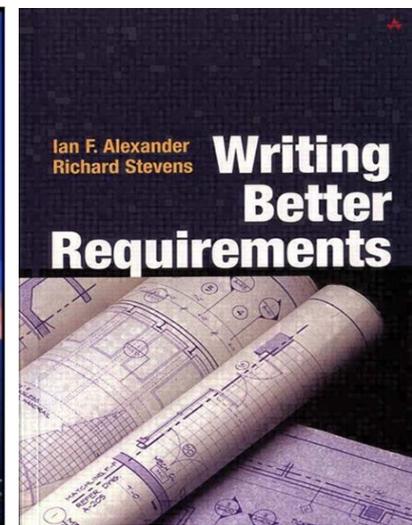
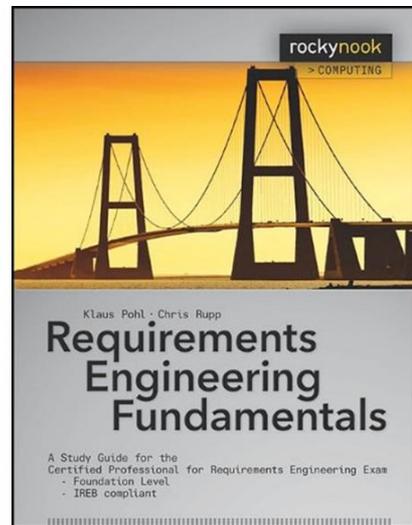
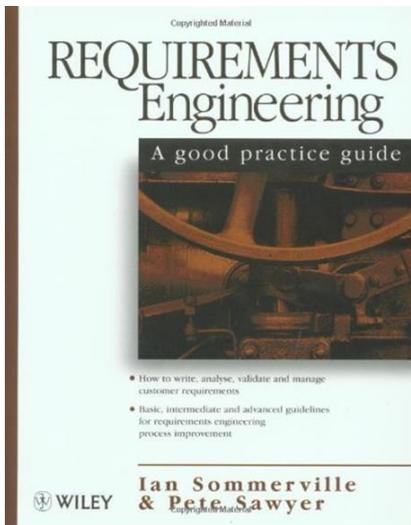
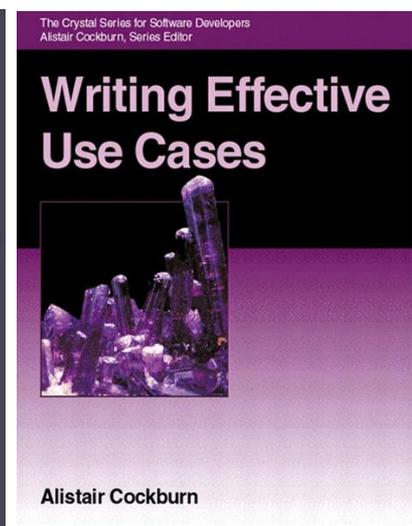
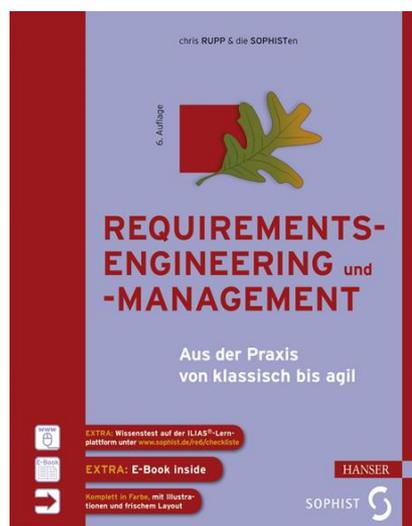
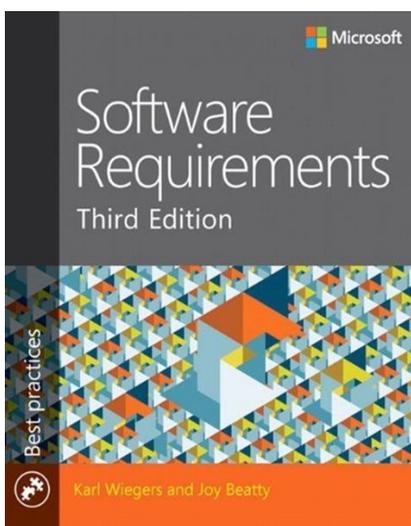
(Trustworthy AI assessment list, 2019)

**Responsible AI needs
responsible humans.**

Be  Responsible.

When and how to inject your AI with ?

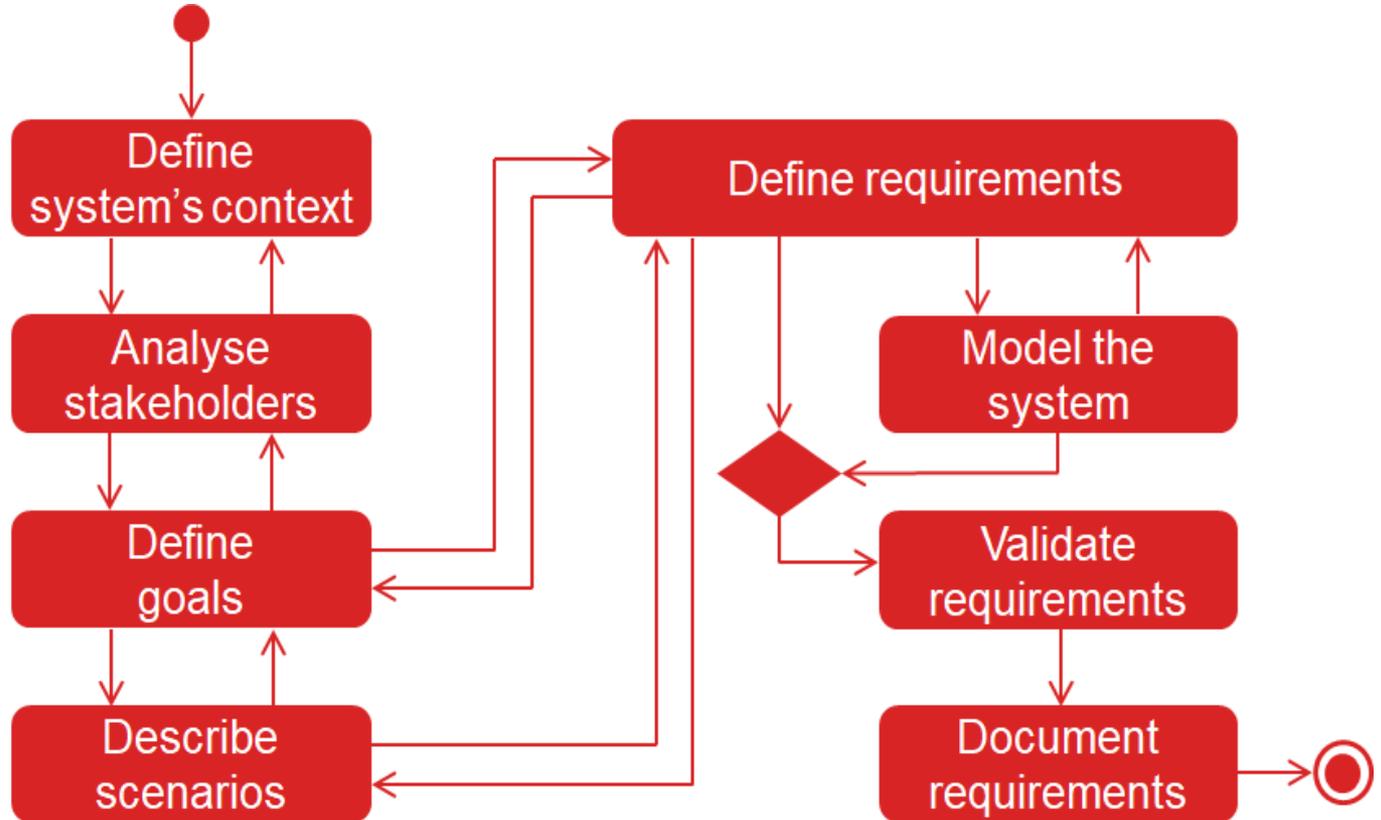
It has been at our fingertips
for a long time...



Among many others!

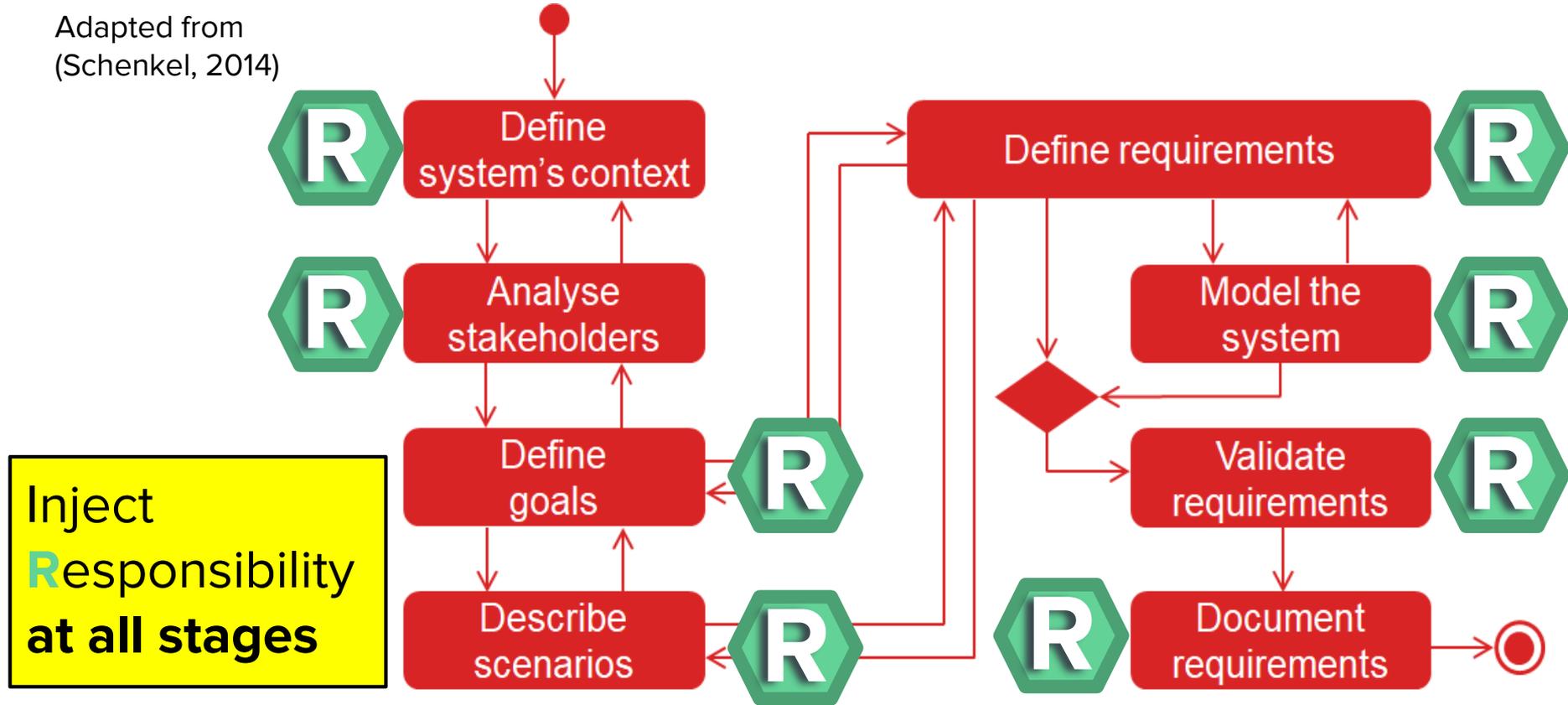
Requirements engineering as a process

Adapted from
(Schenkel, 2014)



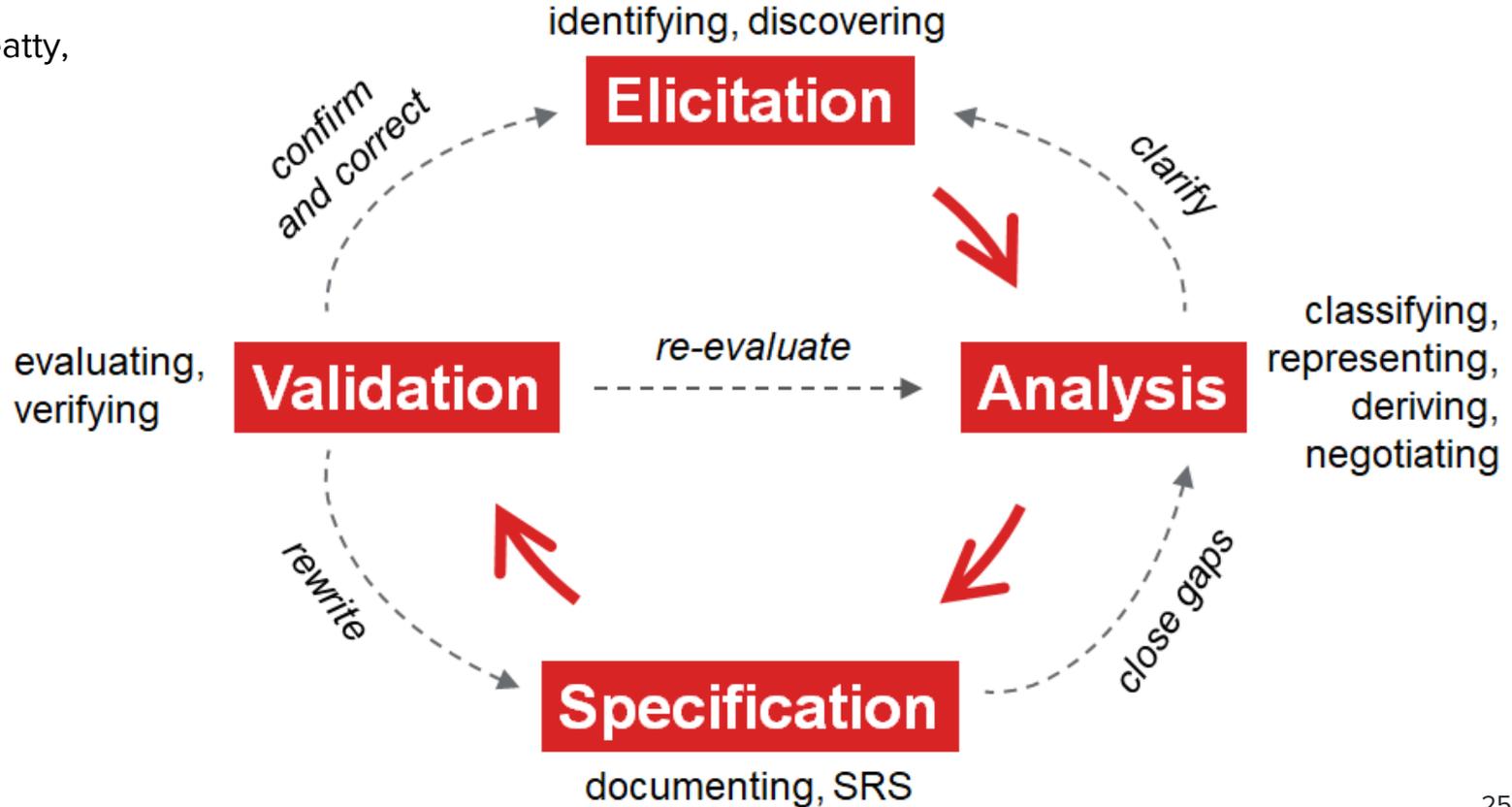
Requirements engineering as a process

Adapted from
(Schenkel, 2014)



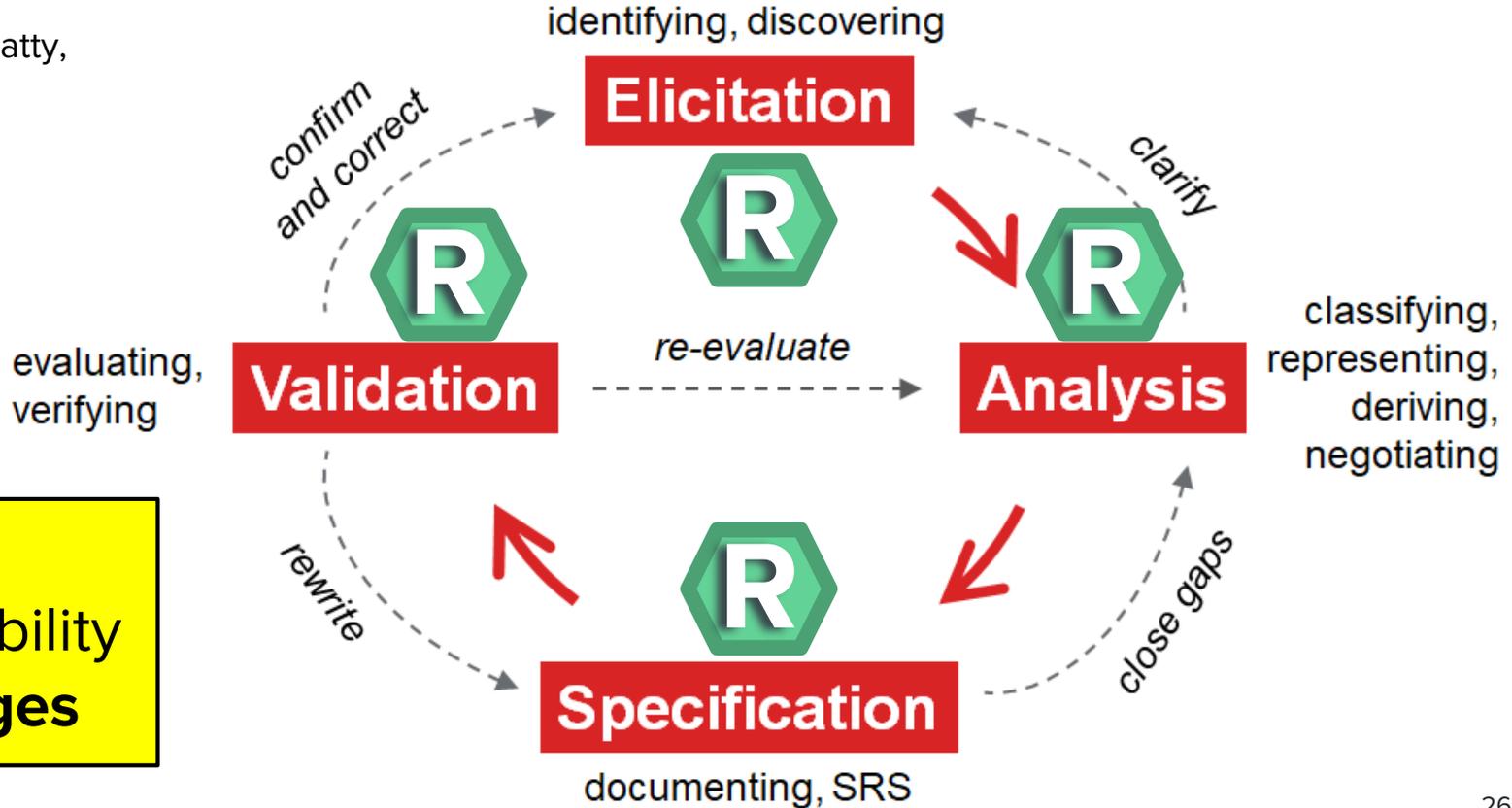
Requirements development in detail

Adapted from
(Wiegiers & Beatty,
2013)



Requirements development in detail

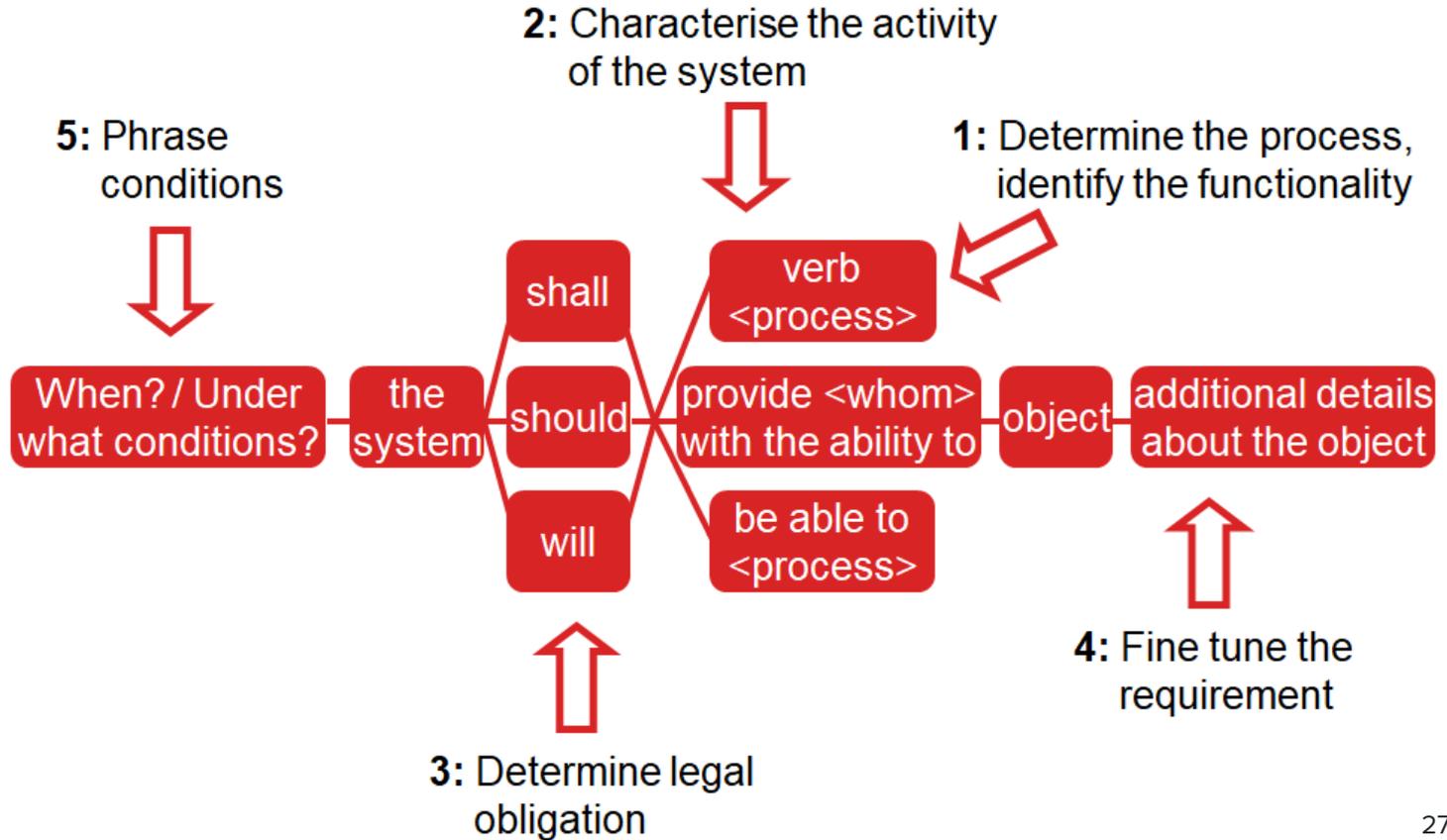
Adapted from
(Wiegiers & Beatty,
2013)



Inject
Responsibility
at all stages

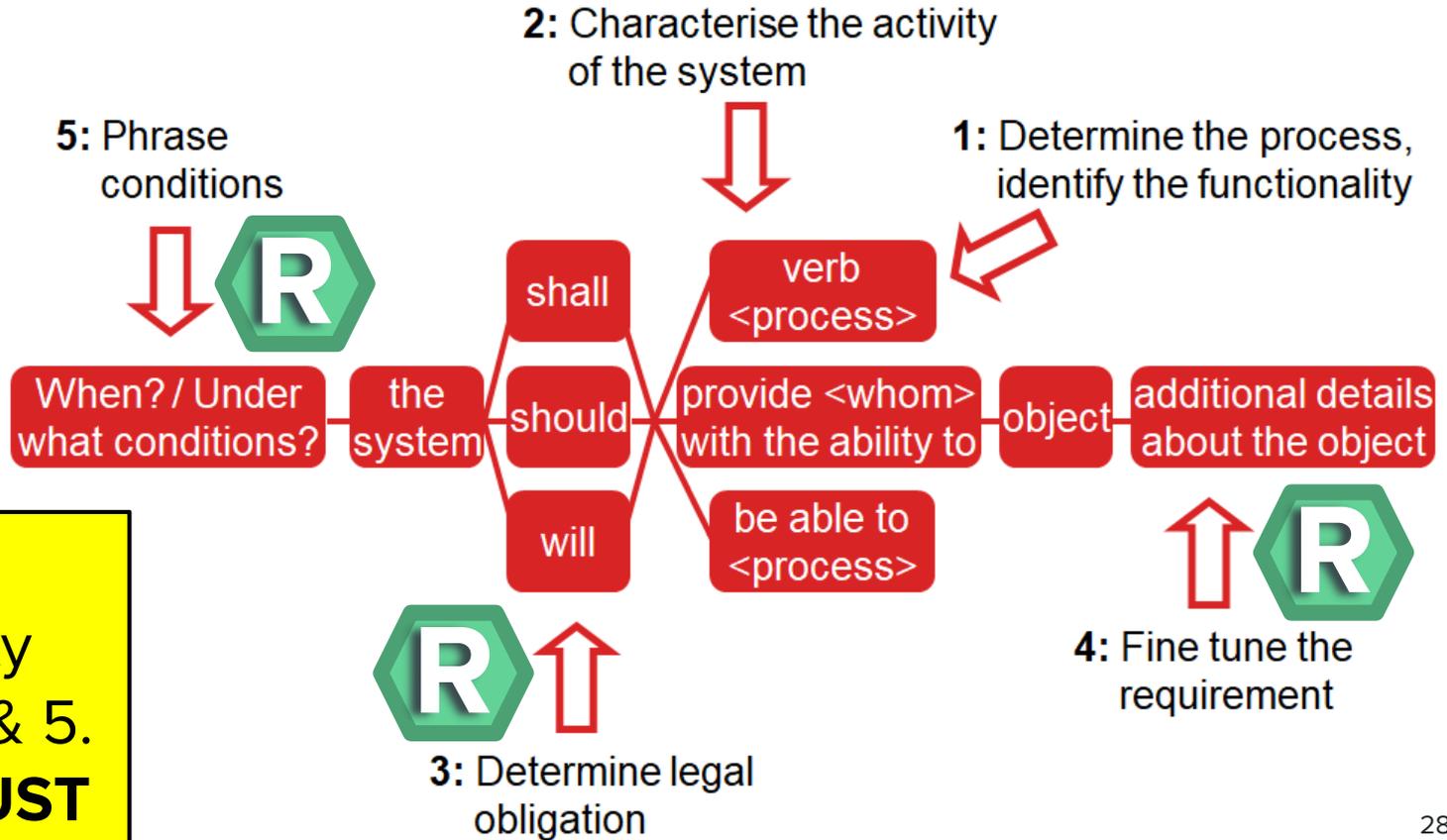
High quality requirements

Adapted from
(Rupp et al., 2014)



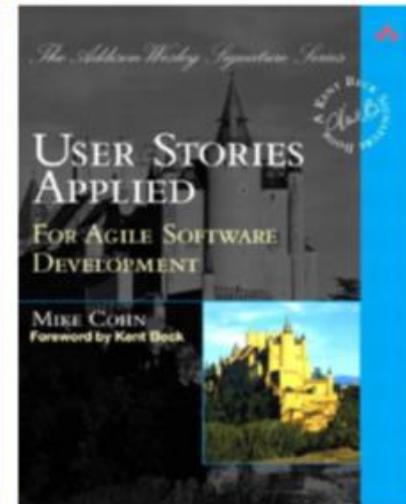
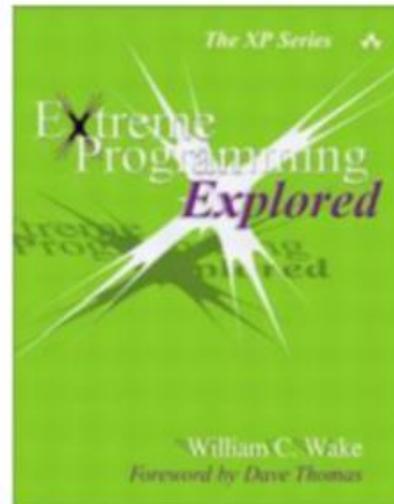
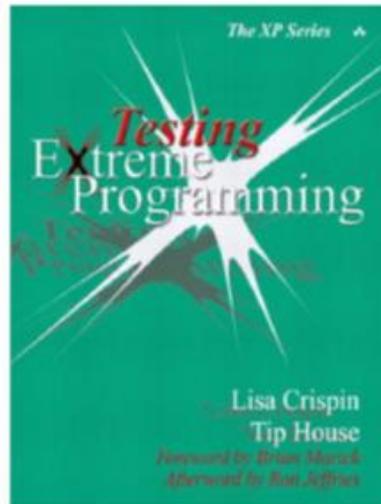
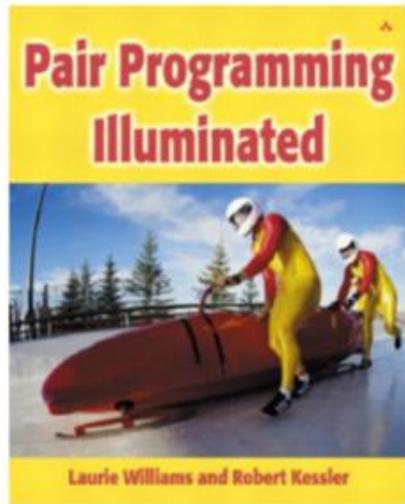
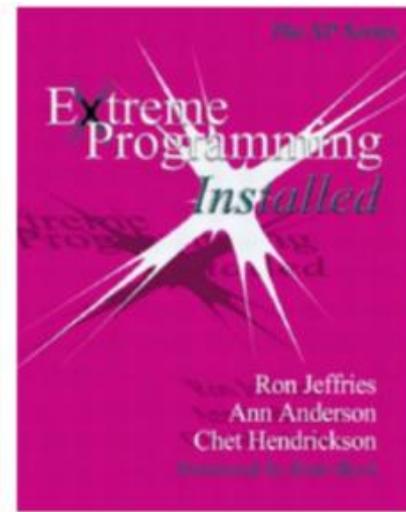
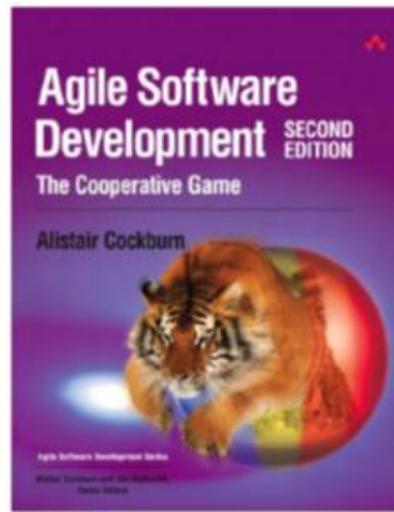
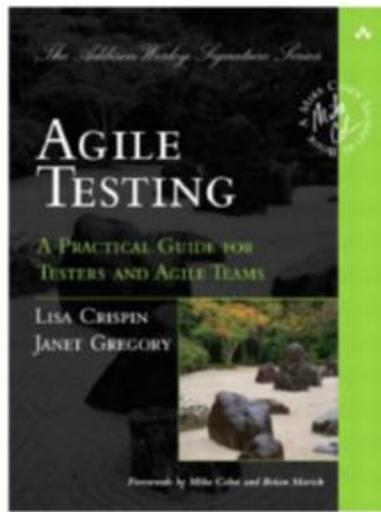
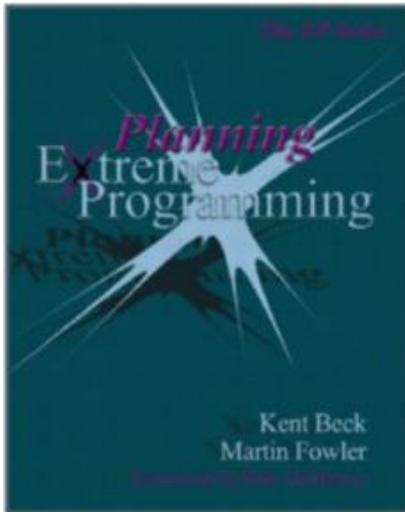
High quality requirements

Adapted from
(Rupp et al., 2014)



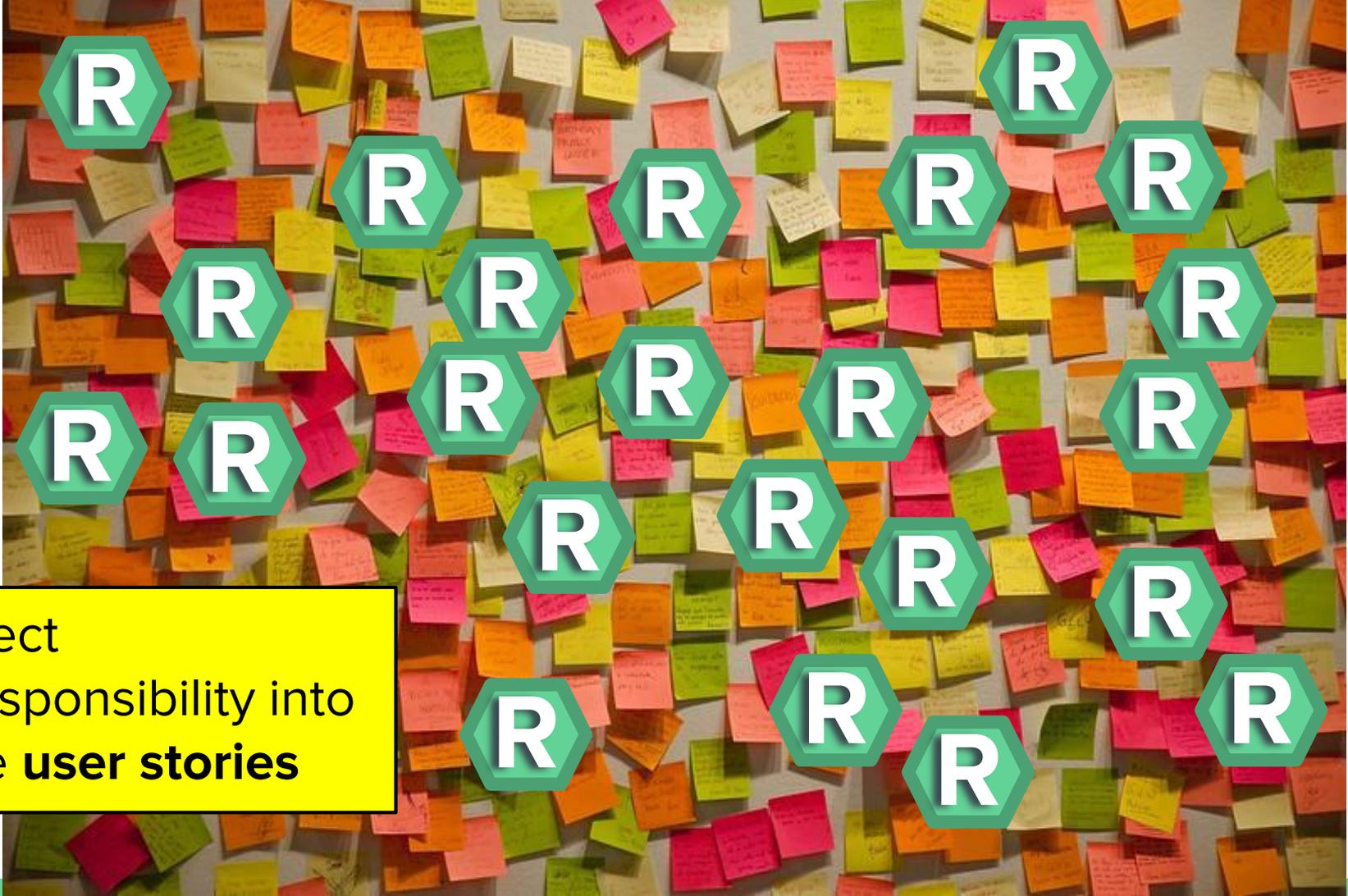
Agile development?

No problem!



Among many others!





Inject
Responsibility into
the **user stories**

**Put it
on your backlog!**

That's how  gets done.

Develop  esponsible AI.

Contact



Hochschule für
Wirtschaft und Recht Berlin
Berlin School of Economics and Law



Prof. Dr. Dagmar Monett

Berlin School of Economics and Law, and AGISI.org



@dmonett



dagmar@monettdiaz.com

AGISI — The AGI Sentinel Initiative



<http://agisi.org>