AI FOR SOCIAL USEFULNESS

EMBEDDING PRINCIPLED GUIDELINES INTO AI SYSTEMS AND RELATED LAWS

<u>Dedication</u>: *To Wills and Olivia Dahl, in loving memory of their father, international Human Rights lawyer Henry Dahl, who devoted his life to ensure Corporate Social Responsibility, and advised me on legally prudent wordings of my writings* ☺.

Veronica Dahl
Researcher, NSERC
Director, Curiessa Consulting, Ltd.
Sci. Advisory Board Member, IMDEA Software
Professor Emeritus, Computing Sciences, SFU

Outline

I- Evolutionary Context that led to AI not necessarily being used first and foremost to benefit the public that funded it, nor to necessarily focus on human values and concerns

II- Constraint Handling Rules- a handy way to *embed principled guidelines into AI systems*

III- Re-evaluating "intelligence", and steering AI to the goal of "meeting the needs of all within the means of the planet" (K. Raworth)

Motivation: un-just uses of AI

- "black box" AIs that cancel our right to information and explanation

- AIs that are not ready to replace humans

- Public forced to work unpaid as debuggers

- Exploitation of users' data without their knowledge/consent

- Robo-signed mass actions

- AIs purported to replace humans don't pay taxes, benefits, unemployment insurance, nor do they contribute to a fund for retraining humans into new jobs -> GROWING UNEMPLOYMENT

AI is enabling informational dictatorships able to manipulate the behaviours of humans and organizations alike, and even to erode representative democracies and world peace

PART I- (Holistic) EVOLUTIONARY CONTEXT:
Shift from partnership to dominator norms (R. Eisler)

- *7000 and 3500 B.C.,* in Old Europe: *long period of peace and prosperity when all the basic technologies on which civilization is built were developed*

  <u>Partnership Societies</u>: -not male dominant, violent, hierarchic.
  - Depictions of divine power are feminine
  - priestesses had spiritual authority.
  - Heads of matrilineal clans held the land as trustees for the people

  From around 5000 B.C., long line of bloody invasions in the old Neolithic cultures, *by warring peoples ruled by a caste of warrior-priests.*

  The old values did not fit their dominator values.

  Subjugation + a millenary process of relentless indoctrination (Oresteia, "sacred" stories like the Bible)
→DEHUMANIZATION OF WOMEN→ DOMINATOR SOCIETIES

"AND MAN CREATED GOD IN HIS OWN IMAGE..."

Shift into emphasizing technologies designed to destroy and dominate- which of its characteristics continue into our time?

- violent reduction of women, and thus also of both their female and male children, to the status of mere male possessions (documented in Kurgan burial practices)

- *appropriation by men of important religious symbols* previously associated with women in the worship of the Goddess.

- *Higher value is placed in the power that takes, vs. gives, life.*

- *male dominance, male violence, and a generally hierarchic and authoritarian social structure becomes the norm.*

- *the way in which they characteristically acquired material wealth was not by developing technologies of production, but through ever more effective technologies of destruction.*

It is this emphasis on domination which threatens all life on our globe (and explains AI being misused).

Where has our emphasis on domination taken us?

The developed world has been hyper "successful" in colonizing
others: *it now consumes its fair share of 8 planet Earth's
resources, and is the planet's self-entitled main polluter.
Main contributor: Cheap energy during 200 years.*

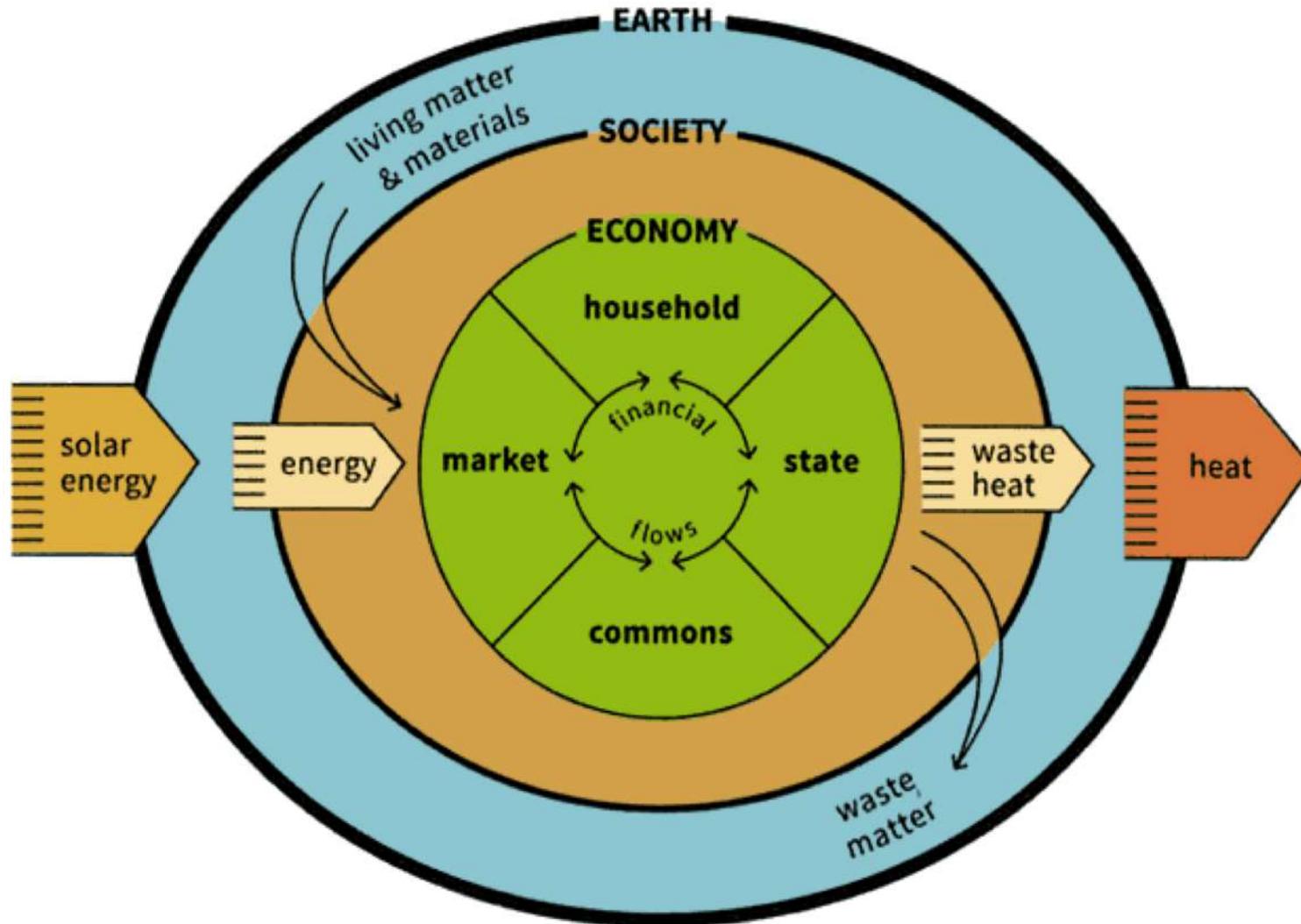*Men* have, *as a group*, been hyper "successful":
- *their work is rewarded with a disproportionate share of wealth,
  power and status,* as if it didn't crucially depend on the (unpaid
  and underpaid) work of others (the female half of humanity);
- *their violence and indoctrinations against that other half have
  been "normalized" and institutionalized*

Male dominance has become foundational to domination mindsets,
making all other dominations unconsciously more "acceptable"

*Technologies of destruction (including uses of AI for domination and
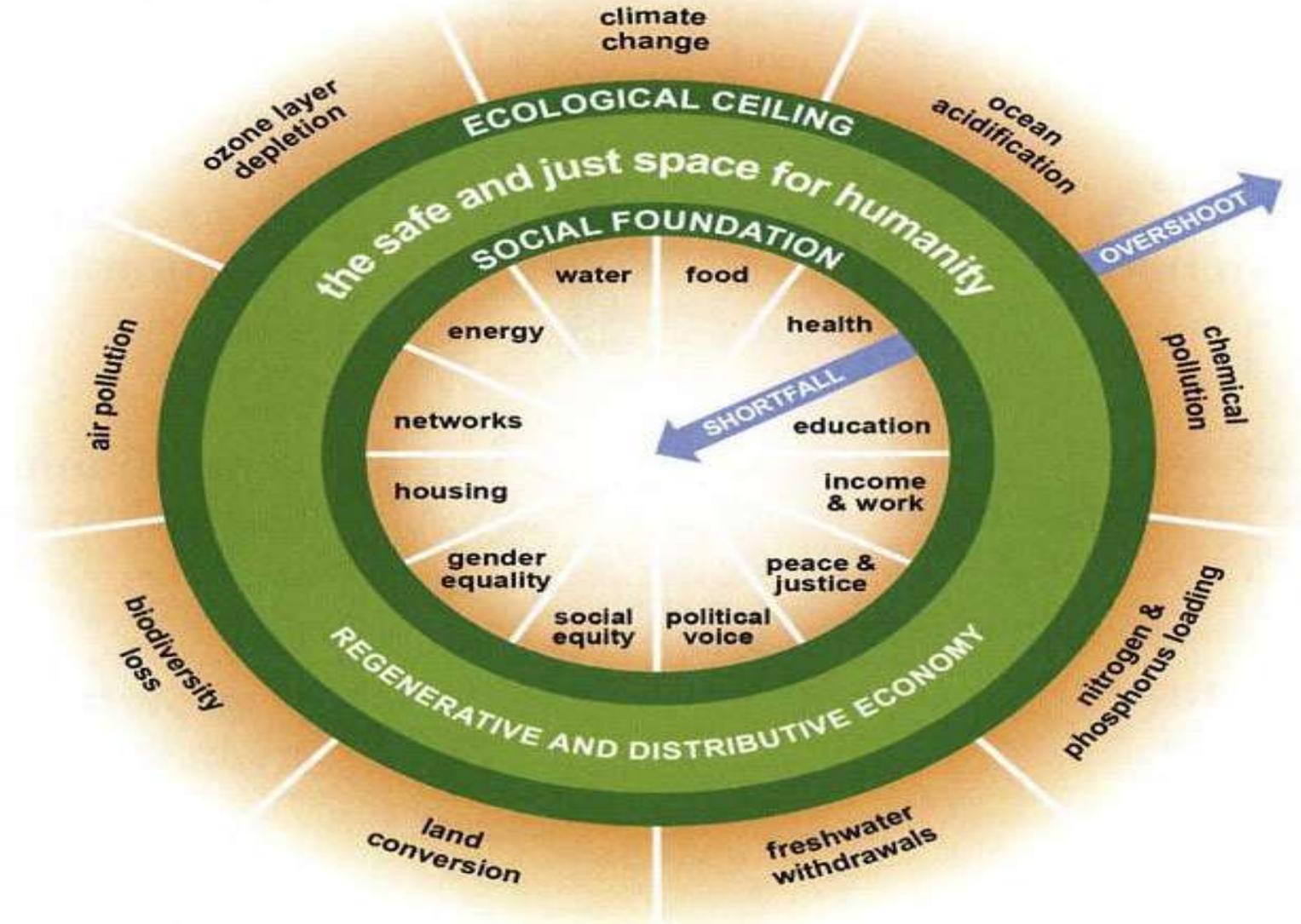exploitation)* have never been as "successful".

Disclaimer ☺: The problem is not men as a sex, but men and women
as they must be socialized (or brutalized) into a dominance system

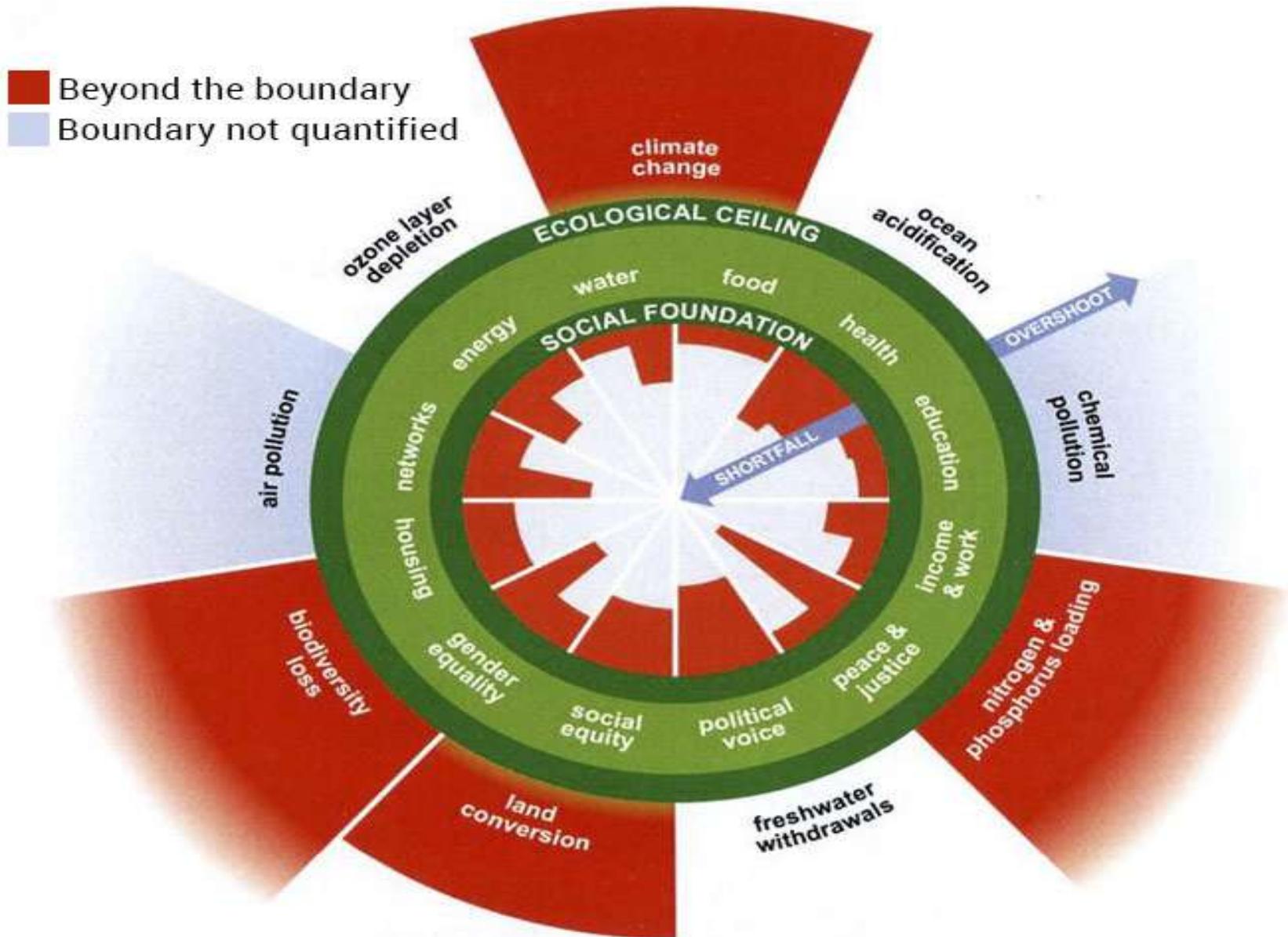*The Embedded Economy- all contributions made explicit from Kate Raworth's "Doughnut Economics"*

We have transgressed both sides of the doughnut's boundaries

PART II- Constraint Handling Rules for principled decisions

Form of a constraint:

t

(where t is a logic term).   *Effect: t is added to the constraint store*

Form of a Rule:

$$c_1, \ldots, c_n ==> \text{Guard} | \text{Body}.$$

Where c_i are constraints,
Guard is a series of Prolog calls,
Body is a new set of constraints.

Effect: if $c_1, \ldots, c_n$ are in the store and Guard holds, Body is added to the constraint store.

AI system that alerts doctors, on the basis of

    - the number of subjects from the article(s) used,
    - the minimum required number

mined(Paper, Medecine, Medical_Area,Number_of_Subjects) ==>
    minimum_required_subjects(Medical_Area,Min),
    Min>Number_of_Subjects |
                alert(Medical_Area, Medecine).

Example 2. thalidomide only had 8 victims in the USA thanks to the heroic resistance of a Canadian pharmacologist, Frances Kelsey.

tested_only_on(mice,Drug) ==>
        do_not_prescribe(Drug, humans).

tested_only-on(men,Drug) ==>
        do_not_prescribe(Drug,women),
        do_not_prescribe(Drug,children).

Field by field, we could create inferential KBs that belonged to all, where fake news and fake studies could be debunked in verifiable manner

SAMPLE AI PROGRAMS FOR RESPONSIBLE DECISION MAKING: 3. **CSR**(Corporate Social Responsibility)

Recommend decisions *taking into account a **user's goals** and **stated values***, and *explain the reasons* for its recommendations.

<u>Example</u>: If I query the system on whether to invest on Dow Chemical having stated:

- a minimum desired return of 5% and
- a high value for the human rights criterion,

I might get the following advice, together with a link:

Do not invest on Dow Chemical because according to Wikipedia, it sells chemicals that damage the human nervous system (and have been banned from the US for that reason), to third world countries that do not yet have protective regulations.

(co-authors: E. Miralles, B. Coleman, E. Maharshak)

Conflicting advice

Of course, *by consulting other sources* (e.g. Dow Chemical's own website), *we could have obtained*:

Do invest on Dow Chemical because it addresses many of the world's most challenging problems such as the need for clean water, renewable energy generation and conservation, and increasing agricultural productivity.

We allow and output the possible contradicting advice and its rationale in each case.

WITH SIMILAR SYSTEMS, we can conjure as much, and as representative, citizen participation as possible, for decision making on, in particular, how to instrumentalize AI for social benefit.

PART 3. *Re-evaluating (Artificial) "Intelligence"- most of it is not AI, it's just SOFTWARE*

"How to prevent Silicon Valley from generating a new global subclass" by Mary L. Gray and Siddharth Suri

Internet giants resort to an army of ghost workers to correct the errors of their algorithms in real time.

Recruited in technological platforms like Mturk (Amazon Mechanical Turk) where millions of microtasks are distributed from a piece of software that requires a job done, to a human who accepts the task.

AI appears to be more successful than it really is thanks to the brute force exploitation of precarious workers and of our private data. Its "intelligence" is not all that artificial…

Re-evaluating "intelligence"

Collective intelligence is determined by three factors [1]:

- " *social sensitivity*" of group members

- the *number of speaking turns taken by members of the group*: where a few people dominated the conversation were less collectively intelligent than those with a more equal distribution of conversational turn-taking.

- the *number of female members*: the more women in the group, the higher the group's IQ.

E.g., teams with an equal gender mix perform better than male-dominated teams in terms of sales and profits [2], … and even in non-intellectual tasks, like the Olympics [3]

(1) https://futureoflife.org/2016/06/13/collective-intelligence-of-women-save-world/?cn-reloaded=1

(2) http://gap.hks.harvard.edu/impact-gender-diversity-performance-business-teams-evidence-field-experiment

(3) https://news.ubc.ca/2014/09/30/gender-equality-olympics

"Only women can stop the apocalypse"(FP News)
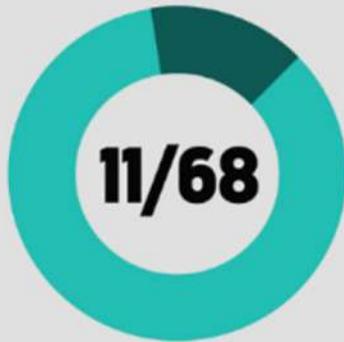## (a cop-out! :

Restore the power they were robbed of, and do your share!)

Research shows that:

- absent women's full participation in nuclear issues,

- *- negotiated agreements are less likely to hold*: when women's organizations participate in peace negotiations, agreements are 64 percent less likely to fail
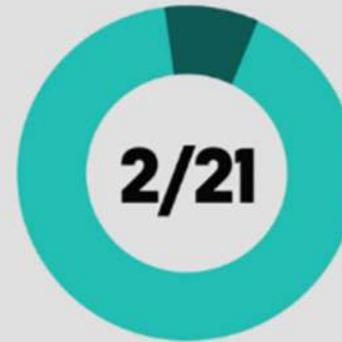
- *- innovative ideas are left unheard*.

- women's status, including their representation in decision-making, is the best predictor of state peacefulness

- More repressive attitudes towards women are predictors of periods of aggressive warfare, militarism, imperialism, totalitarian dictatorships, domestic and international violence.
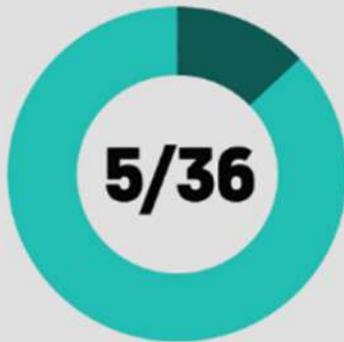
11/68 Total number of women at the Department of State

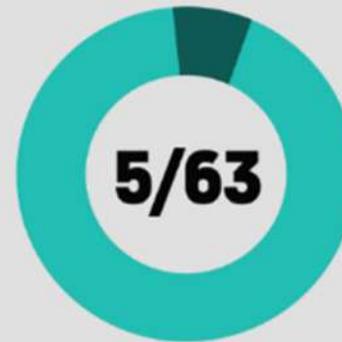**2** Women of color at the Department of State

2/21 Total number of female National Security Advisors

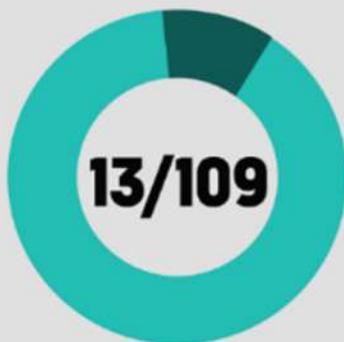**2** National Security Advisors were women of color

5/36 Total number of women at the Department of Energy

1 Woman of color at the Department of Energy

5/63 Total number of women at the Department of Defense

0 Women of color at the Department of Defense

13/109 Total number of women at the Arms Control Disarmament Agency (1961-1999)

* No racial data

Meeting the needs of all within the planet's means –
Redistributive, regenerative Economy

- Present distribution: *determined by* the relative *power* of workers, landlords and capitalists (**not** by the relative power of productivity

- Means for a transition: the ownership revolution already under way:

- employee-owned companies that raise funds by issuing bonds, promising a fixed rate return rather than a slice of ownership- think: AI COMPANIES!

- Transform industrial design from *degenerative to regenerative,* backed by regulation*:* phase out "red list' chemicals and toxic processes, phase-in life-friendly chemistry only, plus net-zero and net-positive industrial standards

- Change perverse tax laws that charge firms for hiring humans (through payroll taxes), subsidize them for buying robots (through tax-deductible capital investments) and levy next to nothing on the use of land and non-renewable resources.

- Tax passive money and the use of non-renewable resources rather than labor, since taxing employees heads us for a jobless economy

- Subsidize renewable energy and resource-efficient investments

- ALREADY HAPPENING: In 2012, the three hundred largest cooperatives worldwide (covering agriculture, retail, insurance and healthcare) generated $2.2 billion in revenue- *equivalent to the world's seventh largest economy*

Meeting the needs of all within the planet's means –
Redistributive, regenerative AI

- <u>Present distribution</u>: *Winner-takes-all dynamics: individual providers (Google, Youtube, Facebook, eBay, Paypal, Amazon) run global social commons in their own interests, guarding their privilege with patents.*

- *internet created commons are being captured by digital monopolies*

- <u>Means for a transition</u>   SHIFT BACK INTO PARTNERSHIP SOCIETIES

- ***Own the ideas***: create a global commons (already under way) and protect it legally from capture by private interests

- ***Own the robots***: instate a "robot dividend" collected e.g. through royalties from co-owned public-private patents, or through public banks owning significant equity in businesses that use robot technologies made possible by publicly funded research

- ***Form AI companies with redistributive structure***

- ***Invest more on skilling people up where they beat robots*** hands down: in creativity, empathy, insight and human contact. Place AI at *their* service.

- ***Develop AI tools to inform policy makers, help law enacting and law enforcing***

Transformation- Stabilize the pendulum movements into partnership

DOMINATOR   GYLANIC

Jesus's original movmt

Suffragettes

Me-Too

LGBTQ

environmental, e.g. Extinction Rebellion

Nobel Women

….

Humanity is *not* doomed to perpetuate patterns of violence and oppression.

But "to the extent that androcracy remains in place, a just political and economic system, and hence a just AI, remains impossible". Let's make it possible!

Main tasks for a life-supporting AI

- Question dominator models and beliefs, debunk *male dominance, and consequently, the idea of dominance.*

- Reprogram our dominator values into gylanic values:

- To be human is to affirm life, not death

- Acknowledge power as enabling: the power to give and to create

- Change destructive myths, e.g. "hero as a killer", "king" and "warrior" as supposedly "normal" male archetypes

- Replace force-based ranking by linking as our organizational structure

- Aim at sexually and socially equalitarian and peaceful societies

- Bring AI into service:

- Participate in/ promote the numerous AI initiatives towards partnership societies

- Demand the necessary laws and participate in AI systems that support them

STRATEGIES

- *Join forces with women* who well know how to split the cake so no one will fight over it. *Promote proportional gender representation* (key to intelligent decision-making, as well as for success for both sexes).

- Join forces *with all possible gylanic movements*

- Question old narratives about an inherently flawed humanity

- Educate in a hurry (your young ones, students, …) about our holistic human history and its having been detoured by thousands of years of dominator societies.

- Brainstorm: what kinds of social systems support the human capacity for consciousness, caring, and creativity, or conversely for insensitivity, cruelty, and destructiveness

GRACIAS!!

References

1. **K. Raworth (2018)** Doughnut economics- 7 ways to think like a 21rst Century economist

Change the goal into: to meet the needs of all within the means of the planet

https://www.youtube.com/watch?v=Mkg2XMTWV4g

2**. R. Eisler (1987) The Chalice and the Blade: Our history, our future**

K. Millet (1970) Sexual Politics (Garden City, NY: Doubleday)

3**. J. Nelson (2018) Bypassing Dystopia. Watershed Sentinel Books**

4. **Monbiot, G.** Out of the Wreckage**.**

User must define values and goals, e.g.

User-defined Values and their priorities:

    priority(environmental_safety, high).
    priority(human_safety, high).
    priority(transparency, medium).
    transparency: medium
    …

User-defined Goals:

goal(minimum-average-yield,5).
…

Sample KB

<u>Knowledge Base</u>: CHR constraints in the form

criterion(Criterion, Company, Rating, Reasons,URL).

Ex.:

criterion(humanSafety,
    'DowChemical',
    low,
    'because it sells chemicals that damage the human nervous
    system and have been banned from the US for that reason, to third
    world countries that do not yet have protective regulations',
    'http://en.wikipedia.org/wiki
     Dow_Chemical_Company#DBCP').

The system's shell, exemplified

To express that a given Company is ok with respect to a given Criterion and Goal, check:

- if the user's value for that criterion coincides with the value the company gets on that criterion, and
- if the goal is met

```
priority(Criterion,Value), goal(Goal,N)   ==>
    criterion(Criterion,Company,Value,Reason,Link),
    achievedHistorically(Goal,Company,N1),
    N1> N,
    print_reason(Company,Reason,Link) |
                                        ok(Company).
```