DOONISDAY:

IS THE WORLD DOOMED

BECAUSE

WE CANNOT EXPRESS OUR

SUSTAINABILITY AND AI GOALS

CLEARLY?

TOM GILB

tom@Gilb.com www.Gilb.com

+47 92066705

Copyright Tom Gilb 2019-2020 March 4 2020 Oslo (from 17:00)

Hosts:

'Digital Innovators Oslo' meetup

Led by

Calvo, Arturo" <arturo.calvo@accenture.com>

Rødsrud, Camilla <camilla.rodsrud@accenture.com>



These slides are placed here:

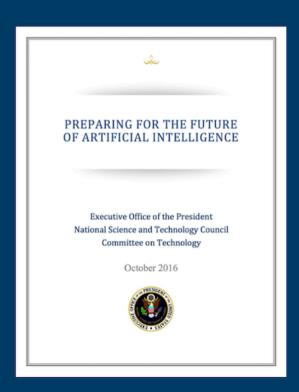
http://concepts.gilb.com/dl964

About 20 minutes of exercises

How clear are Al Principles?

Oct. 2016: "Preparing for the future of Al" from USA





Proposal of Discussion toward Formulation of Al R&D Guideline

Referring OECD guidelines governing privacy, security, and so on, <u>it is necessary to begin discussions and considerations toward formulating an international guideline</u>

<u>consisting of principles governing R&D of AI to be networked</u> ("<u>AI R&D Guideline</u>") as framework taken into account of in R&D of AI to be networked.

Proposed Principles in "AI R&D Guideline"

1. Principle of Transparency

Ensuring the abilities to explain and verify the behaviors of the AI network system

2. Principle of User Assistance

Giving consideration so that the AI network system can assist users and appropriately provide users with opportunities to make choices

3. Principle of Controllability

Ensuring controllability of the Al network system by humans

4. Principle of Security

Ensuring the robustness and dependability of the AI network system

5. Principle of Safety

Giving consideration so that the AI network system will not cause danger to the lives/bodies of users and third parties

6. Principle of Privacy

Giving consideration so that the AI network system will not infringe the privacy of users and third parties

7. Principle of Ethics

Respecting human dignity and individuals' autonomy in conducting research and development of AI to be networked

8. Principle of Accountability

Accomplishing accountability to related stakeholders such as users by researchers/developers of AI to be networked

© 2017 International Business Machines Corporation

Exercise 10 minutes, 2 person teams

- 1. Count undefined but ambiguous words in 1. Transparency, like 'Ensuring'
- 2. **Define**: 'Ensuring' so that it cannot be misunderstood. 100% Unambiguous.
- 3. Quantify 'Ensuring' so that 100% is 'perfect' ensuring, and 0% is none at all. You need some definition here.

Proposal of Discussion toward Formulation of Al R&D Guideline

Referring OECD guidelines governing privacy, security, and so on, <u>it is necessary to begin</u> <u>discussions and considerations toward formulating an international guideline</u> <u>consisting of principles governing R&D of AI to be networked ("AI R&D Guideline"</u>) as framework taken into account of in R&D of AI to be networked.

Proposed Principles in "AI R&D Guideline"

1. Principle of Transparency

Ensuring the abilities to explain and verify the behaviors of the AI network system

2. Principle of User Assistance

Giving consideration so that the AI network system can assist users and appropriately provide users with opportunities to make choices

3. Principle of Controllability

Ensuring controllability of the AI network system by humans

4. Principle of Security

Ensuring the robustness and dependability of the Al network system

5. Principle of Safety

Giving consideration so that the AI network system will not cause danger to the lives/bodies of users and third parties

6. Principle of Privacy

Giving consideration so that the AI network system will not infringe the privacy of users and third parties

7. Principle of Ethics

Respecting human dignity and individuals' autonomy in conducting research and development of AI to be networked

8. Principle of Accountability

Accomplishing accountability to related stakeholders such as users by researchers/developers of AI to be networked

17

Norwegian Government And UN Sustainability Development Goals



Exercise 2: UN Poverty Goal. 10 mins.

Nice but not specific or measurable

Nice but not specific or measurable

Goal 1. End poverty in all its forms everywhere

Target 1.5: By 2030, **build** the resilience of the poor and those in vulnerable situations <u>and</u> **reduce** their exposure <u>and</u> vulnerability to climate-related extreme events <u>and</u> other economic, social <u>and</u> environmental shocks <u>and</u> disasters

- 2 Person exercise.
- 1. Count all ambiguous words in Target 1.5
- (did you miss 'by 2030'?:)
- 2. How would we express the idea of 'resilience' so we know when we have got 100% resilience of the poor (write a suggested scale of measure)
- 3. How many distinctly different objectives are contained in 1.5?

Discuss in your 2 person group

- What is a possible practical consequence of an ambiguous word in a critical objective?
- Do you think the UN has clarified and quantified all 17 critical development goals?
- If not, what are the consequences of unclear objectives in prctice?



End of Exercises.

Now I will share some examples of making Important objectives

Much clearer

Using 'Planguage'

Artificial Intelligence Goals: UX for Al and our Future

Elon Musk:

"Mark my words — A.I. is far more dangerous than nukes"

https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html

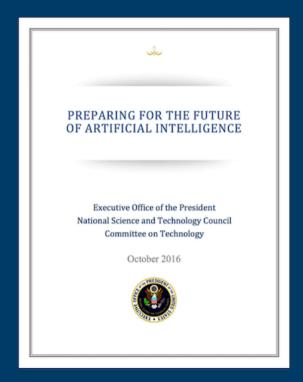


"there needs to be a regulatory body overseeing the development of super intelligence,"

'Principles' = -> 'Qualities' ! * With the usual lack of quantified definition

Oct. 2016: "Preparing for the future of A!" from USA







Proposal of Discussion toward Formulation of Al R&D Guideline

Referring OECD guidelines governing privacy, security, and so on, it is necessary to begin discussions and considerations toward formulating an international guideline consisting of principles governing R&D of Al to be networked ("Al R&D Guideline") as framework taken into account of in R&D of Al to be networked.

Proposed Principles in "A R&D Guideline"

1. Principle of Transparency

Ensuring the abilities to explain and verify the behaviors of the AI network system

2. Principle of User Assistance

Giving consideration so that the Al network system can assist users and appropriately provide users with opportunities to make choices

3. Principle of Controllability

Ensuring controllability of the Al network system by humans

4. Principle of Security

Ensuring the robustness and dependability of the AI network system

5. Principle of Safety

Giving consideration so that the AI network system will not cause danger to the lives/bodies of users and third parties

6. Principle of Privacy

Giving consideration so that the AI network system will not infringe the privacy of users and third parties

7. Principle of Ethics

Respecting human dignity and individuals' autonomy in conducting research and development of AI to be networked

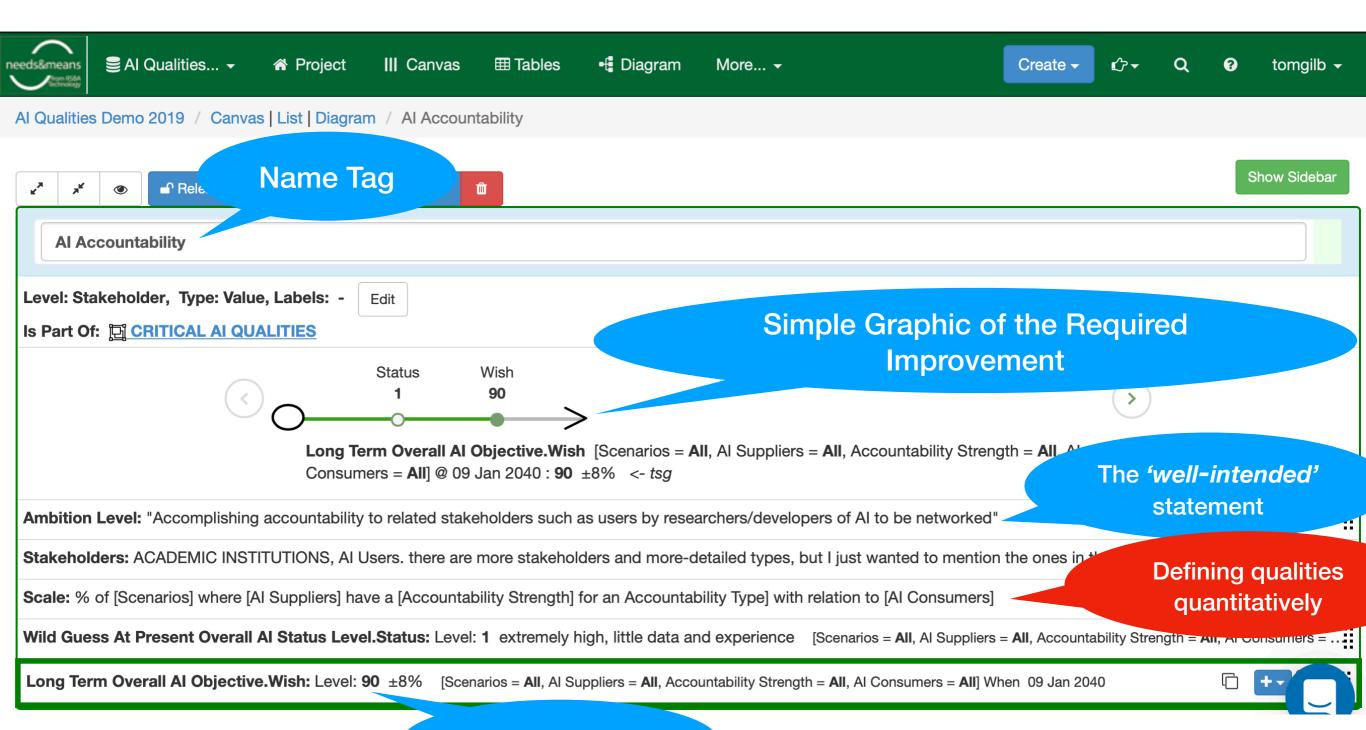
8. Principle of Accountability

Accomplishing accountability to related stakeholders such as users by researchers/developers of AI to be networked

© 2017 International Business Machines Corporation

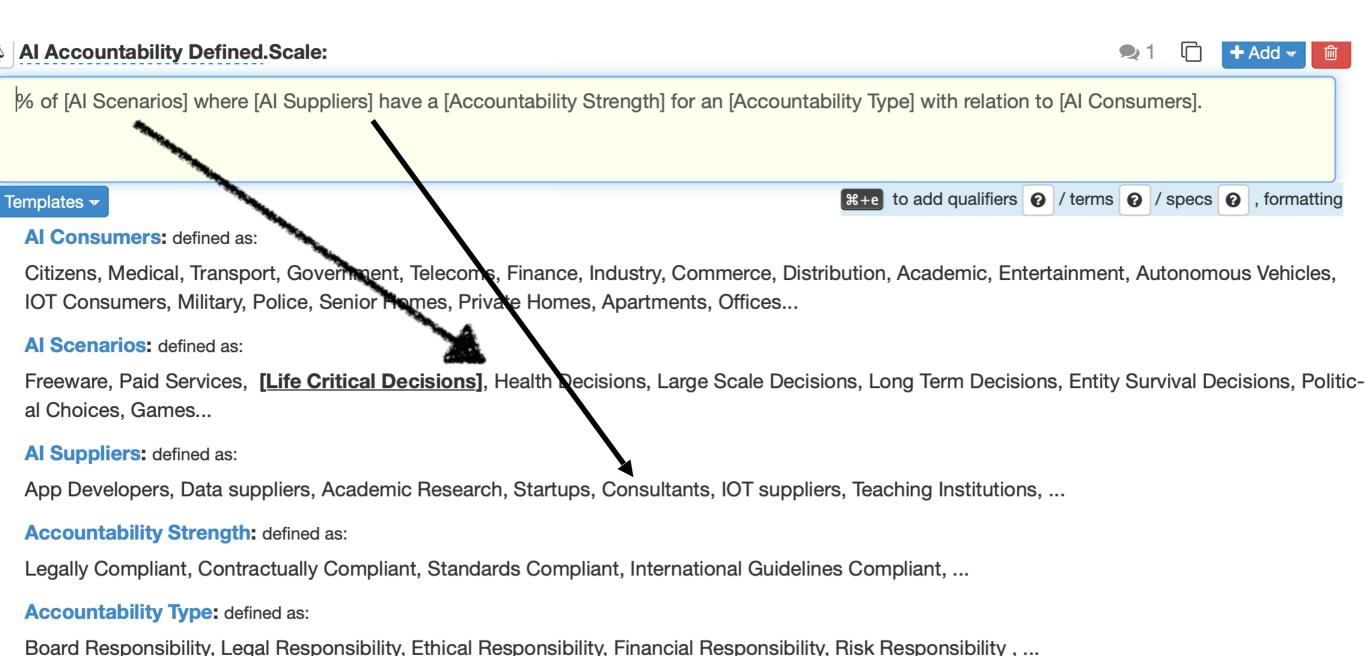
17

Al Accountability Defined Fuzziness -> Quantified. Structured, Enriched Idid this for all XAI Qualities



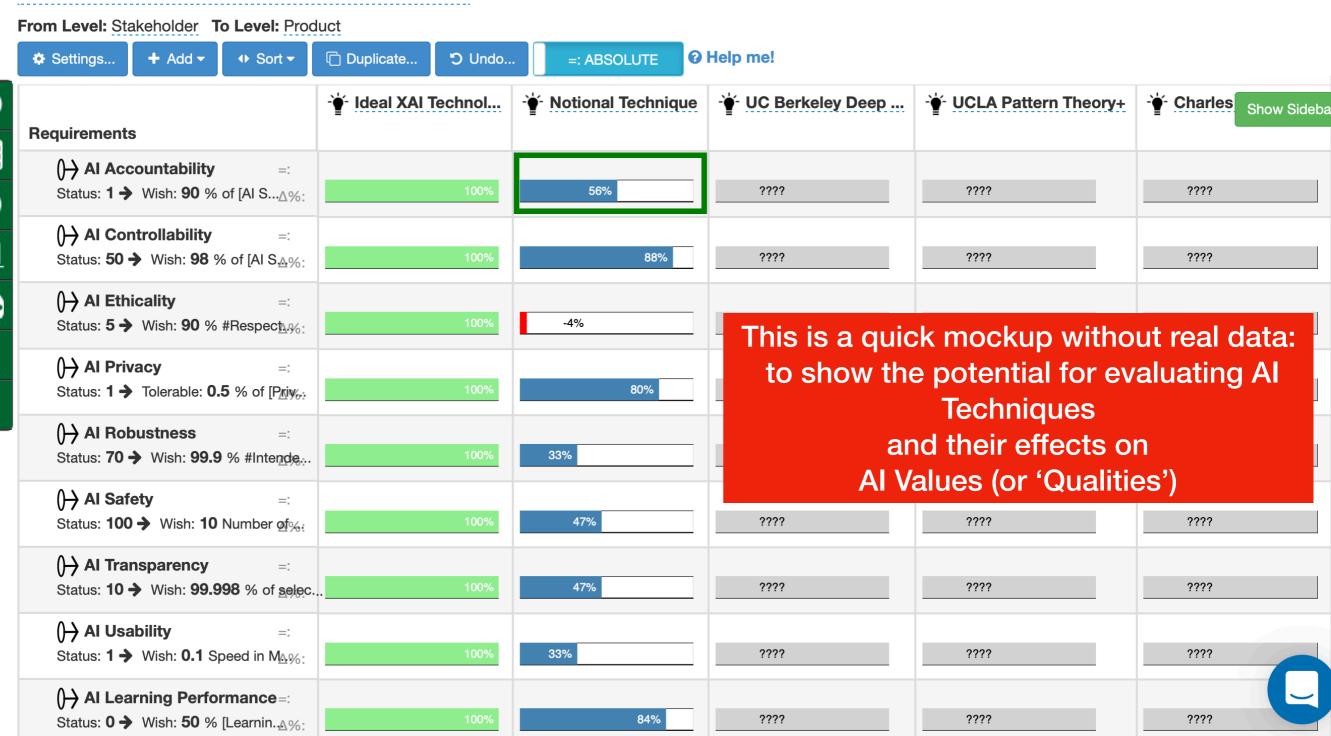
Al 'Accountability' scale Detailed [Scale Parameters] Why is this Useful?

I did this for all XAI Qualities



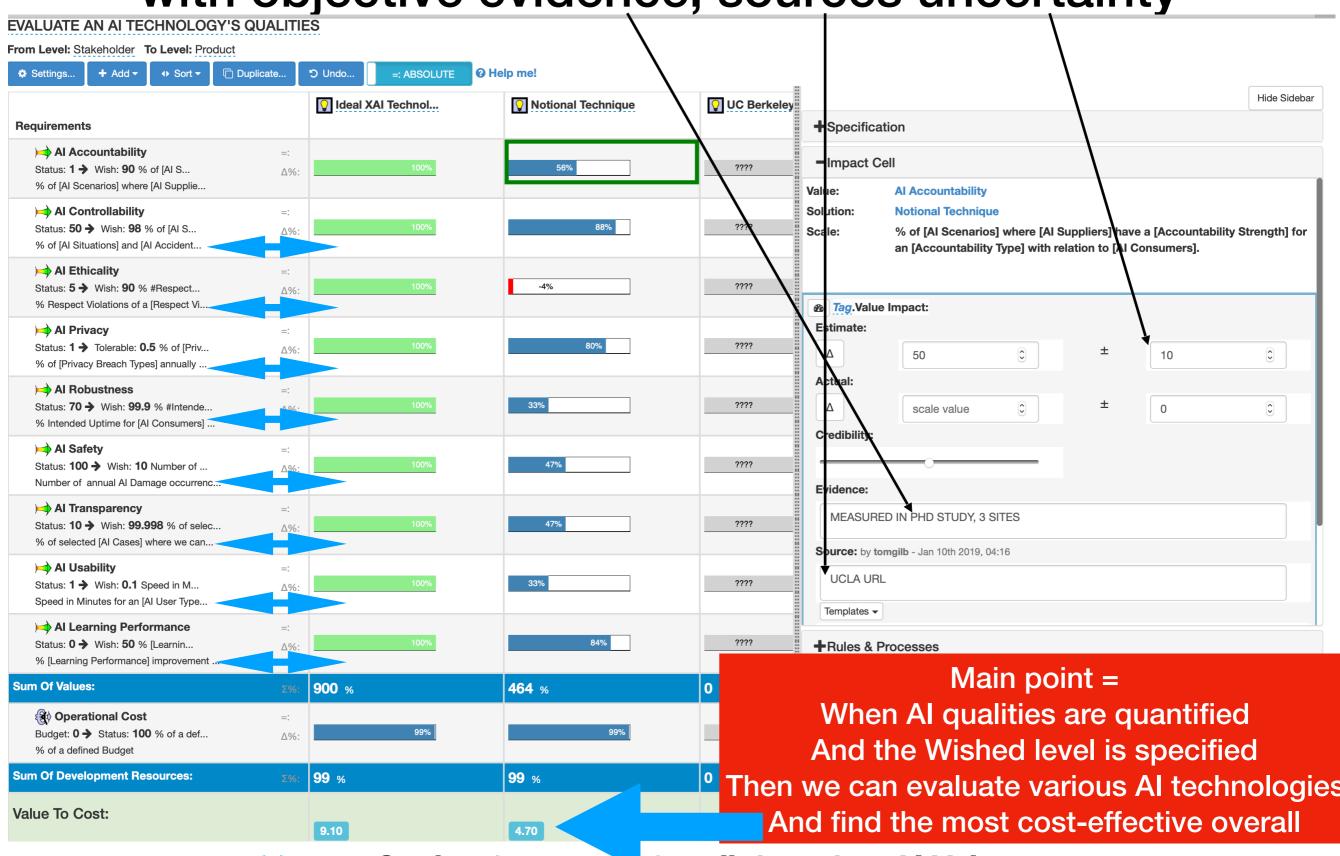
Quantitative Evaluation of Al Technology Qualities using 'Impact Estimation Tables'

EVALUATE AN AI TECHNOLOGY'S QUALITIES



Al 'Technique Evaluation'

with objective evidence, sources uncertainty



= Scale of measure for all the other Al Values

My (TSG) Observations and Conclusions

- The most fundamental obstacle to Al Standards progress is WE NEED TO QUANTIFY AND STRUCTURE-RICHLY all CRITICAL STAKEHOLDER VALUES (ALL > 8)
- Next: we need much-deeper, more-comprehensive identification, and consensus, and detailed knowledge (their values and constraints) about Al/ Al STANDARDS **STAKEHOLDERS**

Proposal of Discussion toward Formulation of Al R&D Guideline

Referring OECD guidelines governing privacy, security, and so on, it is necessary to begin discussions and considerations toward formulating an international guideline consisting of principles governing R&D of Al to be networked ("Al R&D Guideline") as framework taken into account of in R&D of Al to be networked.

Proposed Principles in "AN R&D Guideline"

1. Principle of Transparency

Ensuring the abilities to explain and verify the behaviors of the Al network system

2. Principle of User Assistance

Giving consideration so that the AI network system can assist users and appropriately provide users with opportunities to make choices

3. Principle of Controllability

Ensuring controllability of the AI network system by humans

4. Principle of Security

Ensuring the robustness and dependability of the AI network system

5. Principle of Safety

Giving consideration so that the AI network system will not cause danger to the lives/bodies of users and third parties

6. Principle of Privacy

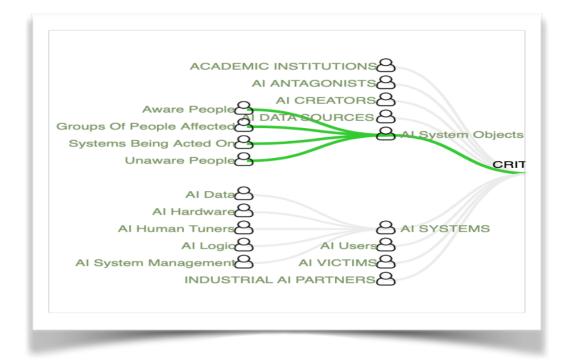
Giving consideration so that the AI network system will not infringe the privacy of users and third parties

7. Principle of Ethics

Respecting human dignity and individuals' autonomy in conducting research and development of AI to be networked

8. Principle of Accountability

Accomplishing accountability to related stakeholders such as users by researchers/developers of AI to be networked



Details of Quantifying all the XAI Objectives

Oct. 2016: "Preparing for the future of AI" from USA

Proposal of Discussion toward Formulation of Al R&D Guideline

Referring OECD guidelines governing privacy, security, and so on, it is necessary to begin

PREPARING FOR THE FUTUR OF ARTIFICIAL INTELLIGENCE

Executive Office of the President
National Science and Technology Council
Committee on Technology
October 2016

discussions and considerations toward formulating an international guideline consisting of principles governing R&D of Al to be networked ("Al R&D Guideline") as framework taken into account of in R&D of Al to be networked.

Proposed Principles in "Al R&D Guideline"

1. Principle of Transparency

Ensuring the abilities to explain and verify the behaviors of the AI network system

2. Principle of User Assistance

Giving consideration so that the AI network system can assist users and appropriately provide users with opportunities to nake choices

3. Principle of Controllability

Ensuring controllability of the AI network system by humans

4. Principle of Security

Ensuring the robustness and dependability of the Al network system

5. Principle of Safety

Giving consideration so that the AI network system will not cause danger to the lives/bodies of users and third parties

6. Principle of Privacy

Giving consideration so that the AI network system will not infringe the privacy of users and third parties

7. Principle of Ethics

Respecting human dignity and individuals' autonomy in conducting research and development of AI to be networked

8. Principle of Accountability

Accomplishing accountability to related stakeholders such as users by researchers/developers of AI to be networked

2017 International Business Machines Corporation

17

XAI Explaining AI Lecture Slides

http://concepts.gilb.com/dl958

A Serious 'Multi-dimensional Metrics Attack' on Poor Al 'Academic and Standards' Thinking & Planning.

An analysis of published Principles for Managing and Standardizing AI,

where about 10 Al Qualities like 'Safety' and 'Transparency' are shown to be quantifiable.

This is prelude to rational thinking about the entire subject.

GilbFest Talk June 25 2019
Presented to the IBM AI Standards Committee Members
(who said they were suitably impressed)



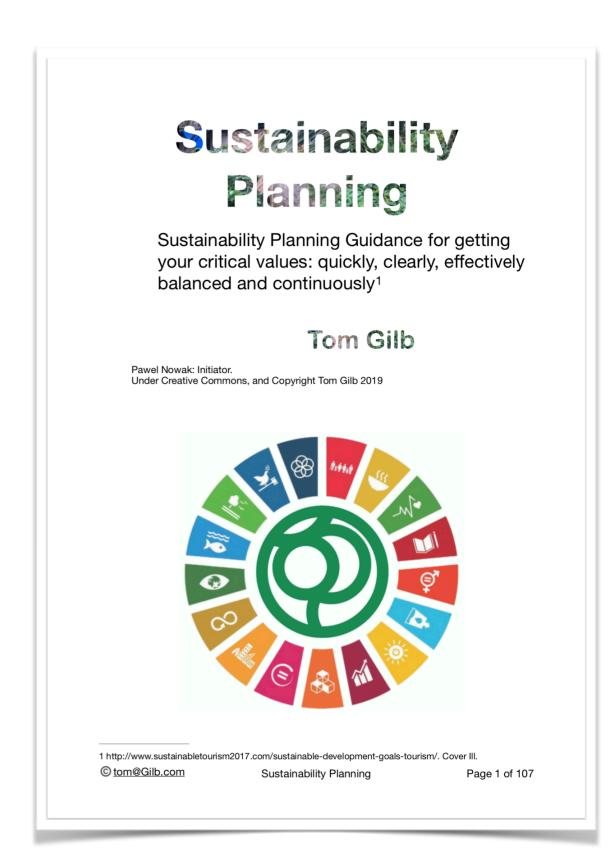
Do politicians like BS?

If you bulls**t an AI system, could you fool it into to making bad decisions?

• The 18 United Nations Sustainable Development Goals

UX for the Planet

- "Sustainability Planning"
- https://tinyurl.com/
 UNGoalsGilb
- Free book, with details of how to analyze, clarify.



Feedback 🗀





Our business Operational performance Future energy Financials Governance UK Power Networks Services

Transparency

Organisational resilience

Board of Directors

Executive Management Team

Transparency

We are acutely aware of the privileged position we occupy as a monopoly provider of an essential service. We are a top performing UK DNO group and accept that this high performance brings with it increased scrutiny from our customers, other stakeholders and the industry at large. We are committed to being open about our operations and, through documents such as this Annual Review and our RIIO-ED1 commitments report, we seek to make it easy for all our stakeholders to understand how we are performing and, most importantly, how My clients are using customers.

We believe that it makes business sense for our em vision. The main way employees are incentivised is In line with our vision, our employees are rewarded for customer service and cost saving.

these UN Goals as a framework. But, do all ess stakeholders have an heme). identical clear iity, understanding of the goals?

Since its inception in 2010 the key focus of UK Power Networks. at our customers want at the lowest possible cost. We have delivered significant improvement me reliability of our

On this page you will see how we have performed against the following Sustainable Development Goals:









90%

of UK Power Networks employees are covered by the Trade Union **Collective Agreements**

End poverty in all its forms everywhere.



Nice but not specific or measurable

Nice but not specific or measurable

Attempt to find something measurable But

IS THIS REALLY THE CORE OF POVERTY CONCERNS? (No!)

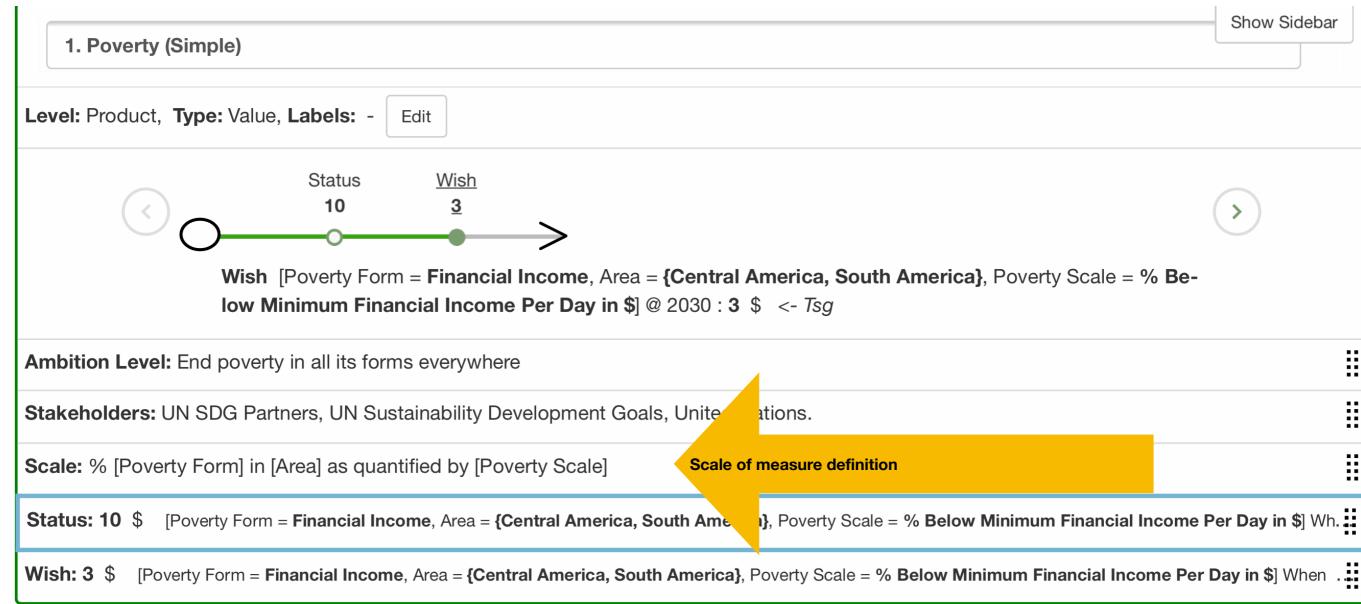
No specific targets set

Goal 1. End poverty in all its forms everywhere

Target 1.5: By 2030, **build** the resilience of the poor and those in vulnerable situations <u>and</u> **reduce** their exposure <u>and</u> vulnerability to climate-related extreme events <u>and</u> other economic, social <u>and</u> environmental shocks <u>and</u> disasters

- Indicator 1.5.1: Number of deaths, missing persons and directly affected persons attributed to disasters per 100,000 population
 - See Metadata:
- Indicator 1.5.2: Direct economic loss attributed to disasters in relation to global gross domestic product (GDP)
 - See Metadata :
- Indicator 1.5.3: Number of countries that adopt and implement national disaster risk reduction strategies in line with the Sendai Framework for Disaster Risk Reduction 2015-2030
 - See Metadata :
- Indicator 1.5.4: Proportion of local governments that adopt and implement local disaster risk reduction strategies in line with national disaster risk reduction strategies
 - See Metadata:

Simple Goal 1(Poverty) Draft



 Problem. This spec is too simple; it cannot serve the purposes of modelling critical aspects of complex systems.



Facts and Figures

Goal 1 Targets

Links

at LT

- **1.1** By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day
- **1.2** By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions
- **1.3** Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable
- **1.4** By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance
- 1.5 By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters
 See next slides, for analysis of this
- **1.A** Ensure significant mobilization of resources from a variety of sources, including through enhanced development cooperation, in order to provide adequate and predictable means for developing countries, in particular least developed countries, to implement programmes and policies to end poverty in all its dimensions
- **1.B** Create sound policy frameworks at the national, regional and international levels, based on pro-poor and gender-sensitive development strategies, to support accelerated investment in poverty eradication actions 22

Values and Resources 1. Poverty (Decomposed) 1.1 Financial Poverty 1.2 Poverty In All Dimensions 1.3 Social Poverty 1.4 Resources. Poverty 1.5 Disaster Protection P 1.A National Poverty Reduct... 1.B Investment Frameworks...

Ok Let's take Poverty Target 1.5 as an example

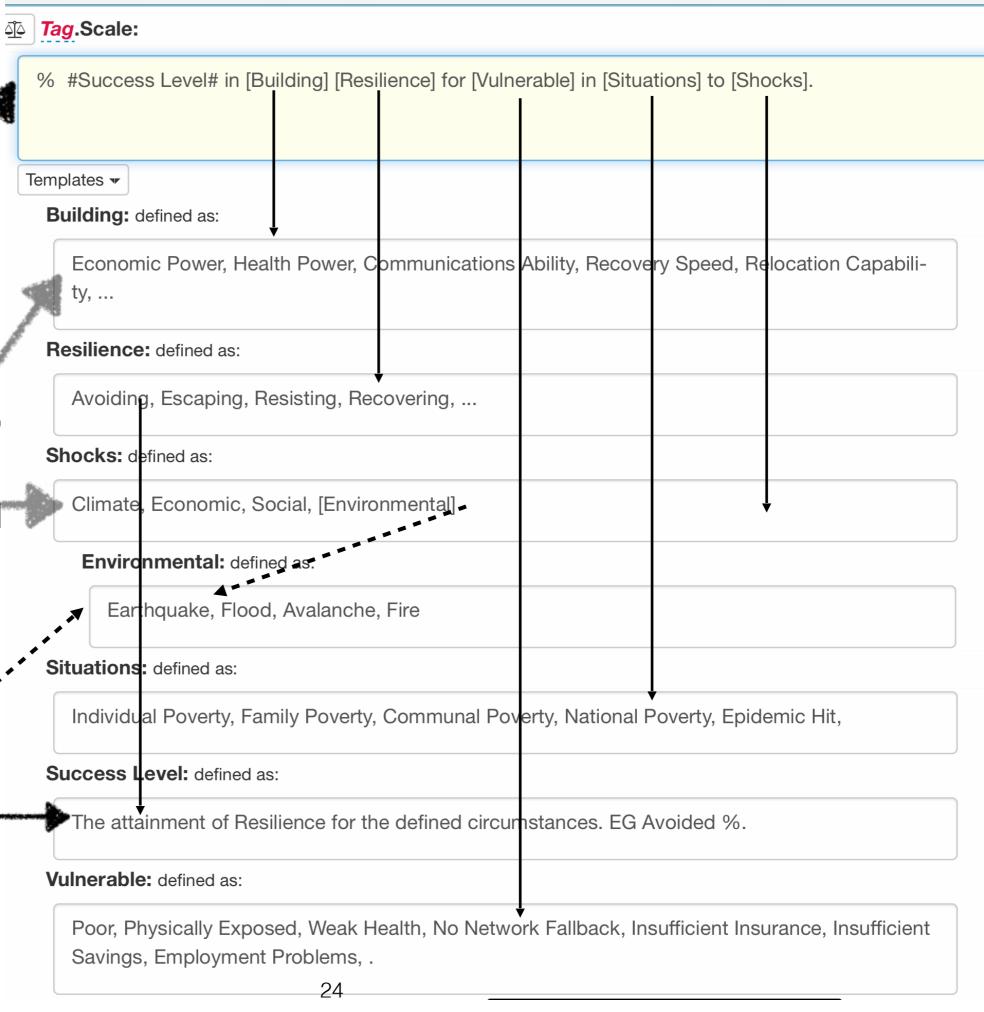
"By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters"



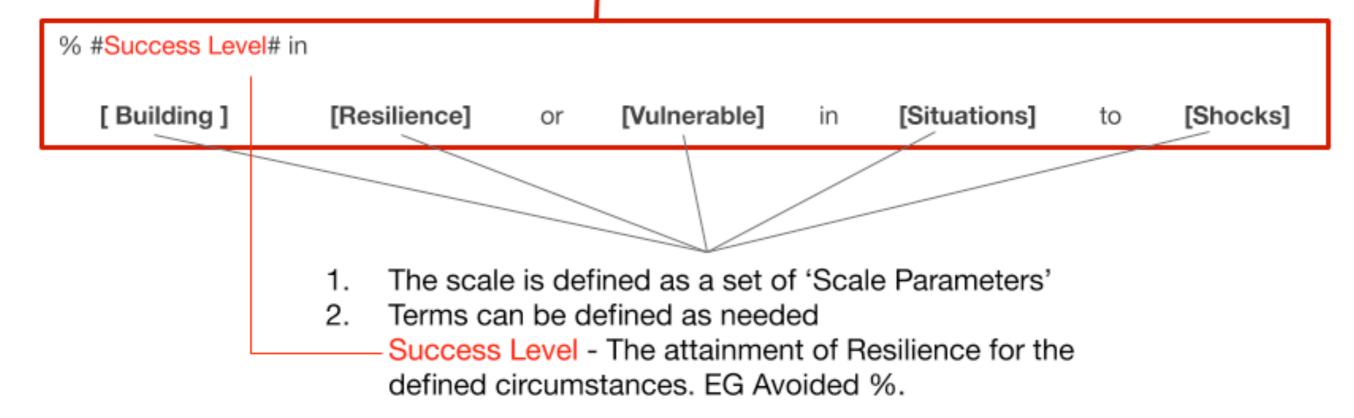
<- see next slides for analysis and re-organization of this

Defining a structured scale of measure

- The Scale is defined as a set of 'Scale Parameters'
- Each SP is defined as a set of conditions
- Each condition can be defined
- Terms can be defined as needed



Defining a structured scale of measure



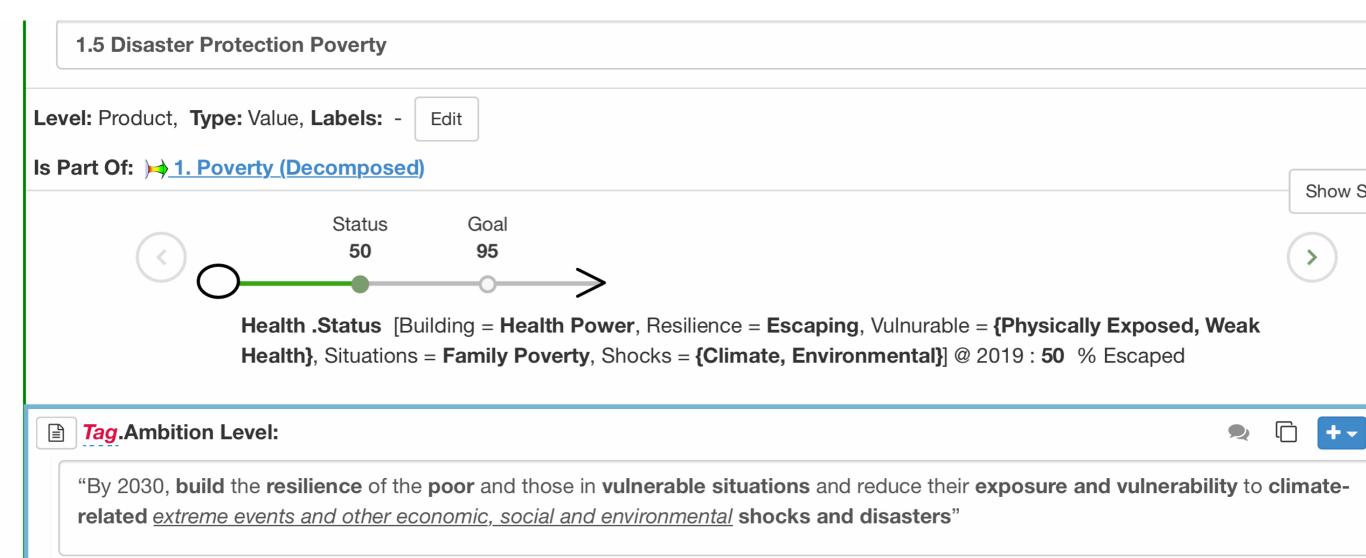
Defining a structured scale of measure



Each Scale Parameters is defined as a set of conditions

The Scale Parameter Structure

Graphic by anna.maria.karlowska@gmail.com, 2019



 Now we can apply the scale and some selected Parameters to define the status quo, and the Goal level

1.5 Disaster Protection Poverty

Show Sidebar

Level: Product, Type: Value, Labels: - Edit

Is Part Of: 1. Poverty (Decomposed)



 $\textbf{Overall Long Term.Tolerable} \hspace{0.2cm} \textbf{[Building = AII, Resilience = AII, Vulnerable = AII, Situations = AII, Shocks = AII]} \\$

@ 2030 : **90** %. ±5%

Ambition Level: "By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-.

Scale: % Success Level in [Building] [Resilience] for [Vulnerable] in [Situations] to [Shocks].

Stakeholders: UN Children's Fund, UN Global Compact, UN International Strategy For Disaster Reduction.

Health .Status: 50 % Escaped [Building = Health Power, Resilience = Escaping, Vulnerable = Weak Health, Situations = Family Poverty, Shocks = {Climate, Env

Health .Goal: 95 [Building = Health Power, Resilience = Escaping, Vulnerable = Weak Health, Situations = Family Poverty, Shocks = {Climate, Environmental}] Wh.

Overall Long Term.Status: 5 % [Building = All, Resilience = All, Vulnerable = All, Situations = All, Shocks = All] When 2019

Overall Long Term.Goal: 98 % ± 1% [Building = All, Resilience = All, Vulnerable = All, Situations = All, Shocks = All] When 2030

Setting long and short term Goals, re-using the Scale of measure Using a variety of scale *levels* (Wish, Goal, Stretch, Tolerable, Past, Status), and [Scale parameters].

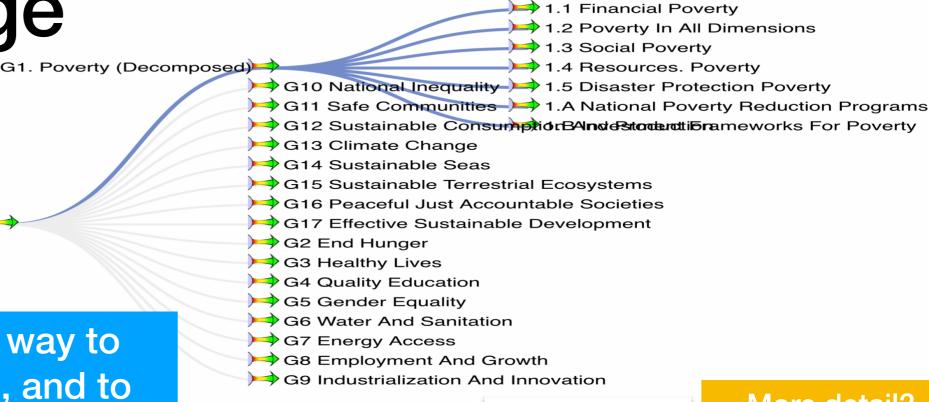
To articulate: <u>How much</u> 'poverty' where, when, for whom, which type, under which conditions.

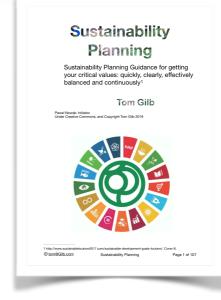
I am done with my main message

"There is a smarter way to clarify, to structure, and to quantify all critical value objectives

TOP 17 Goals

If you do not clarify, you will for sure waste your time trying to get the results you dream of"





More detail?
"Sustainability
Planning"
Digital Book 2019
https://
tinyurl.com/
UNGoalsGilb
Later see
Gilb.com

Now if I have time I'd like to give you one insight into some advantages of this kind of goal quantification and structuring

Main idea

You can now evaluate your

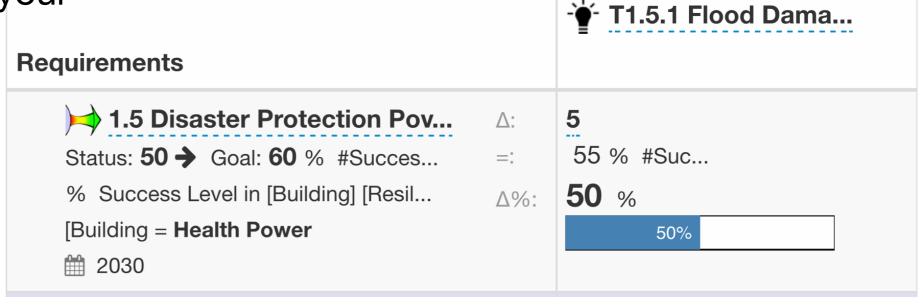
<u>strategies</u>

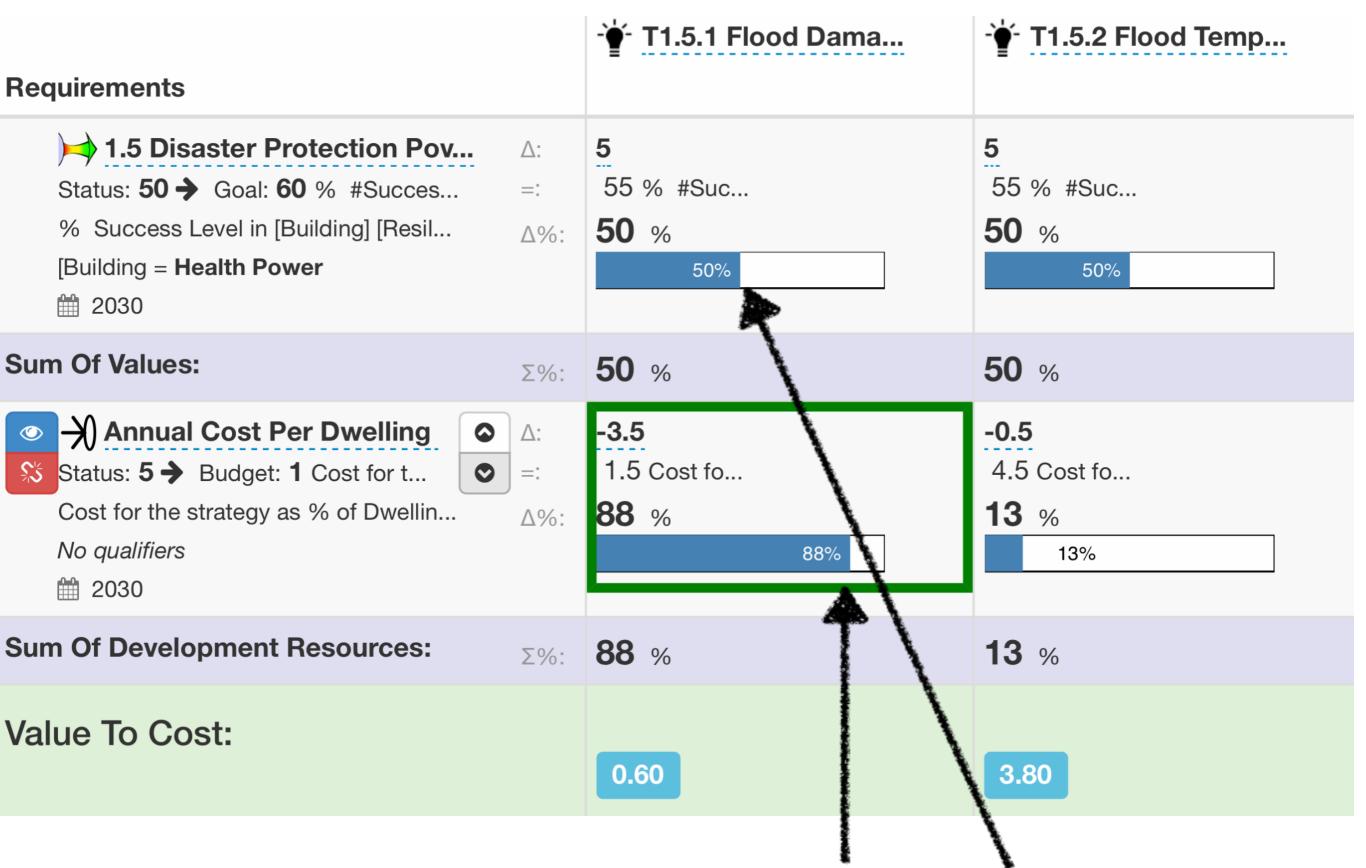
Quantitatively

In many dimensions

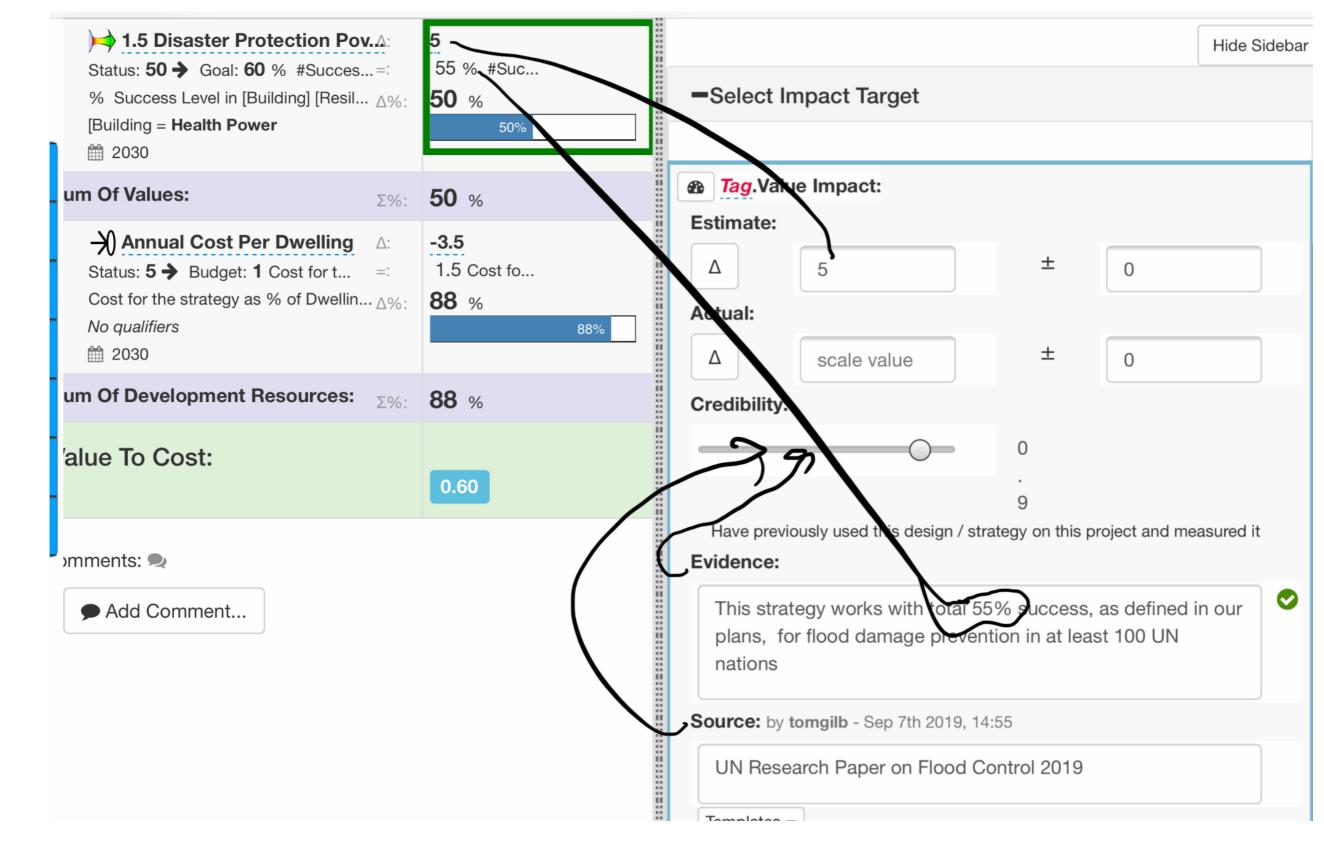
With regard to risks

And costs





We can now estimate how cost-effective our strategy options will probably be.

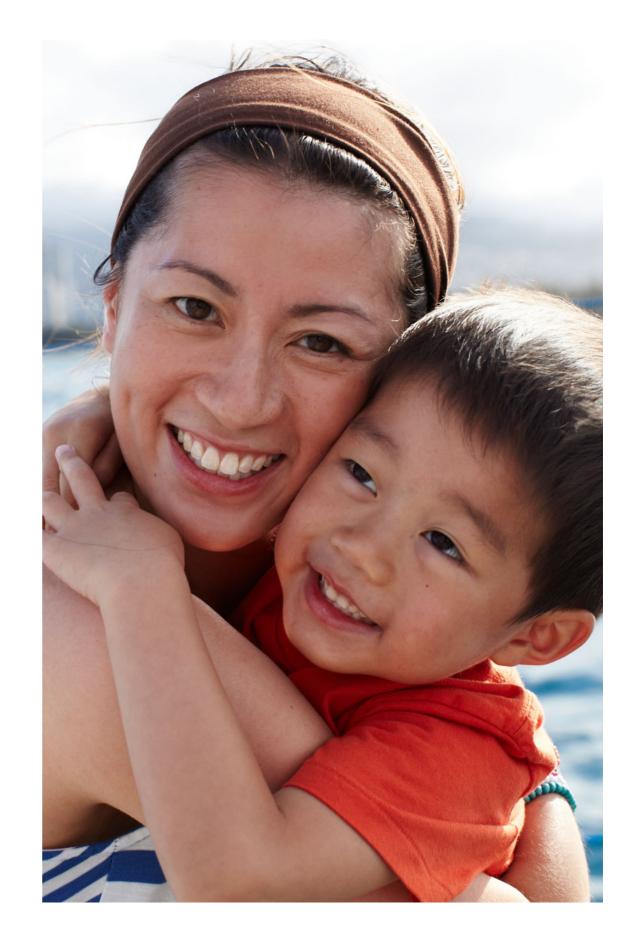


And if there is time for one more slide.

This is how we estimate the credibility of our cost-effectiveness estimates, so we know what to trust, and what to prioritise for good results.

The Concrete Steps for You are

- NO BLA BLA BLA: Decide that important ideas need great clarity
 CLARIFY THE CRITICAL: Always clarify and quantify the critical values
- 3. **IDEAS MUST DELIVER VALUES**: Evaluate all designs and strategies for how well they satisfy your critical values



More detail?

More detail?

"Sustainability Management"

Digital Book 2019

https://tinyurl.com/UNGoalsGilb

Later see

gilb.com

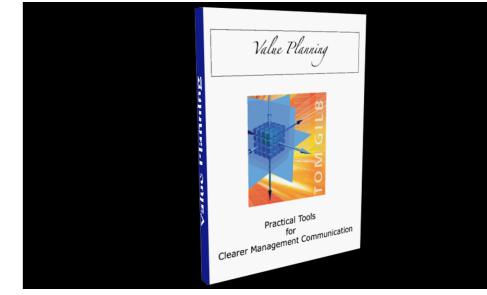
XAI Explaining AI

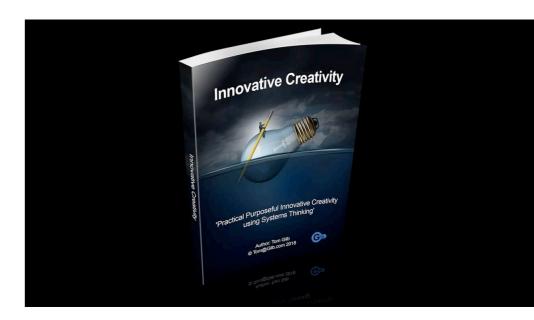
Lecture Slides http://concepts.gilb.com/dl958

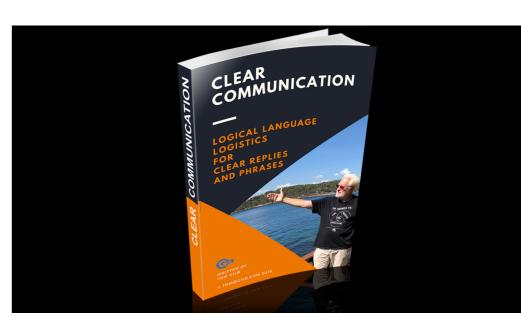
Training? www.Gilb.com

OSWA

https://www.meetup.com/Oslo-Software-Architecture/ (free Oslo course March 5-6 2020) In English







https://www.gilb.com/about

Now maybe you do not care about saving the human race (yet)



The good news is that these exact same methods (Gilb, Planguage) can be used for your **Usability**Planning!

This is a child who knows nothing about science and shouldn't be given a platform



The other is Greta Thunberg