



GLOBAL INNOVATION DESIGN

**DIAGRAMING
PRACTISE**

RCA

FERNANDO GALDON



Royal College of Art

This text represents the submission for the degree of Doctor of Philosophy at the Royal College of Art. This copy has been supplied for the purpose of research for private study, on the understanding that it is copyright material, and that no quotation from the thesis may be published without proper acknowledgement.

PhD

DESIGNING TRUST

**EVOLVING MODELS
AND FRAMEWORKS
TOWARDS
PROSPECTIVE
DESIGN FUTURES IN
HIGHLY
AUTOMATED
SYSTEMS**

© Fernando Galdon
Royal College of Art
Kensington Gore
SW2 2EU
London
www.rca.ac.uk



Royal College of Art

A thesis submitted in partial fulfilment of the requirements of the Royal College of Art
for the degree of Doctor of Philosophy

Examined by

Professor / doctor, University of ... ,

Professor / doctor, University of ... ,

on

20 April 2016 at the Royal College of Art in London.

www.rca.ac.uk

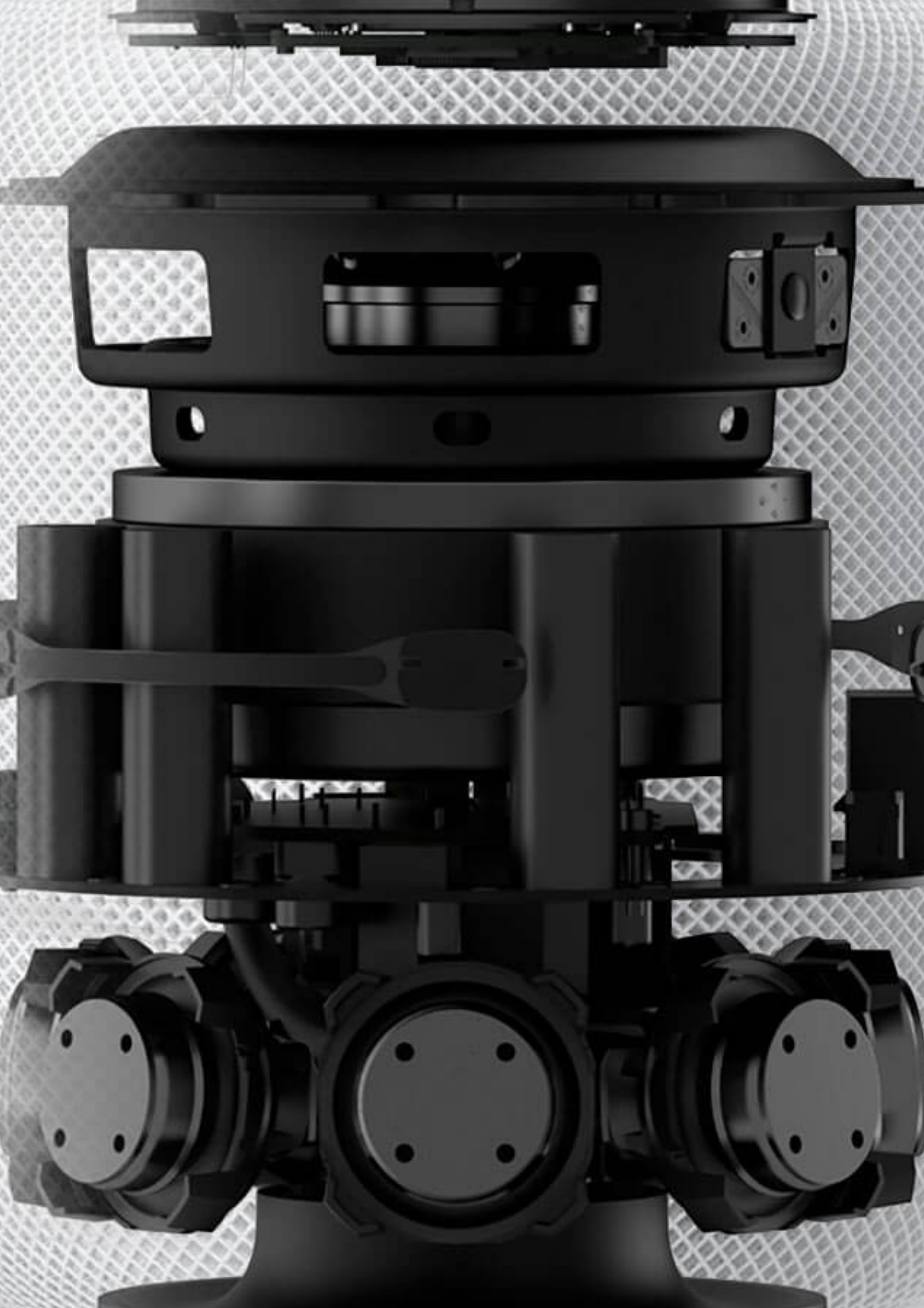
PhD

INTRODUCTION

In this thesis, the author proposes Prospective design as a future-led mixed-methodology to mitigate unintended consequences. This framework combines systems analysis with extrapolations and constructivist perspectives to reconcile confronted models of designing futures in the context of Highly Automated Systems.

It does so by conducting a case on virtual assistants. Although virtual assistants are still in their infancy, they are expected to dominate digital interactions in the coming years. Investigating the prospective developments of this type of interaction device reveals the particular challenges of highly automated interactions for scholarly research. In this context, the intersection between the critical issues of automation and accountability acts as a focal point. Departing from authored multi-dimensional strategies and modes of calculation in ethical computing, this research examines how design decisions affect interactions, how these decisions may be made accessible in design frameworks and how prospective design strategies are better suited to address the rising concerns of these systems. This thesis contributes a new understanding of the implications of designing highly automated systems and provides practical and conceptual means for making this knowledge accessible and usable.

The developmental process consisted of a combination of archive research, surveys, experiments, case studies and co-design workshops. In this process, diagramming became a fundamental tool for practice. Diagrams have been traditionally used in computer science as schematic tools to explain the internal functioning of a system (circuit boards). This approach is translated to explain interactive elements of the system functioning in the context of AI - schematics of interaction -. This technique facilitated the understanding and communication of dematerialised systems. In this process, diagrams also become reflective tools. They helped me to structure knowledge in a manageable way to implement critical analysis via comparative or relational studies. As a synthetic tool, they represent a reduction of reality, but this reduction facilitated understanding. Furthermore, this tool was particularly helpful to facilitate cross-disciplinary enquiry, and this element allowed me to find relationships among disciplines and fields.





Royal College of Art

1

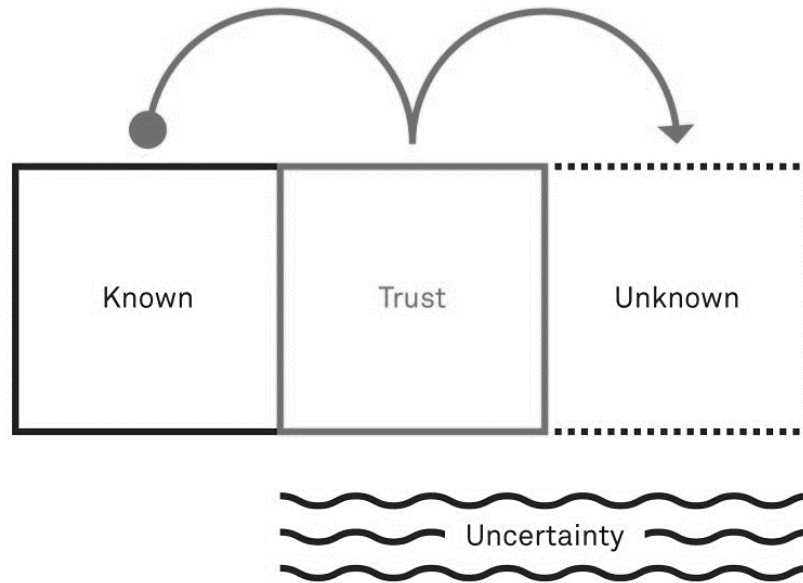
CHAPTER

DIAGRAMMING

TRUST

1.1 WHAT IS TRUST?

According to trust expert Rachel Botsman, trust is “trust is a confident relationship with the unknown” (Botsman, 2018). She builds her definition from social psychologist Morton Deutsch, who wrote in his seminal 1973 book *The Resolution of Conflict*: “Trust involves the delicate juxtaposition of people’s loftiest hopes and aspirations with their deepest worries and darkest fears.” (Botsman, 2018). Nevertheless, is this definition correct?



RACHEL
BOTSMAN

Trust. Rachel Botsman. 2017

According to the *Handbook of Research Methods on Trust*, trust is a multi-faceted phenomenon difficult to underpin (Lyon, 2015). This concept cannot be easily observed or even defined. This topic can be traced back at least to work in the 1960s and 1970s with a range of influential exploratory pieces (such as Deutsch, 1973; Garfinkel, 1967; Rotter, 1967; Zand, 1972). Furthermore, In the 1980s and 1990s, research was implemented on conceptual aspects. It was followed by a wide range of empirical and experimental studies from the late 1990s to the present (see Bachmann and Zaheer, 2006; Möllering, 2006). Seppanen et al. (2007) presented an excess of 70 definitions of the concept of trust (see also Castaldo, 2007). Trust is a very elusive concept. It is placed in the middle of two entities interacting. It is multi-level. According to Blobaum, trust is referential to a system (media), an organisation (newspaper), a person (journalist) or a product (article) (Blobaum, 2016 pp.8-9). You can trust and distrust at the same time, one or more elements in the hierarchy. For instance, you can distrust the media, and the guardian, but can trust a specific journalist and the content he creates for the newspaper. You can trust it today and distrust the same thing tomorrow. Trust is constituted in an act in which the trustor makes himself or herself

vulnerable to the trustee. One of the most cited definitions of trust in the organisational science literature is Mayer et al. They define trust as “The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party. (Mayer et al., 1995, p. 712)

This point is capital and a missing point in Botsman definition. Trust is very fragile and can be easily broken. Preliminary knowledge, past experiences, personality and personal situation, affect trust from a trustee perspective. Positive experiences reinforce trust. In this regard, the work of Giddens defines trust as “confidence in the reliability of a person or system, regarding a given set of outcomes or events, where that confidence expresses a faith in the probity or love of another, or in the correctness of abstract principles (technical knowledge).” (Giddens on Lane and Bachman, 1998, pp.35). It is the “Belief in someone or something, which is nurtured through positive experiences”. Media, K. (2016).

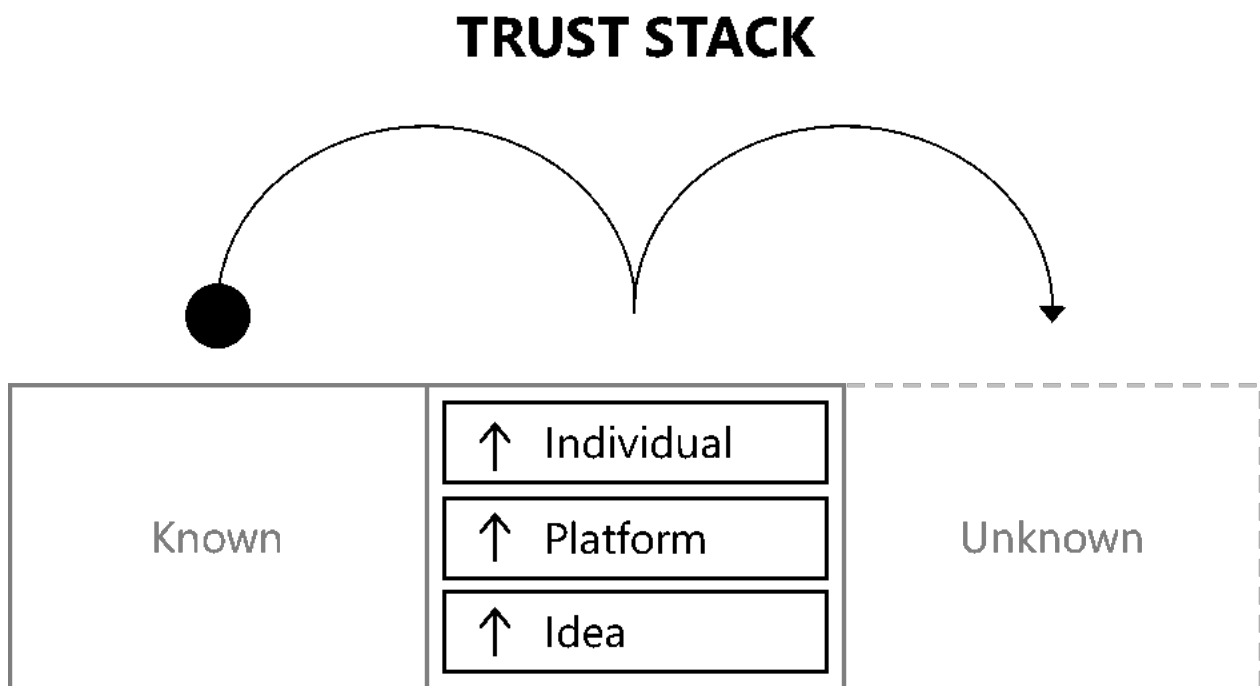
In the technological context we are living, evangelists make their promises through technologies. In this scenario, Botsman states that “Technology is only the means” (Botsman, 2018). However, Sophia Jasanoff, a Harvard professor, presents in her seminal book, the ethics of invention, quite a contrary approach. She builds from examples of animals’ behaviour to demonstrate that technology as an end to satisfy needs is not even a distinctively human action (Jasanoff, 2016. P. 2--)

For the author, “Technology, in short, is not merely about achieving ends that we already foresee but an open the door to an uncharted, often uncertain future where current social understandings and practices may be fundamentally transformed. Uncertainty, moreover, can deter as much as it entices. The bright gleams of promise that invite human societies to invest in technology march hand in hand with darker misgivings about what could go wrong if the promises fail and the unexpected breakdown happens on a grand scale.” (Jasanoff, 2016. P. 214). And concludes, “Neither practicality nor predictability captures the evolving relations between human beings and their technologies. Human technological wizardry extends far beyond performances of repetitious tasks to serve simple, predetermined purposes. Artistry, imagination, and the desire to probe the unknown have long dominated the will to make and use technology” (Jasanoff, 2016. P, 212). Therefore, I will redefine trust as confidence in somebody’s promise to navigate the unknown through design. It is structured in four steps; Fear is created by uncertainty. Trust is created when somebody promises you how to navigate (which leads to hope). A promise is enabled by design. And trust is reinforced when the promise is delivered.

Rachel is right in pointing to the unknown but misses the void created by this uncertainty, which is always occupied by somebody through the delivery of a promise to navigate this terrain. What is weird is that she makes references to it. For instance, on page 124, where she states that “we need to think of trust as trusting someone to do something”. (Botsman, 2017. P.124).

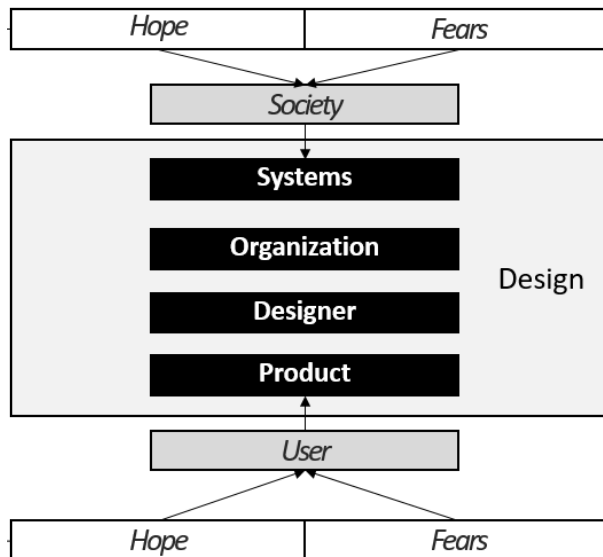
1.2 HOW DO WE STRUCTURE TRUST?

As we have seen earlier, trust is multi-level. Botsman presents a trust stack as a model to understand how trust works in the digital domain. She fills the trust void with three levels; an idea, a platform and an individual. These elements are the foundation for a user to take a trust leap.



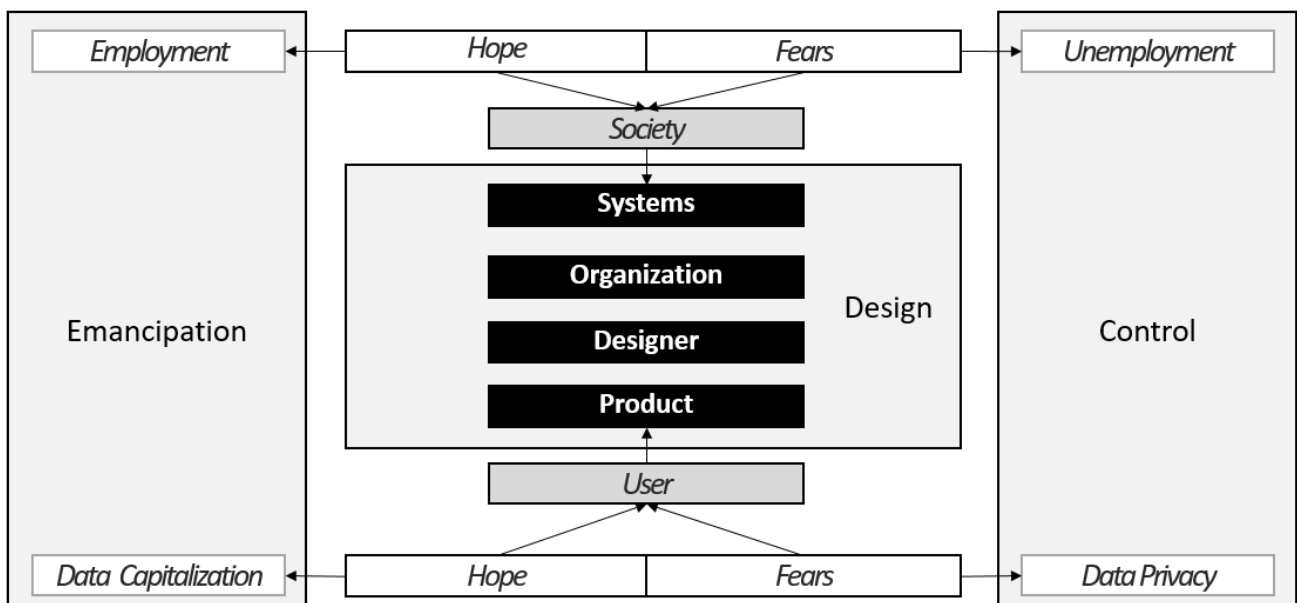
Type to enter a caption.

However, according to Blobaum, trust is referential to a system (media), an organization (newspaper), a person (journalist) or a product (article) (Blobaum, 2016 pp.8-9). For instance, system (media), an organization (newspaper), a person (journalist) or a product (article). Or system (e-commerce), an organization (Amazon), a person (search algorithm) or a product (book). This approach differs from Botsman with the integration of the system as a level to consider. Again, she makes references on page 188, but do not integrate these elements in the stack system presented. Besides, and building from the definition provided earlier, the dynamics and actions of the ecosystem are also missing. We need to integrate the drivers; fear and hope. And the enabler; design.



Type to enter a caption.

However, which are the driver? Sir Martin Sorrell, CEO of WPP, the most significant advertising and public relations company in the world, stated at the world economic forum, that the most prominent drivers rising fear are unemployment and privacy. These results were based on a range of surveys conducted by his company. Due to the contextual nature of these drivers, likely changing over time, in this area, further research is needed to identify the essential surveys to establish a cross-referenced output. By playing the opposite, we could build a model for hope where fears of unemployment and data privacy could be transformed into employment and data capitalisation. Thus, leading to freedom and emancipation, instead of precariat and control.



Type to enter a caption.

Botsman structure trust around the dimensions of competence, reliability and honesty. Is this true? In this area, due to a lack of clear distinction amongst the factors that constitute trust, trust itself, and the outcomes of trust, its research has been complicated. The main model from which all contemporary research underpins is Mayer's dimensional model. Who after an extensive revision on the topic, proposed a generic typology on the subject of enquiry consisting fundamentally on three dimensions; ability, benevolence, and integrity. (Mayer, Davis & Schoorman, 1995)

They are conceptually distinct since they address different elements of cognitive and affective abstraction of trust. However, collectively, they represent a comprehensive multi-dimensional space for trust. Their multi-dimensional model of trust is one of the most widely accepted (Rousseau, Sitkin, Burt & Camerer, 1998; Wasti et al., 2007). In Mayer's model, three dimensions underpin the process of trust.

Ability - This area refers to "the trustor's perception of trustee's competencies and knowledge salient to the expected behaviour". They can be based on "prior (first-hand or second-hand) experience or institutional endorsements". (Bhattacharjee, 2014)

Integrity – this area refers to the perception a trustor will follow a set of principles or rules.

Benevolence – this area refers to the intentionality and behaviour of the trustee. From who is believed to "intend doing good to the trustor, beyond its own profit motives" (Bhattacharjee, 2014)

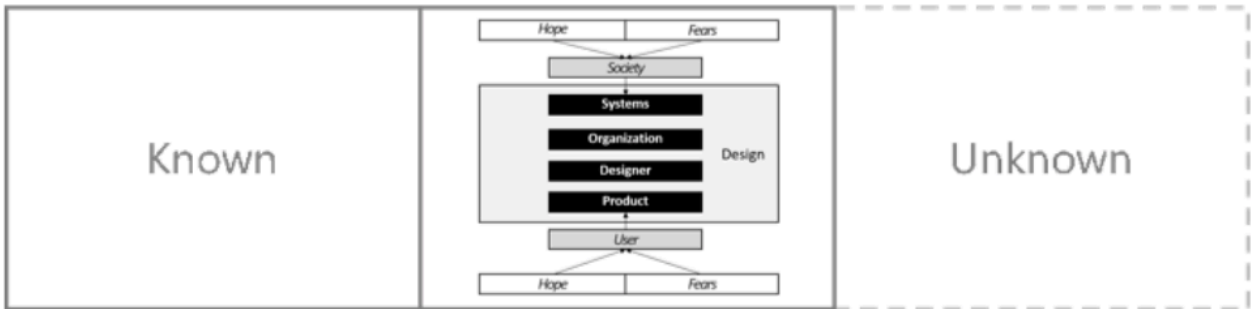
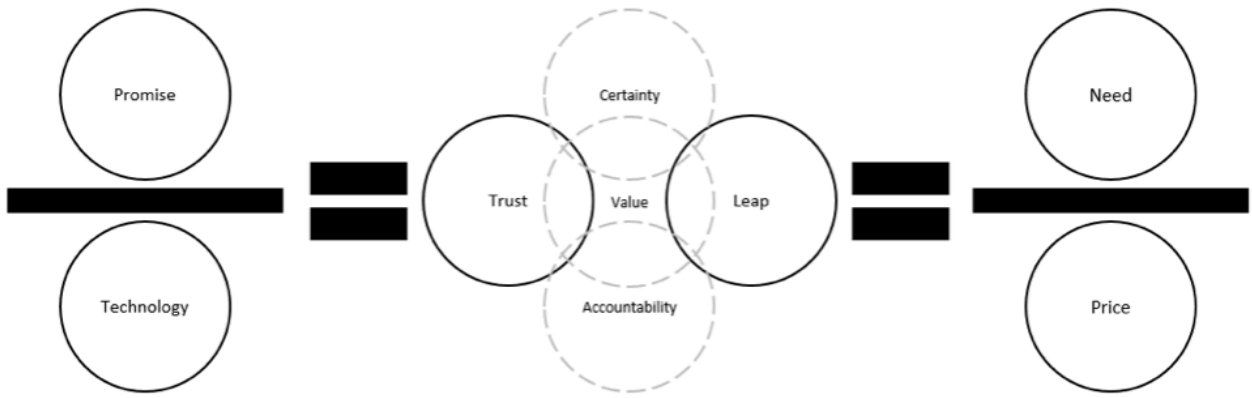
If we translate into Botsman model, we could pair ability with competence, integrity with reliability and honesty with benevolence. Therefore, positioning Botsman as correct.

Finally, Botsman presents value, accountability and certainty as to the main dimensions of trust.

From what we have seen, trust is confidence in somebody's promise to navigate the unknown. It is structured in four steps; Fear is created by uncertainty. Trust is created when somebody promises you how to navigate (which leads to hope). A promise is enabled by design. And trust is reinforced when the promise is delivered.

In this process, the key aspect is enabling a promise, which we do through design. This step is structured in four levels; systems, organisations, designers and products. The development of technology is affected by top-down dynamics coming from society and bottom-up dynamics coming from the individual. And two drivers; hope and fear.

A fundamental area is how and why we take a trust leap. In this context, certainty (reliability or integrity) value (competence or ability) and accountability (honesty or benevolence) play a fundamental role in the final decision. However, an increase in the need or competitiveness of price can overrule a reflective analysis and increase instinctual decisions.



Final model



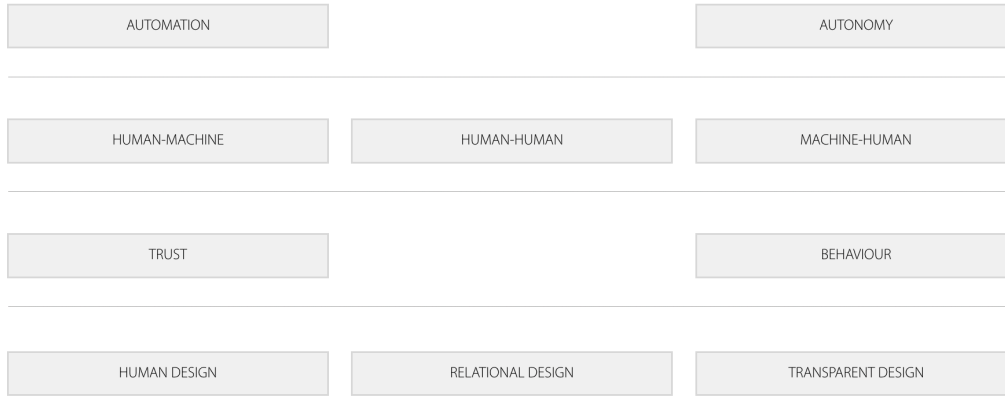
Royal College of Art

2

CHAPTER

DIAGRAMMING

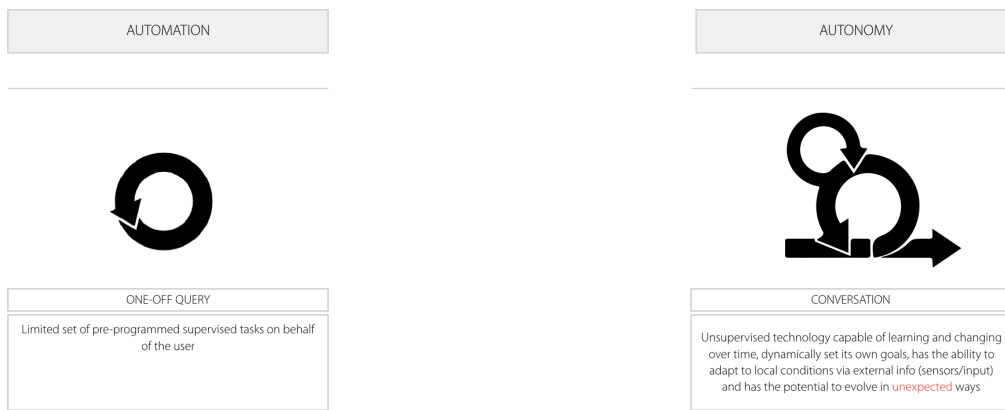
AUTOMATION



2018

Slide / 7

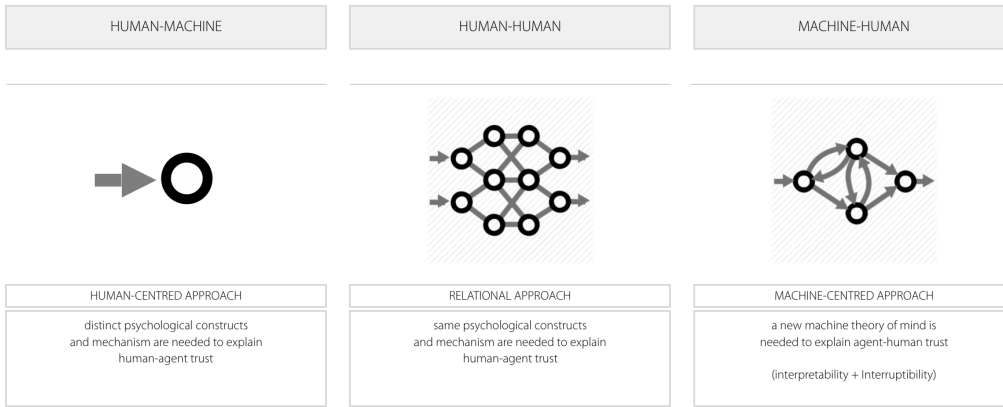
Type to enter a caption.



- de Visser, E (2018). From 'automation' to 'autonomy': the importance of trust repair in human-machine interaction. Ergonomics 2018 vol: 0139 pp: 1-19

Slide / 8

Type to enter a caption.



2018

Slide / 10

Type to enter a caption.

Research

Slide / 11


Machine problems

Trust design

Building safe artificial intelligence: specification, robustness, and assurance.

By Pedro A. Ortega, Vishal Maini


<https://medium.com/@deepmindsafetyresearch/building-safe-artificial-intelligence-52f5f75058f1>

 Royal College of Art

Specification (Define purpose of the system)	Robustness (Design system to withstand perturbations)	Assurance (Monitor and control system activity)
Design Bugs & inconsistencies Ambiguities Side-effects High-level specification languages Preference learning Design protocols	Prevention and Risk Risk sensitivity Uncertainty estimates Safety margins Safe exploration Cautious generalisation Verification Adversaries	Monitoring Interpretability Behavioural screening Activity traces Estimates of causal influence Machine theory of mind Tripwires & honeypots
Emergent Wireheading Delusions Metalearning and sub-agents Detecting emergent behaviour	Recovery and Stability Instability Error-correction Failsafe mechanisms Distributional shift Graceful degradation	Enforcement Interruptibility Boxing Authorisation system Encryption Human override
Theory (Modelling and understanding AI systems)		

Type to enter a caption.

TRUST



CONSCIOUS PROCESS

understanding human behaviour is key for **HAI**
- a conscious decision
- method = calibration

BEHAVIOUR



SUBCONSCIOUS PROCESS

understanding decision-making is key for **HAI**
- a sub-conscious decision -
- physiological correlates = method -

- Drnek K (2016) From Trust in Automation to Decision Neuroscience. *Frontiers in Human Neuroscience* 2016. Vol: 10 (June) pp: 1-14

Slide / 13

Type to enter a caption.


Research

Slide / 14

Behaviour Problem

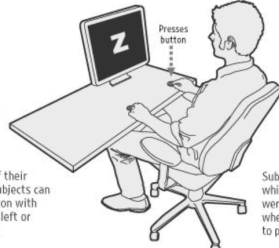
Trust design

- Soon, C.S., Brass, M., **Heinecke, H.**, & Haynes, J. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11, 543-545.



Act First, Think Later

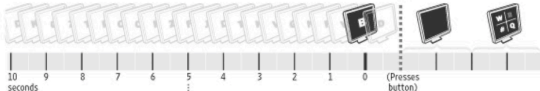
By scanning the brains of people performing simple decision-making exercises, scientists found that brain regions involved in making choices activate before people are consciously aware they've made a choice



Subjects identify which letter they were looking at when they **decided** to push the button

On the screen
Subjects watch a screen that flashes a random sequence of letters at half-second intervals

At a time of their choosing, subjects can press a button with either their left or right hands



Beneath the surface
Throughout the process, scientists are recording the subjects' brain activity

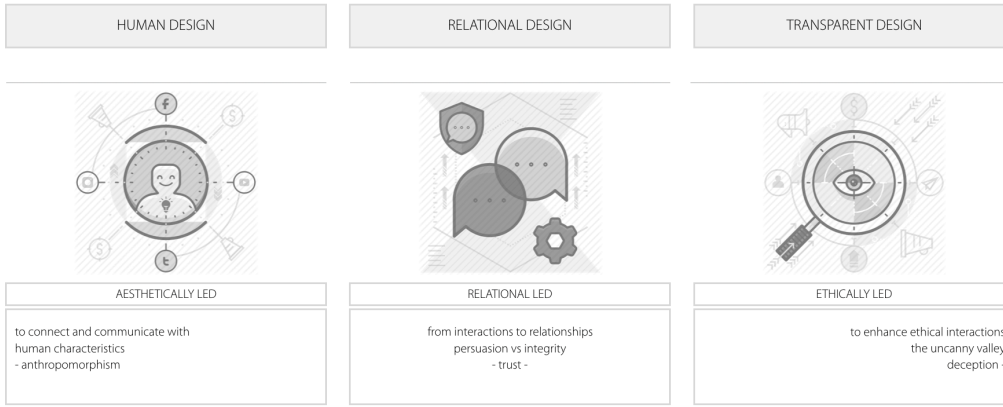
They found that regions involved in decision making became active up to 10 seconds before the subjects consciously decided to press the button

Motor cortex

They also found that the motor cortexes became active five seconds before deciding to press the button. The brain scans also allowed them to predict whether subjects used their left or right hand.

Source: Nature Neuroscience

Type to enter a caption.



2018 - Hoff and Bashir's (2015) model illustrating factors that affect automation trust

Type to enter a caption.


Research

Slide / 17

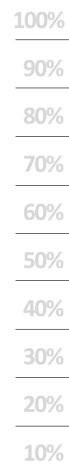
Design Problem

Trust design

- [Driscoll, K.](#), [Marathe, A. R.](#), [Lukos, J. R.](#), Metcalfe, J. S. (2016). From Trust in Automation to Decision Neuroscience: Applying Cognitive Neuroscience Methods to Understand and Improve Interaction Decisions Involved in Human Automation Interaction

 Royal College of Art

Success rate



30%

Overuse > user stops monitoring

70%

Disuse > user stops using the device

Failure rate

Type to enter a caption.


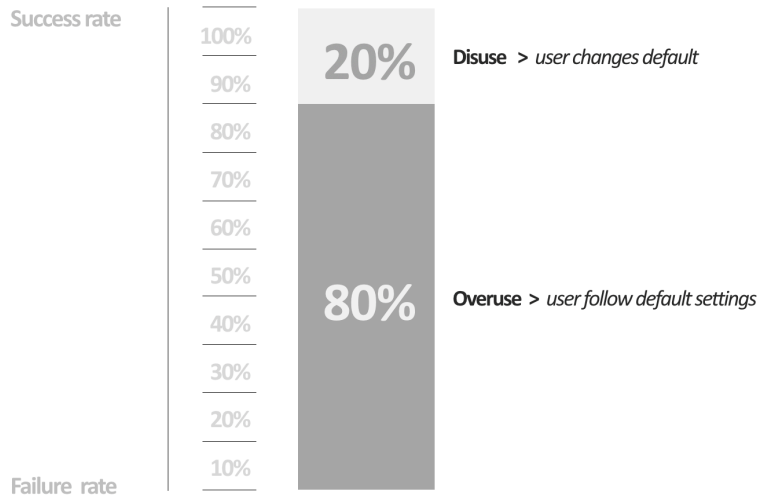
Research

Slide / 18

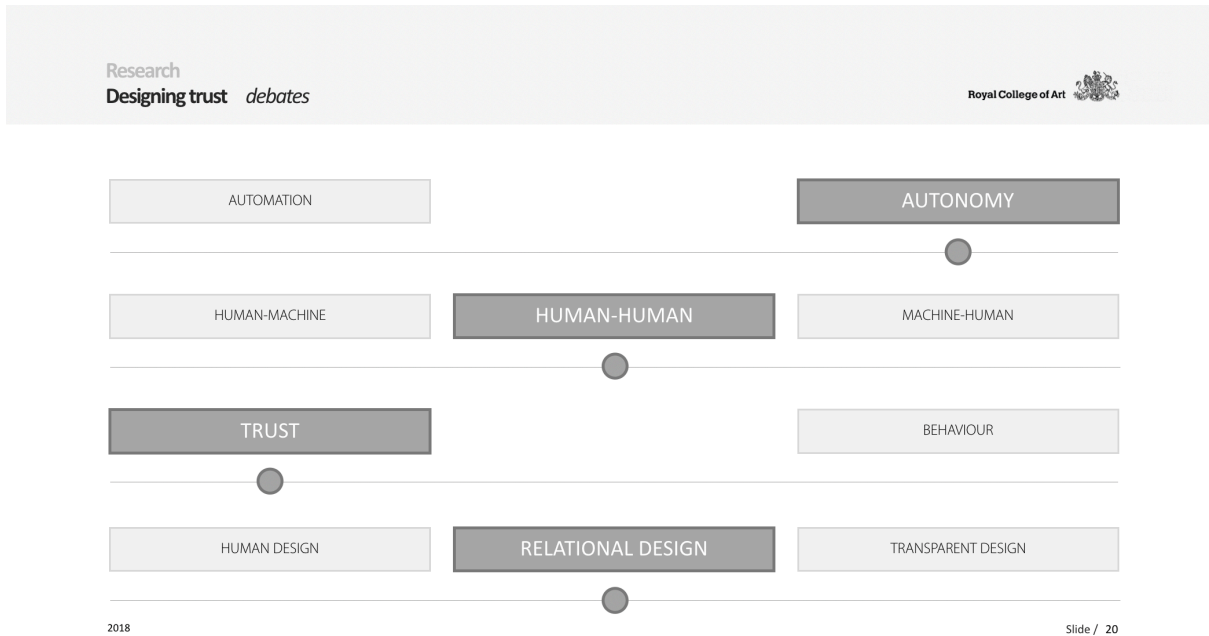
Design Problem

Trust design

- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Type to enter a caption.



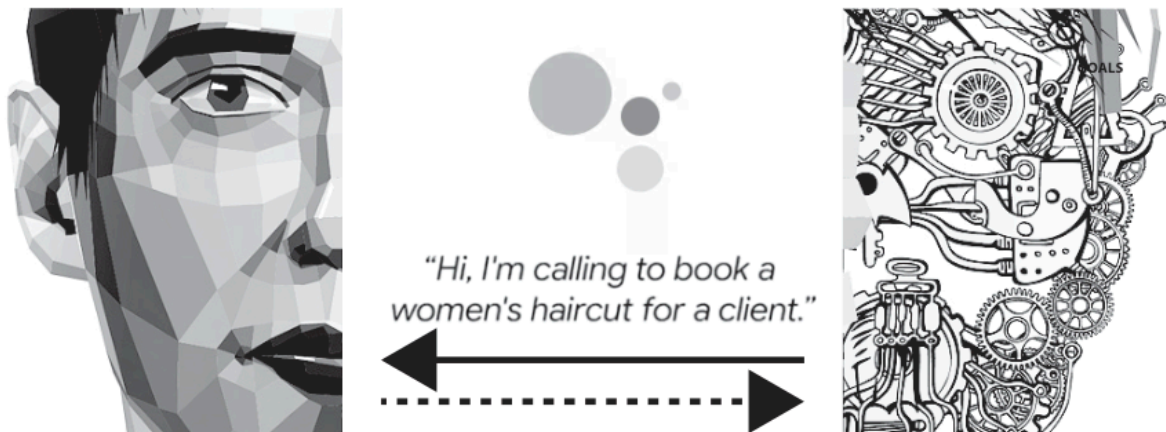
Type to enter a caption.

ONE-OFF COMMAND INITIATED BY THE USER



Limited set of pre-programmed supervised tasks on behalf of the user

A CONVERSATION INITIATED BY THE SYSTEM

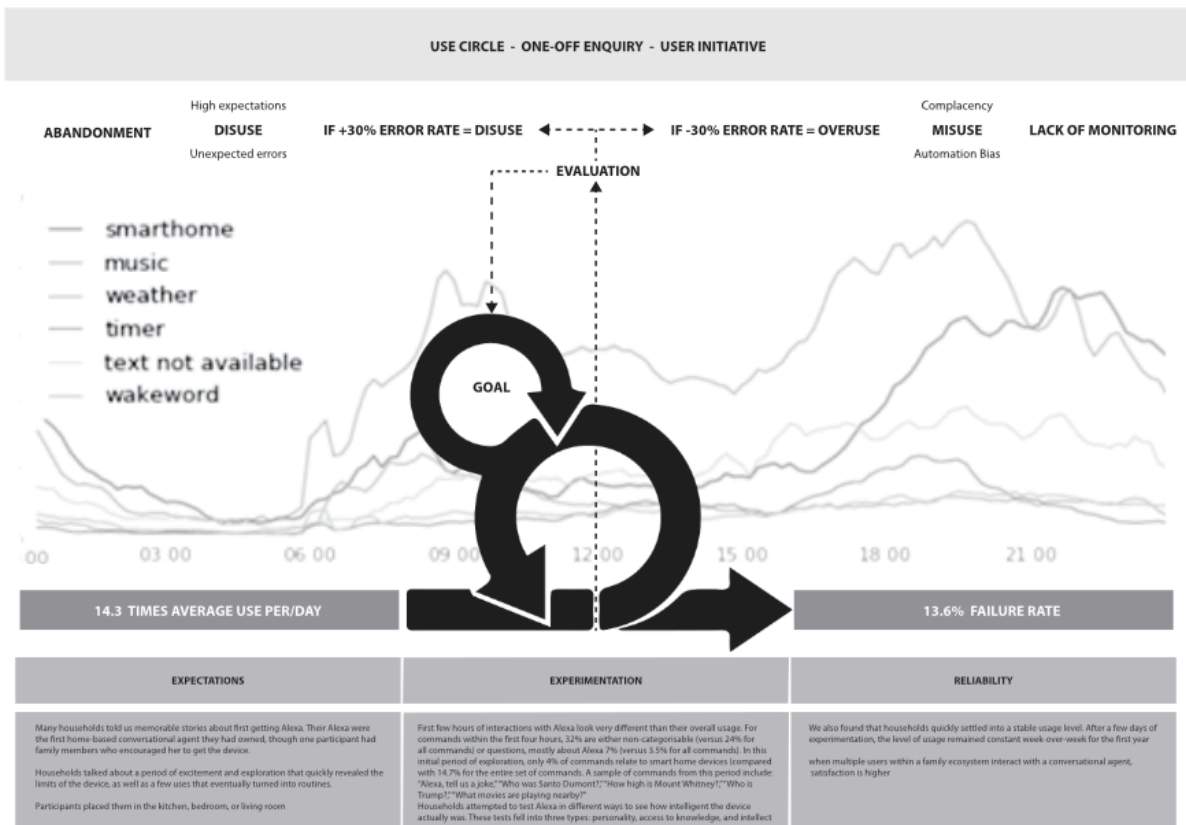


Technology designed to carry out a user's goals without supervision

capable of learning and changing over time, dynamically set its own goals, has the ability to adapt to local conditions via external info (sensors/input) and has the potential to evolve in unexpected ways

Type to enter a caption.

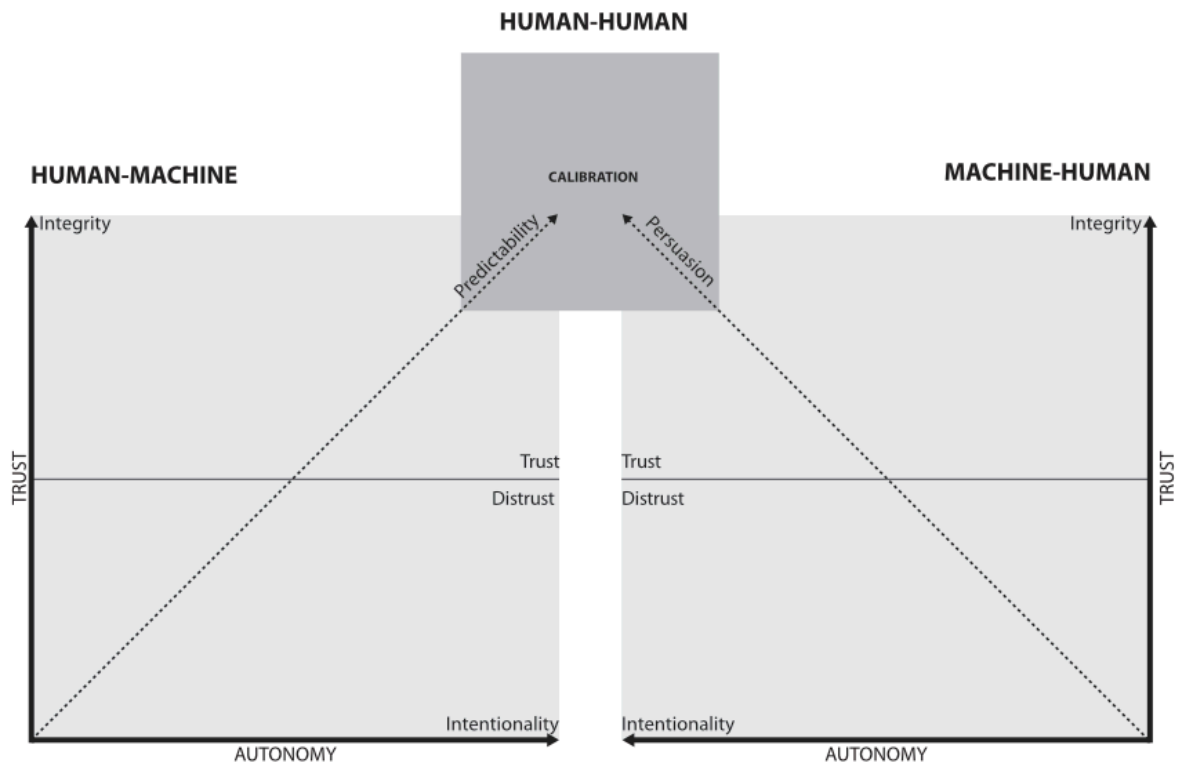
(Wickens & Dixon, 2007)



DESIGN STRATEGIES

PRINCIPLE 5 Desensitisation	PRINCIPLE 5 Desensitisation	PRINCIPLE 5 transparency Communicating intent Desensitisation		
PRINCIPLE 4 Certification	PRINCIPLE 4 Certification	PRINCIPLE 4 Calibration Certification		
PRINCIPLE 3 Recommendation Inner circle Recommendation	PRINCIPLE 3 Recommendation	PRINCIPLE 3 Recommendation		
PRINCIPLE 2 Pre-cognitive	PRINCIPLE 2 Pre-cognitive fluidity and autonomy: pitch and porosity intonation and wording sounds comfortable and natural	PRINCIPLE 2 Pre-cognitive Politeness		
PRINCIPLE 1 Familiarity Typological design - Speakers as voice assistants: - Reminiscent of a vase (neutral colours) - Reminiscent of a Speaker (black) - Reminiscent of a Gadget (black and blue lights) Female Name	PRINCIPLE 1 Familiarity specific accents - localism	PRINCIPLE 1 Familiarity Apologies		
ACCEPTATION Designing for intimacy - Voice driven	CLARITY Designing for usability - Technically driven	PROTECTION Designing for safety - Safety driven		
APPEARANCE	EASE-OF-USE	COMMUNICATION STYLE	TRANSPARENCY-FEEDBACK	LEVELS OF CONTROL
Familiar look remediation is a design concept that refers to when an aspect of an older media is somehow present in a newer media, and it exists aesthetically, culturally, and experientially in different forms of media and technology. A popular example of remediation can be seen in the first personal computer, which was an entirely new invention, but imitated or mimicked the typewriter and television (Levinson, P. 1997, pp. 105-114)		Female Voice According to psychologist Albert Mehrabian 38% of first impressions depend on the tone of voice, and only 7% on the words used. This includes using silences between sentences so that the person listening has time to process the information given, and having fillers just like "mmhm" or "well", all to create the feeling of talking to a human being and therefore developing trust (Perez, 2018) Telefonica		

Type to enter a caption.



PRINCIPLE	5	DECENTRALISATION
PRINCIPLE	4	CERTIFICATIONS
PRINCIPLE	3	RECOMMENDATIONS
PRINCIPLE	2	PRE-COGNITIVE ELEMENTS
PRINCIPLE	1	FAMILIARITY

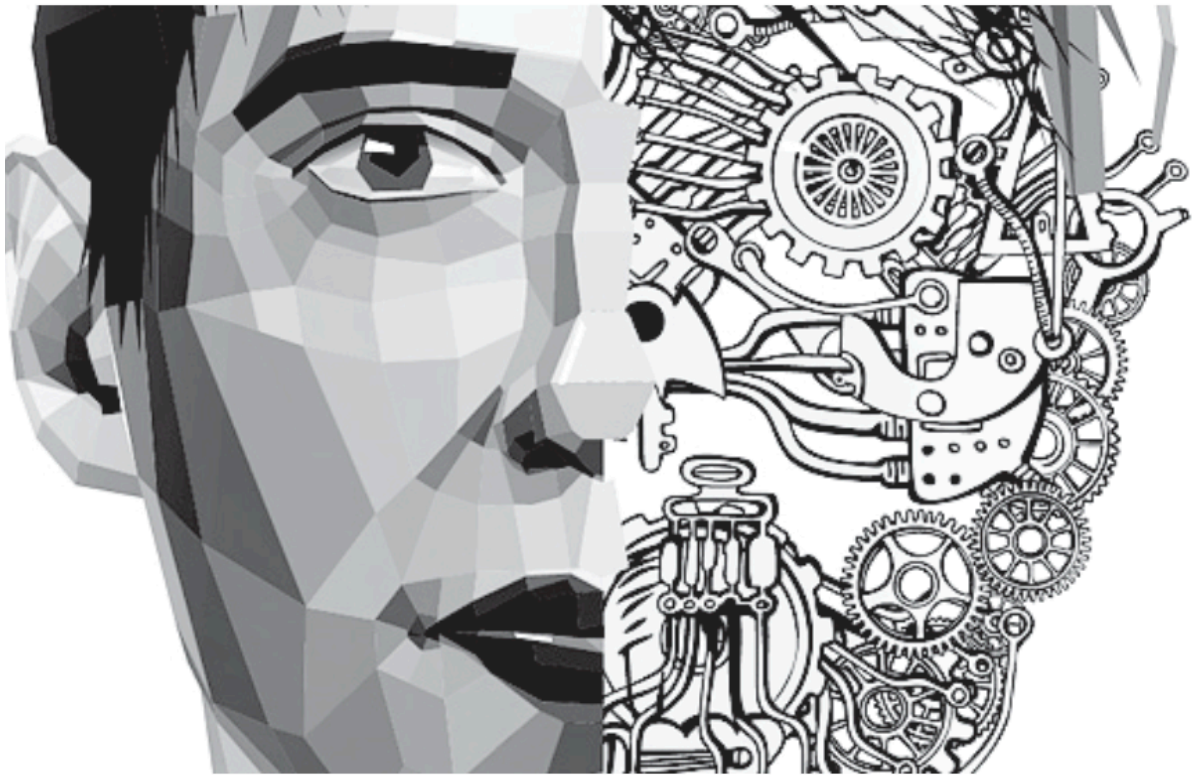
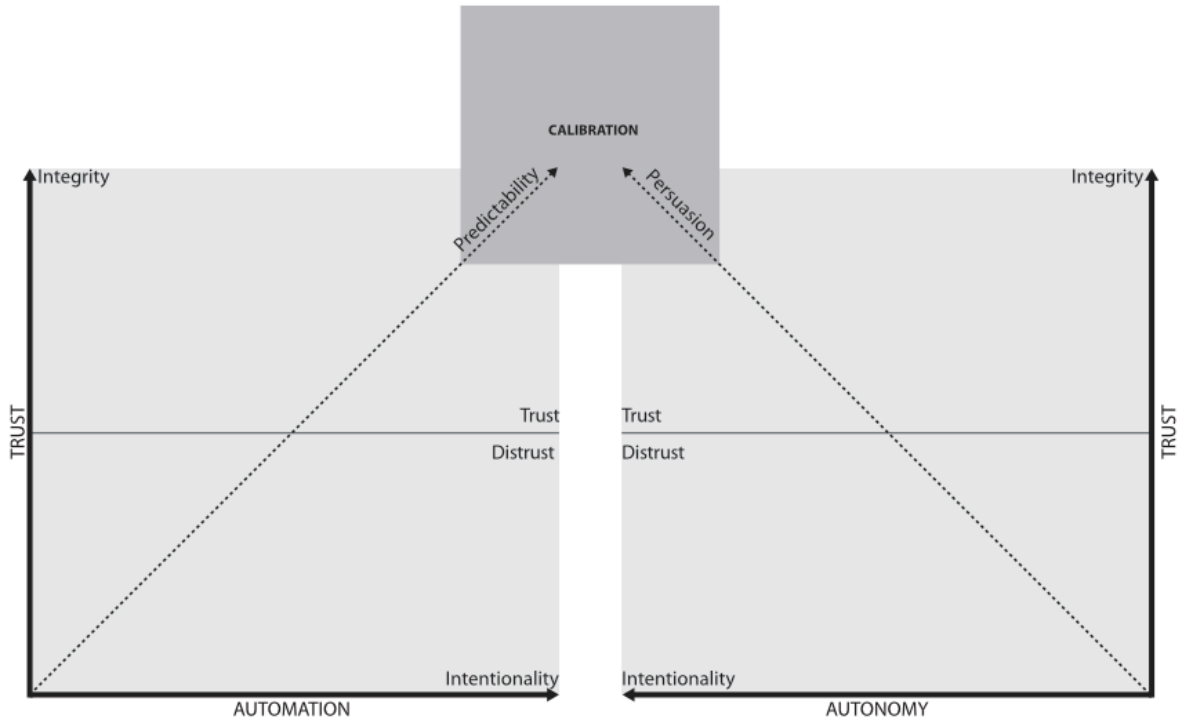
PRINCIPLE	5	LEVELS OF CONTROL
PRINCIPLE	4	TRANSPARENCY-FEEDBACK
PRINCIPLE	3	EASE-OF-USE
PRINCIPLE	2	COMMUNICATION STYLE
PRINCIPLE	1	APPEARANCE

Hoff and Bashir's (2015) model illustrating factors that affect automation trust.

Type to enter a caption.

HUMAN-MACHINE

MACHINE-HUMAN



Type to enter a caption.

GOOGLE USER PROFILE

PREFERENCES

SEARCH HISTORY

LOCATION HISTORY

YOUTUBE HISTORY

TRENDS HISTORY

Type to enter a caption.

I KNOW ABOUT GOOGLE

**35
Bytes**

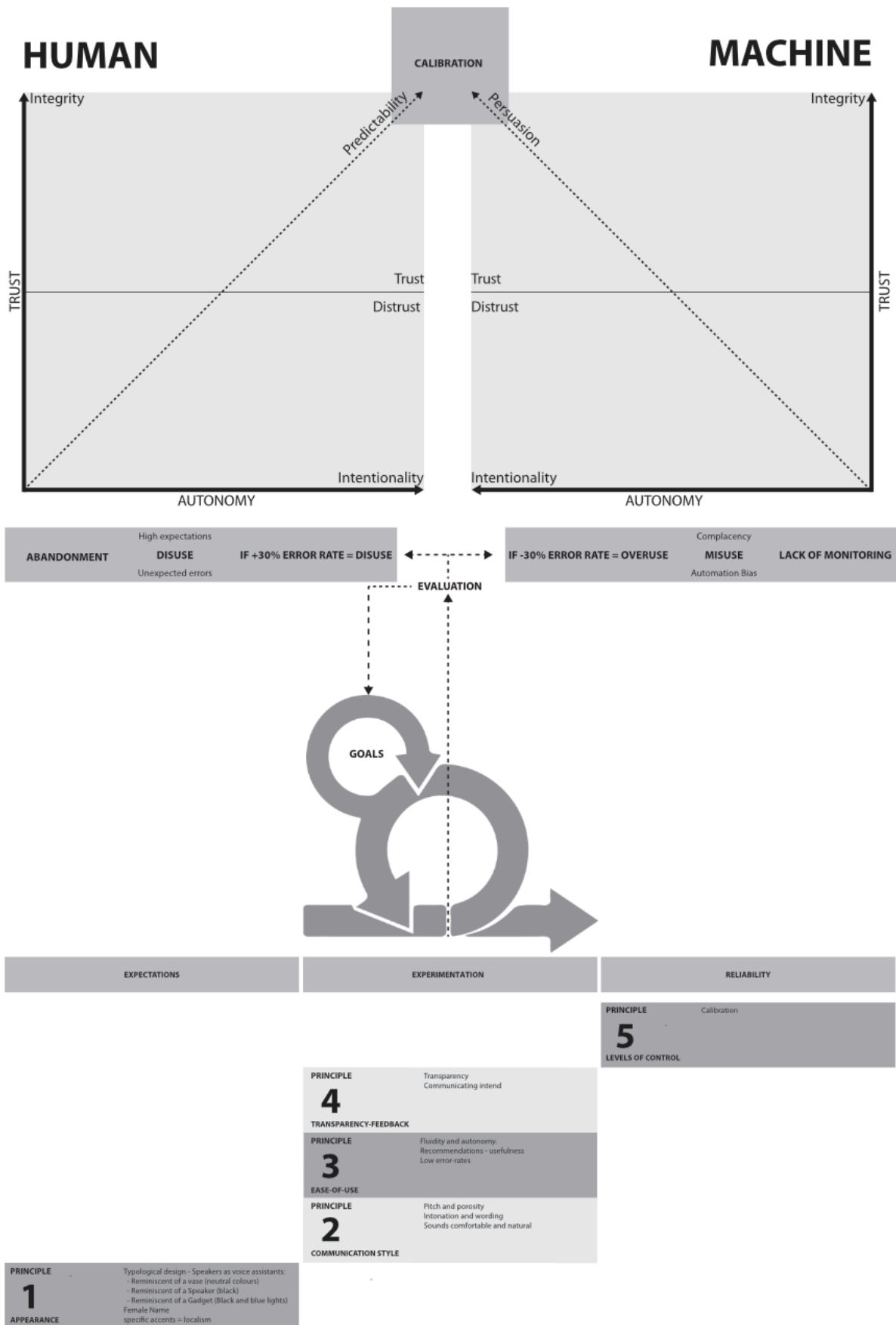
1 Documents

GOOGLE KNOWS ABOUT MYSELF

**55.000.000.000
Bytes**

3.000.000 Documents

Type to enter a caption.



Type to enter a caption.



LEVELS OF AUTONOMY

LEVEL 1	NO AUTONOMY The VA does not implement the action unless requested by the user The user request an action.
LEVEL 2	ASSISTANCE The VA assists determining a range of options related to user's query. The user request an action.
LEVEL 3	PARTIAL AUTONOMY The VA engage in conversation and suggests one option. The user doesn't need to follow.
LEVEL 4	CONDITIONAL AUTONOMY The VA selects action and implements it if human approves. The user needs to respond to requests to confirm.
LEVEL 5	RELATIONAL AUTONOMY The VA selects action, informs human with plenty of time to stop. The user sets time-frame to respond to requests to intervene.
LEVEL 6	HIGH AUTONOMY The VA can perform decisions solely on its own and necessarily tells human what it did The user does not need to respond to requests for the VA to intervene.. The VA must always report
LEVEL 7	FULL AUTONOMY The VA can perform decisions solely on its own without reporting to the user. The user does not need to respond to requests for the VA to intervene.

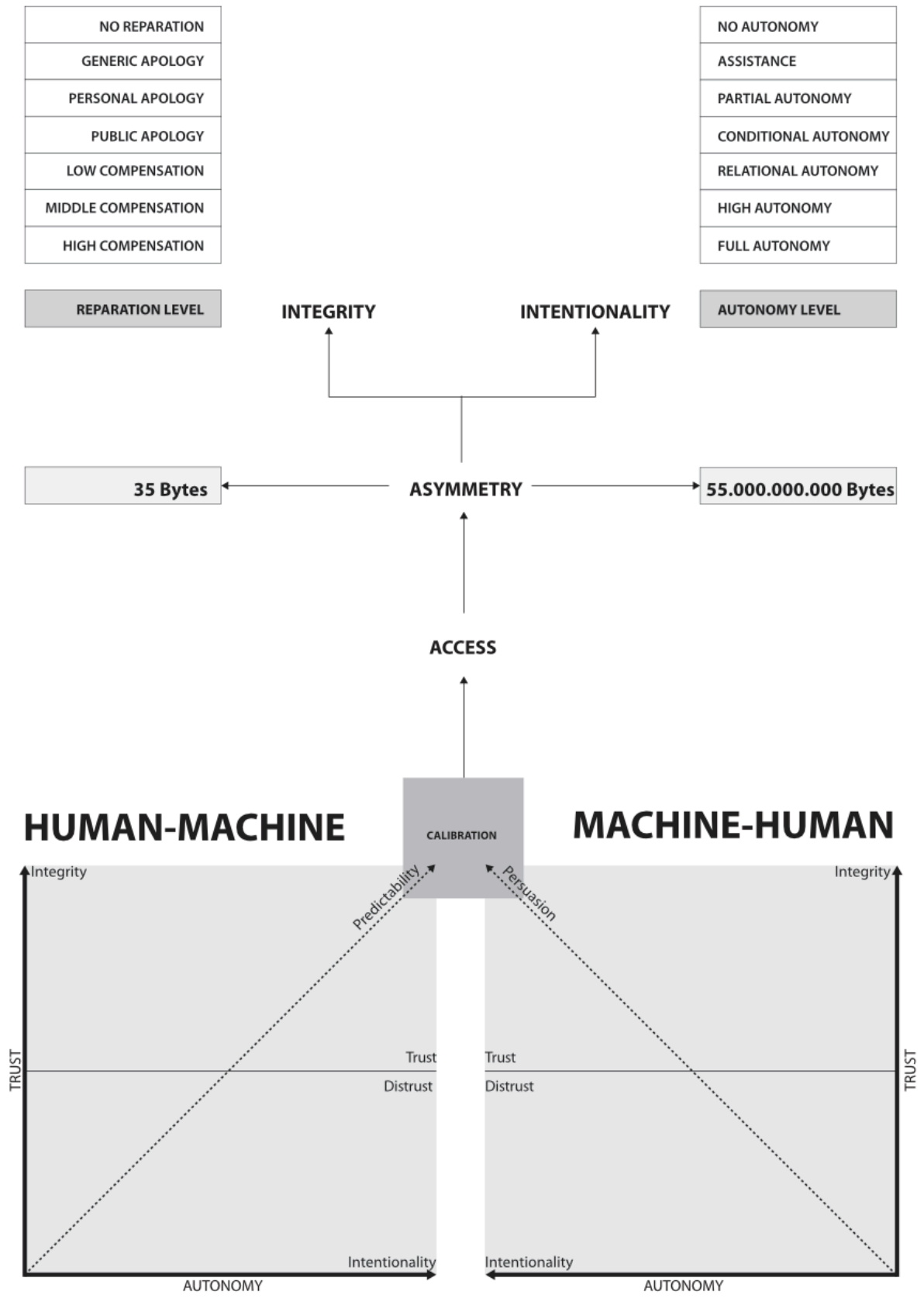
Type to enter a caption.



LEVELS OF REPARATION

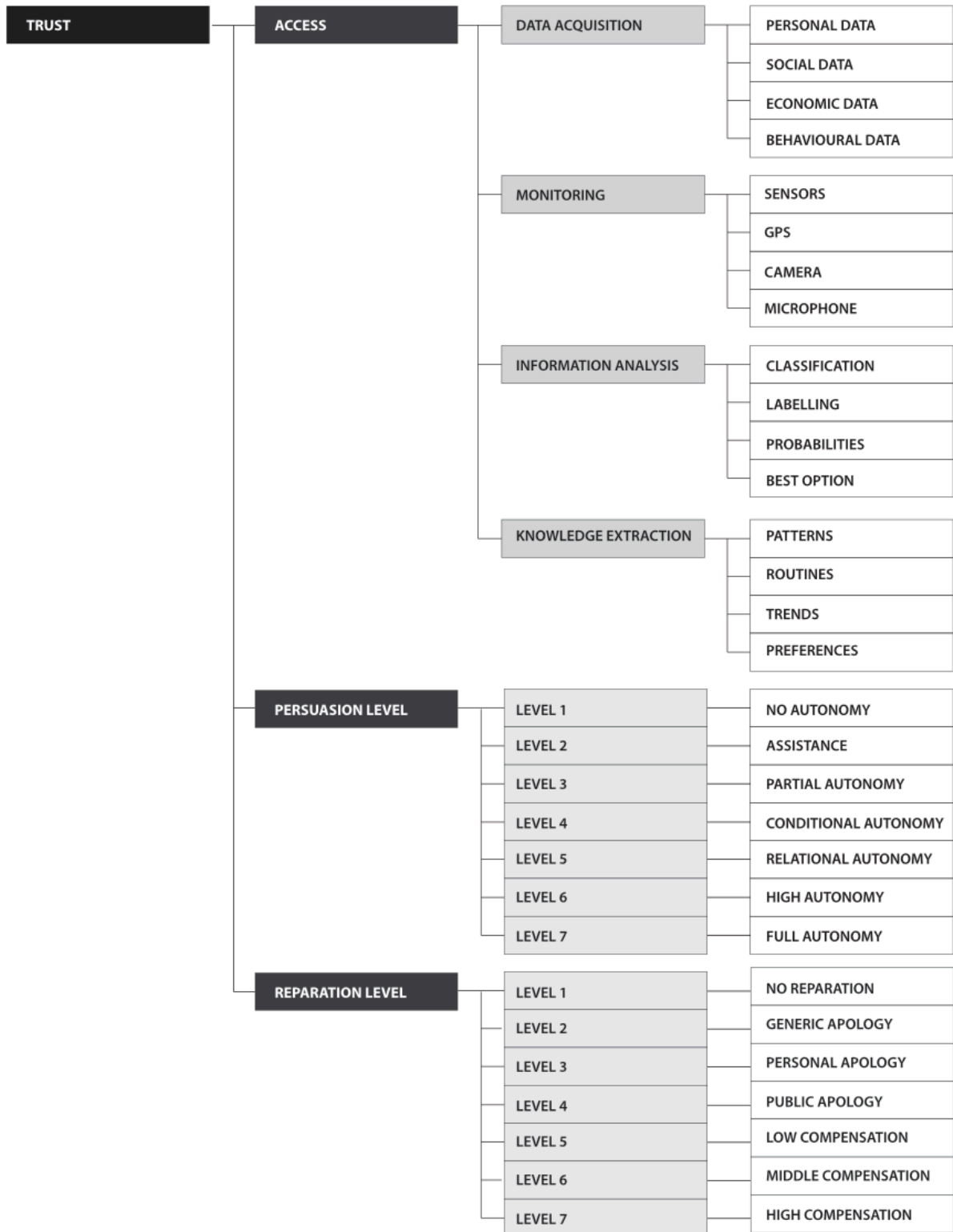
LEVEL 1	NO REPARATION
	The action performed by the VA causes no side effects
	Routinary activities without side effects
LEVEL 2	GENERIC APOLOGY
	The action performed by the VA causes minor discomfort
	A generic apology acknowledging the error
LEVEL 3	PERSONAL APOLOGY
	The action performed by the VA causes discomfort
	A personal apology acknowledging the error
LEVEL 4	PUBLIC APOLOGY
	The action performed by the VA causes offence
	A press release acknowledging the error
LEVEL 5	LOW COMPENSATION
	The action performed by the VA causes harm
	Legal action - Monetary compensation - Thousands
LEVEL 6	MIDDLE COMPENSATION
	The action performed by the VA causes injuries
	Legal action - Monetary compensation - Hundred of thousands
LEVEL 7	HIGH COMPENSATION
	The action performed by the VA causes death
	Legal action - Monetary compensation - Millions

Type to enter a caption.

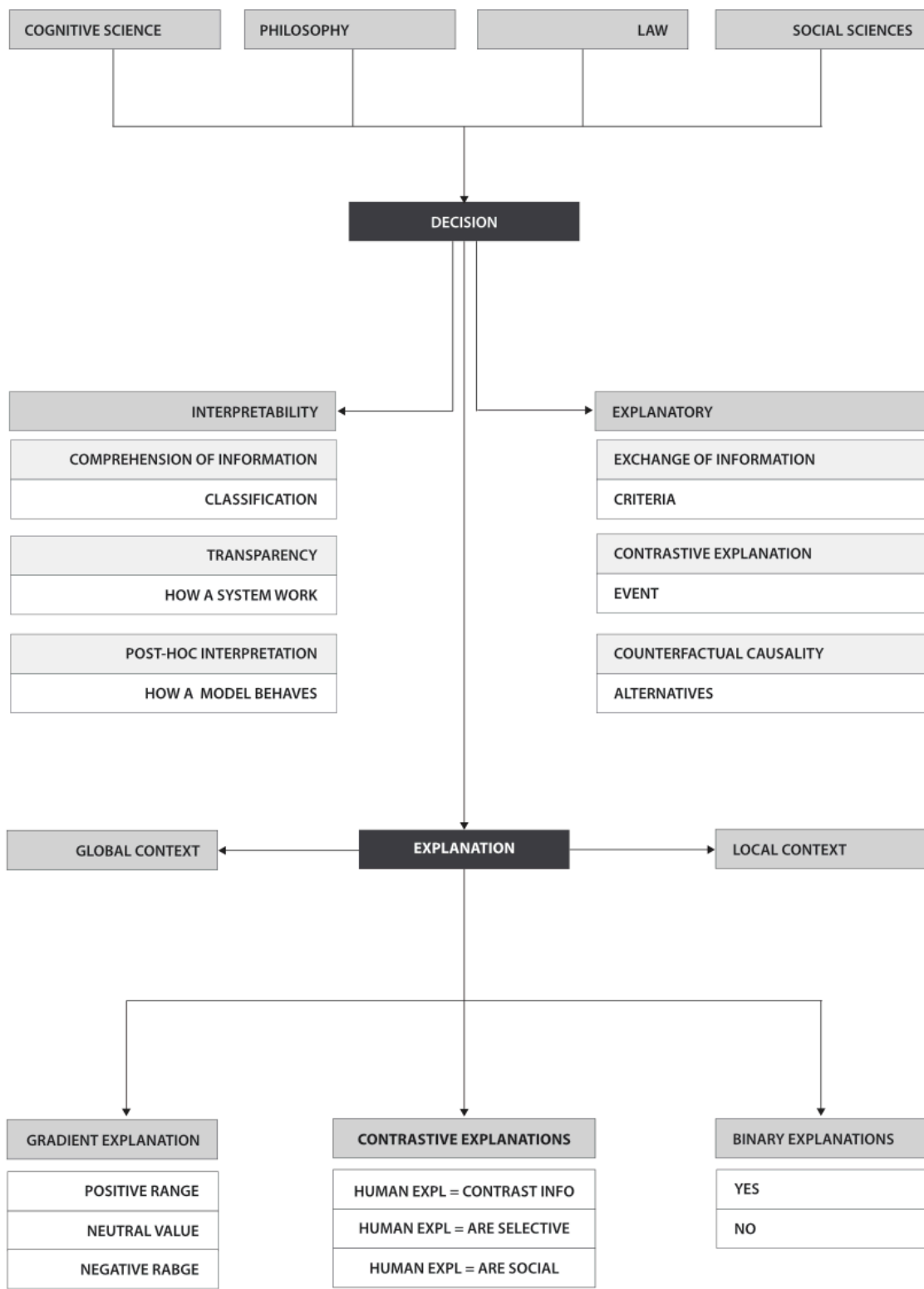


Type to enter a caption.

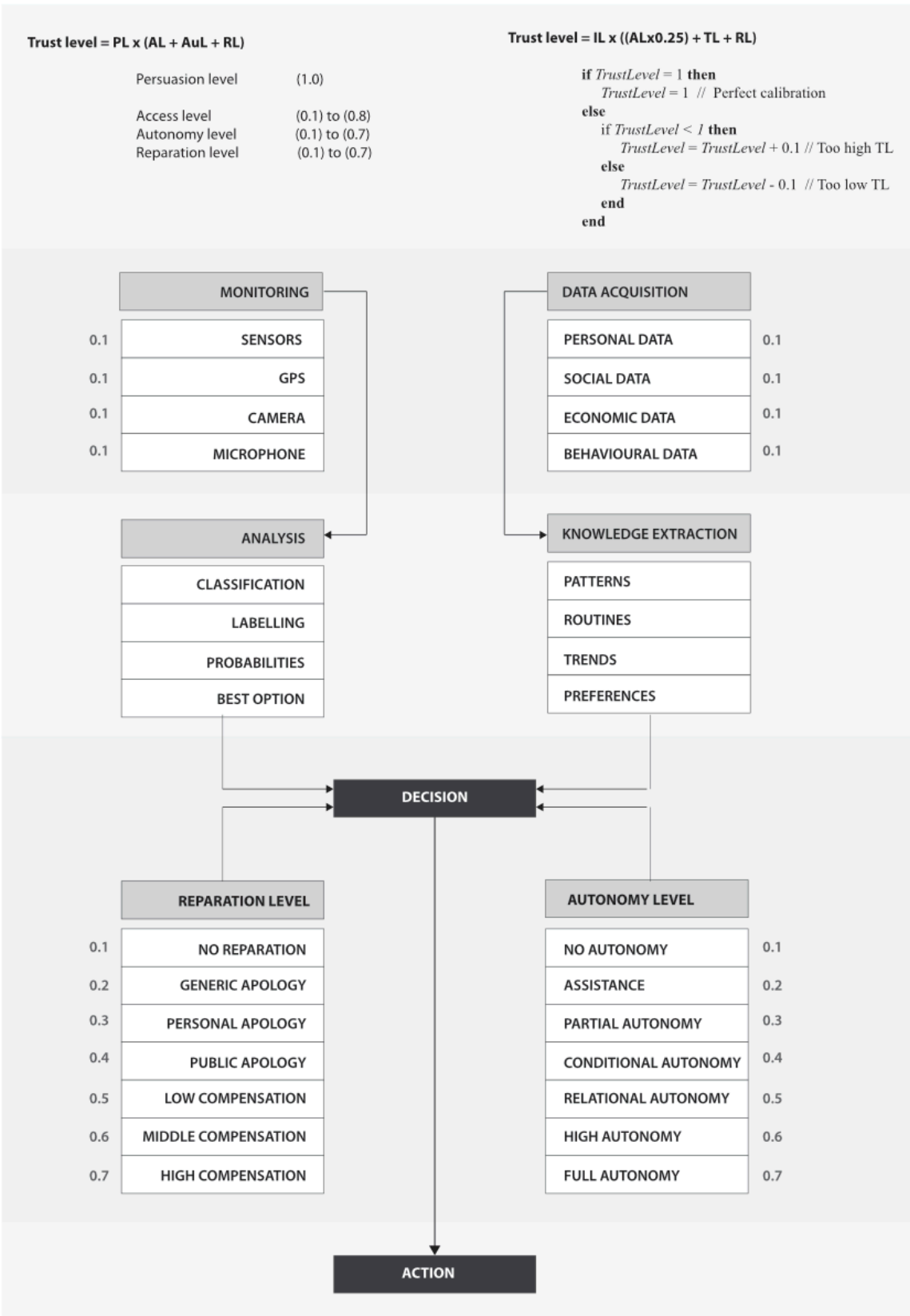
TAXONOMY



Type to enter a caption.



Type to enter a caption.



Type to enter a caption.



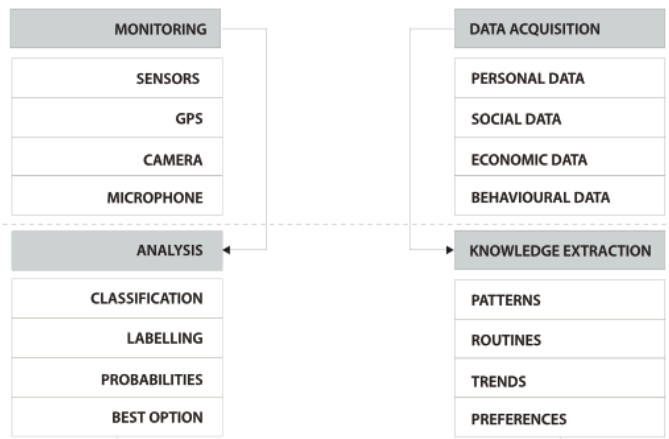
LEVELS OF PERSUASION

LEVEL 1	NO ACCOUNTABILITY
	The user performing the action
LEVEL 2	PLATFORM
	The company who owns the technology
LEVEL 3	DEVELOPER
	The designer/developer who designed the algorithm
LEVEL 4	THIRD-PARTY
	The third-party licencing the technology or delivering a service
LEVEL 5	ALGORITHM
	The artificial system

Type to enter a caption.

ACCESS

ACCESS



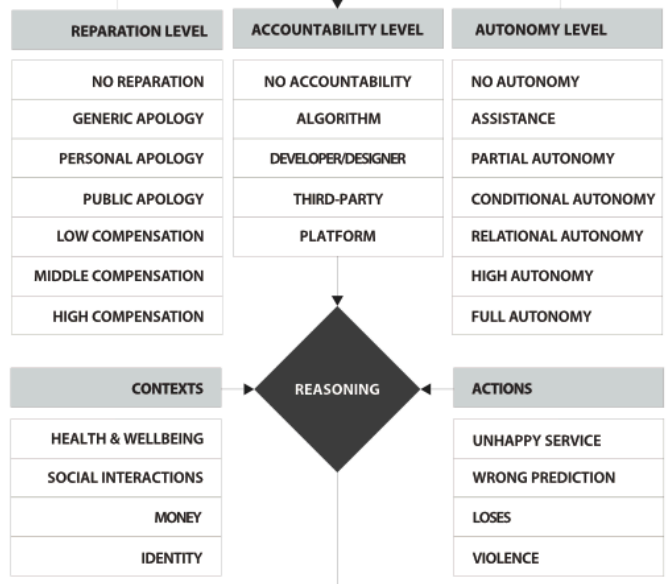
INFERENCES

INFERENCES



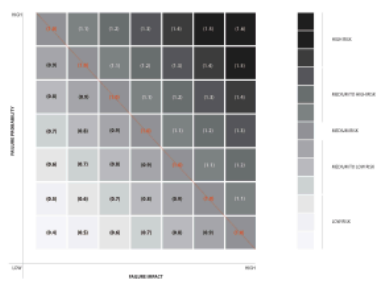
REASONING

REASONING



CALIBRATION

CALIBRATION



ACTION

ACTION



Type to enter a caption.



Royal College of Art

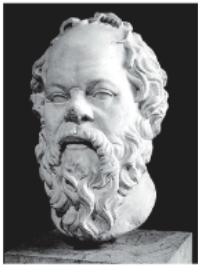
3

CHAPTER
DIAGRAMMING
ETHICS



Type to enter a caption.

VIRTUE - PERSONAL



Socrates

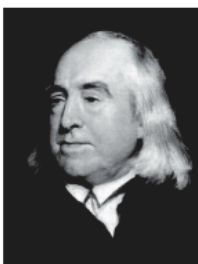
A virtue is generally agreed to be a character trait, such as a habitual action or settled sentiment. Specifically, a virtue is a positive trait that makes its possessor a good human being. A virtue is thus to be distinguished from single actions or feelings

Practical reason results in action or decision

THE USER

PROBLEM - LIMITED IN SCOPE

CONSEQUENTIALISM - INDIVIDUAL



Bentham

The consequences of one's conduct are the ultimate basis for any judgment about the rightness or wrongness of that conduct. Thus, from a consequentialist standpoint, a morally right act (or omission from acting) is one that will produce a good outcome, or consequence

The consequences are more important than the actions

THE CONSEQUENCE

DEONTOLOGY - RULES



Kant

Deontological ethics or deontology (from Greek δέον, deon, "obligation, duty") is the normative ethical theory that the morality of an action should be based on whether that action itself is right or wrong under a series of rules, rather than based on the consequences of the action

The action is more important than the consequences.

THE INTENTION

PRAGMATISM - SOCIAL



Dewey

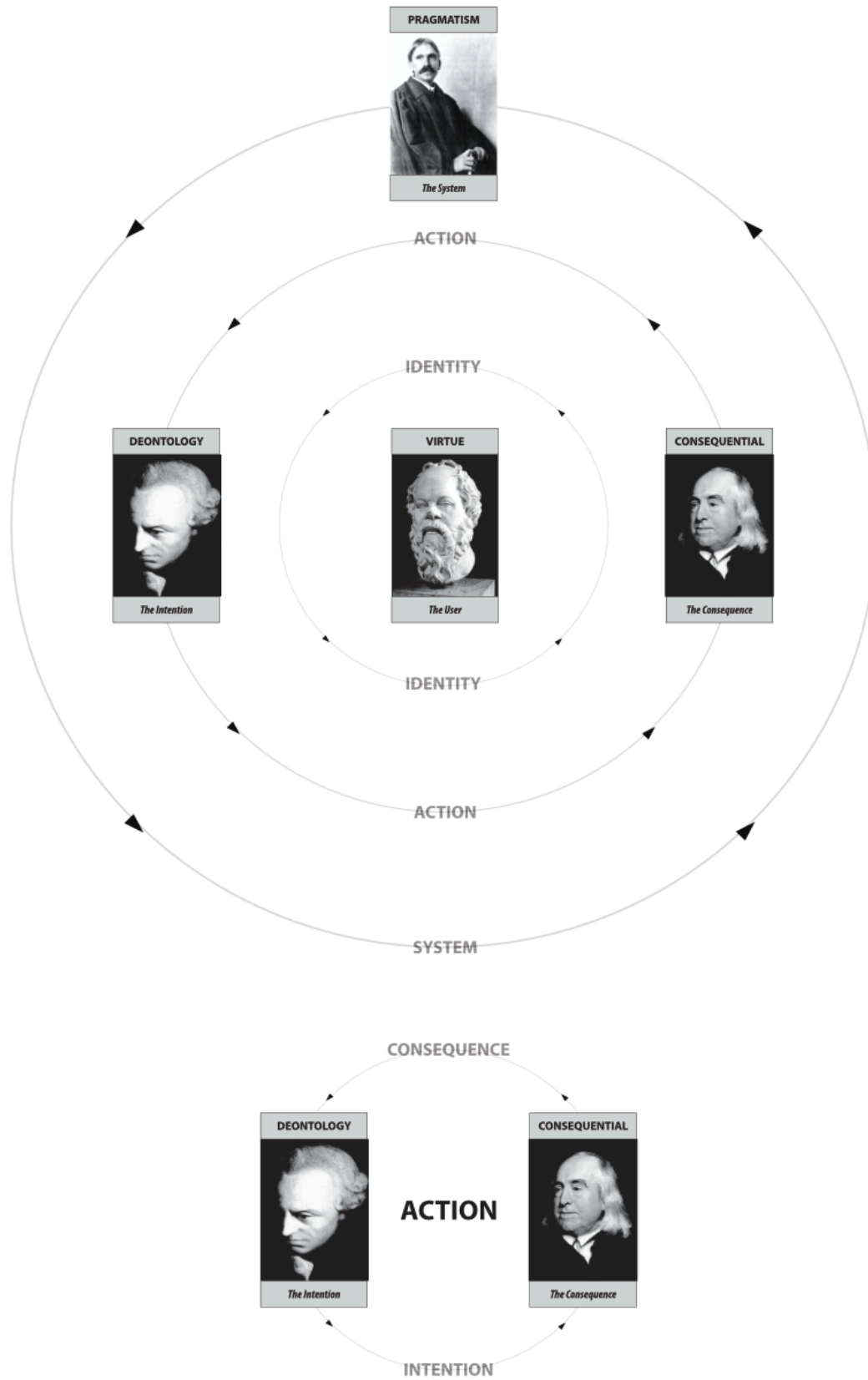
Acknowledge the need for mechanisms which allow society to advance beyond such approaches. Aimed at social innovation. We should prioritise social reform over concern with consequences, individual virtue or duty.

The system is the most important element. And determines actions and consequences

THE SYSTEM

PROBLEM - NEEDS GLOBAL CONSENSUS

Type to enter a caption.



Type to enter a caption.



Royal College of Art


4


CHAPTER


DIAGRAMMING

THE FUTURE

DIMENSIONS


 REALITY


 TIME


 SPACE

APPROACHES

TERM	AUTHOR	YEAR	DESCRIPTION
DETERMINED	ST. AUGUSTINE	426 CE	THE FUTURE IS PASSIVE AND DEPENDENT OF GOD
OPEN	JOACHIM OF FIORE	1135 - 1202	THE FUTURE IS TRANSFORMED BY HUMANS THROUGH ACTION
CONDITIONAL	LOUIS DE MOLINA	1535 - 1600	THE FUTURE ISN'T DETERMINED, NEITHER IS FULLY FREE FOR HUMANS


MATERIALISM



 ORAL
 Priests



 WRITEN
 Plato


 OBJECT
 Da Vinci

NARRATIVES


 PROGRESS


 DECLINE


 CIRCULAR


STRUCTURE


TERM	AUTHOR	YEAR	DESCRIPTION
ALTERNATIVE	PLATO (THE REPUBLIC)	426 CE	DIFFERENT PLACE AT THE SAME TIME
PROJECTIVE	18 Cth WRITING	1700 - 1800	SAME PLACE AT DIFFERENT TIME
ALIEN	GEORGE LUCAS	1977 - ONGOING	DIFFERENT TIME AND DIFFERENT PLACE


KNOWLEDGE


TERM	AUTHOR	YEAR	DESCRIPTION
FUTURA	MARCUS TULLIUS CICERO	106 - 43 BCE	SHALL COME INTO BEING, YET UNDONE
PROBABILISTIC	BERTRAND DE JOUVENEL	1903 - 1987	FUTURE CANNOT BE TRUE/FALSE BUT ABOUT THE REALM OF THE POSSIBLE


TIMEFRAMES

HUNTERS

 TIMELESS
 Ever-present

AGRICULTURAL

 CYRCLE
 Mithology

GREEKS

 LINEAR
 Rational

INDUSTRIAL

 ARROW
 Progress

POST-INDUSTRIAL

 EXPONENTIAL
 Disruptive

Type to enter a caption.

ANTIQUITY

ANTIQUITY

1000 BC - 1400 - PROPHECIES AND ALTERNATIVE FUTURES

TERM	AUTHOR	YEAR	APPROACH	FUNCTION	TIMEFRAME
PROPHECY	PRIESTS	1000 BCE	PRE-RATIONAL	PREDESTINE BY GOD	ETERNAL
LOGIC	PLATO	380 BCE	MENTAL-RATIONAL	HUMAN-CENTRED AROUND PROBLEMS	LINEAR
LOGIC	VIRGIL	42 BCE	MENTAL-RATIONAL	BETTER WORLD BASED ON HUMAN ACTIVITY	LINEAR
LOGIC	CICERO	106-43 BCE	ANALYTICAL	FUTURA: WHAT SHALL COME INTO BEING	LINEAR
THEORY	KHALDUN	1377	THEORETICAL	FUTURE IS A PLACE FOR PROGRESS OR DECLINE	CYCLICAL

RENAISSANCE

RENAISSANCE

1400 - 1800 - PLANETARY EXPLORATION - UTOPIAS OF ANOTHER PLACE

TERM	AUTHOR	YEAR	APPROACH	FUNCTION	TIMEFRAME
VISIONARY	DA VINCI	1452-1519	CONSTRUCTIVE	VISIONS PROVIDED PROTOTYPES FOR INVENTIONS	ETERNAL
UTOPIA	MORE	1516	POLITICAL	COMMUNITY OVER INDIVIDUAL VALUES	ALTERNATIVE
IMAGINATIVE PROPHECY	NOSTRADAMUS	1555	PRE-RATIONAL	FUTURE EVENTS	PROJECTIVE
NEW ASTRONOMY	COPERNICUS	1543	SCIENTIFIC	FROM GEOCENTRIC TO HELIOCENTRIC UNIVERSE	ETERNAL
FUTURA	DE MOLINA	1589	POLITICAL	FUTURE NEITHER FULLY DETERMINED, NOR FREE	CONDITIONAL

SCIENTIFIC

SCIENTIFIC REVOLUTION

1600 - 1700 - OBSERVATION AS METHOD

TERM	AUTHOR	YEAR	APPROACH	FUNCTION	TIMEFRAME
EMPIRICISM	BACON	1627	SCIENTIFIC	FROM IDEALISM TO SCIENCE AND PROGRESS	PROSPECTIVE
CARTESIANISM	DESCARTES	1637	PHILOSOPHY	THE FUTURE IS A MENTAL ACTIVITY	CONDITIONAL
SCIENCE FICTION	GODWIN	1638	LITERARY	FROM UTOPIA TO LITERARY FANTASY	ALTERNATIVE
EVOLUTIONARY	BOYLE	1662	SCIENTIFIC	THE FUTURE IS EVOLUTIONARY	INTERACTIVE
DETERMINISM	NEWTON	1687	SCIENTIFIC	MOVEMENTS CAN BE PREDICTED BY MATHS	PROJECTIVE

ENLIGHTENMENT

ENLIGHTENMENT

1700 - 1900 - THEORIES OF PROGRESS

TERM	AUTHOR	YEAR	APPROACH	FUNCTION	TIMEFRAME
PARTICIPATORY	ROUSSEAU	1783	PHILOSOPHY	SOCIALLY ENGAGED FUTURES	ALTERNATIVE
IDEALISM	SCHILLING	1800	PHILOSOPHY	HUMANISTIC IDEAS OF SOCIAL PROGRESS	PROJECTIVE
SOCIAL PROGRESS	TURGOT	1750	SOCIOLOGY	THE IDEA OF HUMAN PROGRESS	PROSPECTIVE
POSITIVISM	COMTE	1830 - 1860	SOCIOLOGY	POLITICAL SCIENCE CAN BE PREDICTED	PROJECTIVE
MARXISM	MARX	1848	POLITICAL	COLLECTIVE FUTURE	PROSPECTIVE
EVOLUTIONARY	DARWIN	1859	SCIENCE	THE FUTURE EVOLVES FROM INTERACTION	INTERACTIVE
SOCIAL ENGINEERING	SPENCER	1870	SOCIOLOGY	SURVIVAL OF THE FITTEST	PROSPECTIVE
SOCIALISM	MORRIS	1890	POLITICAL	ASPIRATIONAL FUTURES OF WORK	PROSPECTIVE

MODERNISM

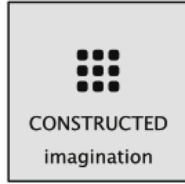
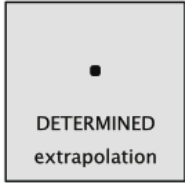
INDUSTRIAL

1900 - 2020 - KNOWLEDGE-BASED FUTURES

TERM	AUTHOR	YEAR	APPROACH	FUNCTION	TIMEFRAME
ANTICIPATION	H. G. WELLS	1901	SCIENCE/CYBERNETICS	MULTIPLICITY AND OPENNESS	LONG-TERM
FORECAST	C. K. OGDEN	1920	SCIENTIFIC	PREDICTIVE EXTRAPOLATION OF TRENDS	SHORT-TERM
FUTUROLOGY	O. K. FLECHTHEIM	1950	SOCIAL SCIENCE	PROJECTION OF HISTORY INTO NEW DIMENSION	MEDIUM-TERM
PROSPECTIVE	G. BERGER	1957	HUMAN AGENCY	FROM SEEING THE FUTURE TO TAKING ACTION	INTERACTIVE
SCENARIO PLANNING	H. KAHN	1960	COMMERCIAL	MAPPING THE FUTURE	VARIABLE
STRATEGIC FORESIGHT	R. SLAUGHTER	1990	PARTICIPATORY	NON-ACTIVIST STRATEGIC MANAGEMENT	LONG-TERM
TREND SPOTTING	2000	2000	COMMERCIAL	AGGREGATIONS OF PAST INFORMATION	SHORT-TERM
CRITICAL FUTURES	A. DUNNE, J. AUGER	2010	SOCIAL SCIENCE	ACTIVIST SCENARIO INTERVENTION	SHORT-TERM

Type to enter a caption.

PARADIGMS



POSITIVISM

1950-1960 - RAND CORPORATION - Computer-based

EMPIRICAL METHODS BASED ON NEWTONIAN PHYSICS

the systematic practise of repeating laboratory experiments and controlling variables to establish proof of our hypothesis

Mathematics
Modelling
Simulation
Gaming

main methods

- extrapolation of historical data
- utilisation of analytical models
- the systematic use of experts as forecasters opinion and speculation

extrapolative approach

- + perceived objectivity and values neutrality
- narrowness in focus (only one possible future) and lack of contextual awareness

PLURALISM

1960-2000 - MANKIND 2000 CONFERENCE - Sociology-based

HUMAN-CENTRED METHODS BASED ON SOCIOLOGY

The social and critical practise of mapping a wealth of possibles futures

Timelines
Mind maps
Future wheels
Flow-scapes

main methods

- contextual data analysis
- interpretative analytical methods
- the systematic use of participatory methods

extrapolative approach

- + perceived inclusivity and values neutrality
- loose in focus (too many possible future) and too dependent of contextual awareness

BEHAVIOURISM

2000-2020 - SOCIAL NETWORKS - Biology-based

RELATIONAL METHODS BASED ON BIOLOGY

The behavioural and ethical practise of predicting and modelling behaviour

Pattern recognition
Simulation
Nudging
Profiling

main methods

- biometric data analysis
- psychological analytical methods
- the systematic use of data mining methods

extrapolative approach

- + perceived accuracy and values neutrality
- manipulative in focus (one possible future) and narrow of contextual awareness

Type to enter a caption.

PLURALISM

1960-2000 - MANKIND 2000 CONFERENCE - Sociology-based

HUMAN-CENTRED METHODS BASED ON SOCIOLOGY

The social and critical practise of mapping a wealth of possibles futures

Timelines
Mind maps
Future wheels
Flow-scapes

main methods
- contextual data analysis
- interpretative analytical methods
- the systematic use of participatory methods

extrapolative approach
+ perceived inclusivity and values neutrality
- loose in focus (too many possible future) and too dependent of contextual awareness

INPUTS	GATHERING INFORMATION	METHODS
		WORKSHOPS
		QUESTIONNAIRES
		INTERVIEWS
		SURVEYS
		TECHNOLOGY ASSESSMENTS

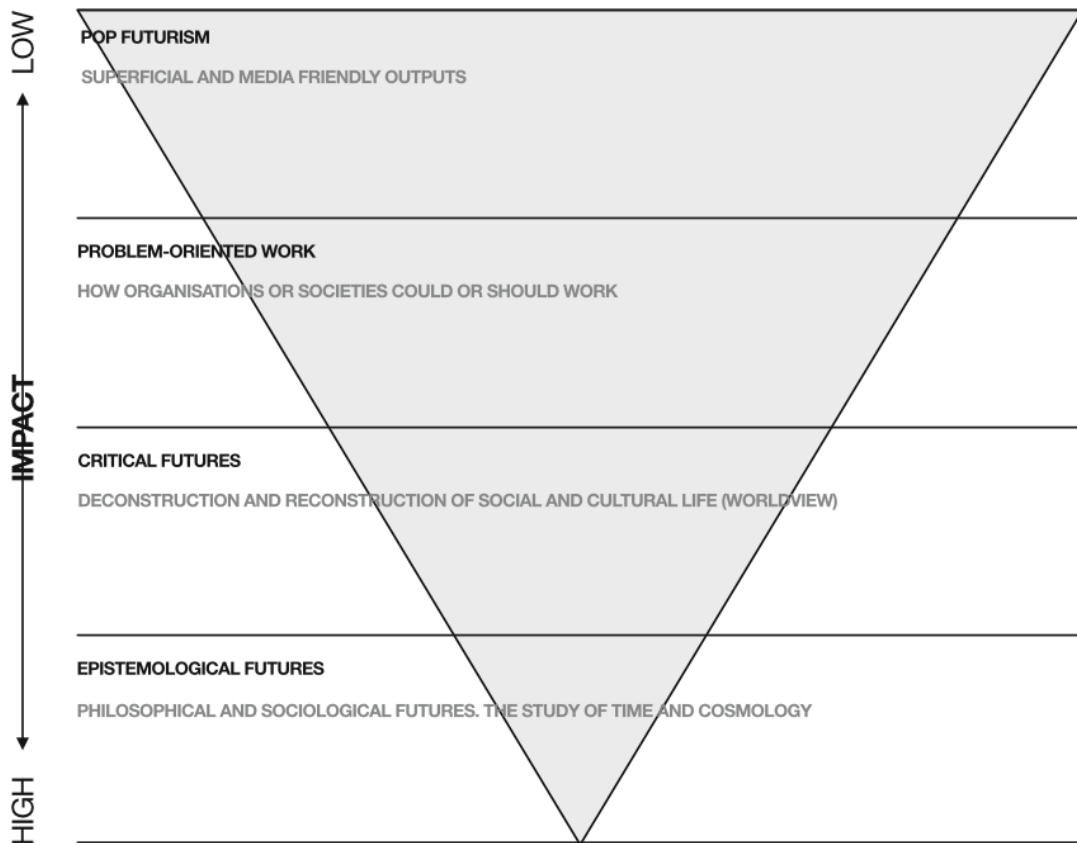
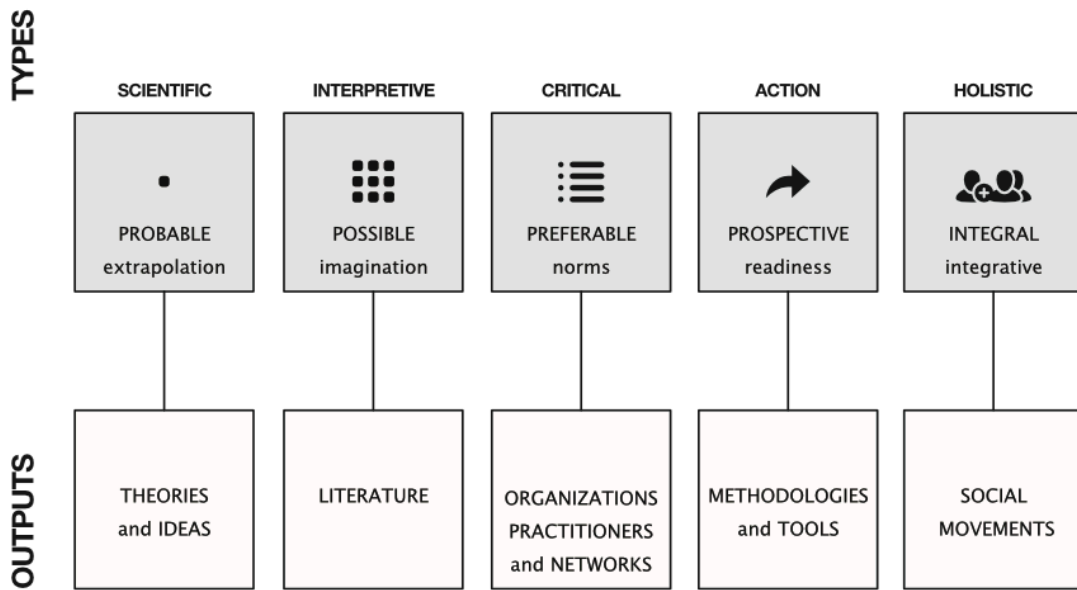
ANALYSIS	MEANING-MAKING	METHODS
		ISSUES ANALYSIS
		TRENDS ANALYSIS
		EXTRAPOLATION
		CROSS-IMPACT ANALYSIS
		TEXTS ANALYSIS
		DIALOGUE

INTERPRETATIVE	GAINING DEEPER INSIGHT	METHODS
		MACRO-HISTORY
		CAUSAL LAYERED ANALYSIS
		SYSTEMS THINKING
		HERMENEUTICS
		MIX-METHODS - BRICOLAGE
		ETHNOGRAPHY
		MEDIA CRITIQUE
		CULTURAL ARTEFACTS

OUTPUTS	TO PRODUCE FUTURE IMAGES	METHODS
		VISIONING
		IMAGINATION
		CREATIVITY
		SCENARIO PLANNING
		VISION-ACTION
		ACTION RESEARCH
		PARTICIPATORY FUTURES
		QUESTIONNAIRES

Type to enter a caption.

ACADEMIC STUDIES

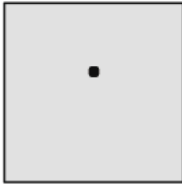


Type to enter a caption.

DEVELOPMENT STRATEGIES

TYPES

TRANSFORMATION

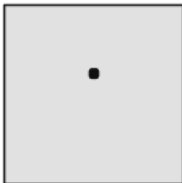


Every so often, a technology emerges which changes how the whole of society works - electricity, for example, or more recently the internet. Lately, there's been much discussion of how artificial intelligence or the blockchain might have sweeping effects.

What might the next "general purpose technology" be?

OUTPUTS

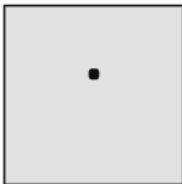
MULTIPLIER



Many significant innovations effect change by modifying how people work. An invention - the PC made mathematics quicker. Or it might be a way of doing something: open-source allowed coders to collaborate on programs.

What new "enabling" technology could help people work far more effectively in future?

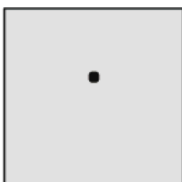
MASH-UPS



Sometimes it's the combination of existing or emerging technologies that makes for the most powerful social change. For instance the press and cheap paper impacted journalism.

Which technologies, new or old, might come into their own if combined with others?
unlikely pairing

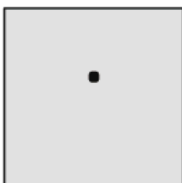
NEW SOLUTIONS



The past decade has seen many championing digital technologies as useful tools for tackling such issues.

Are there other, more tangible, technologies we should be developing to make progress on social challenges, or discoveries in parallel fields which could be applied to live social problems?

IN FROM THE COLD

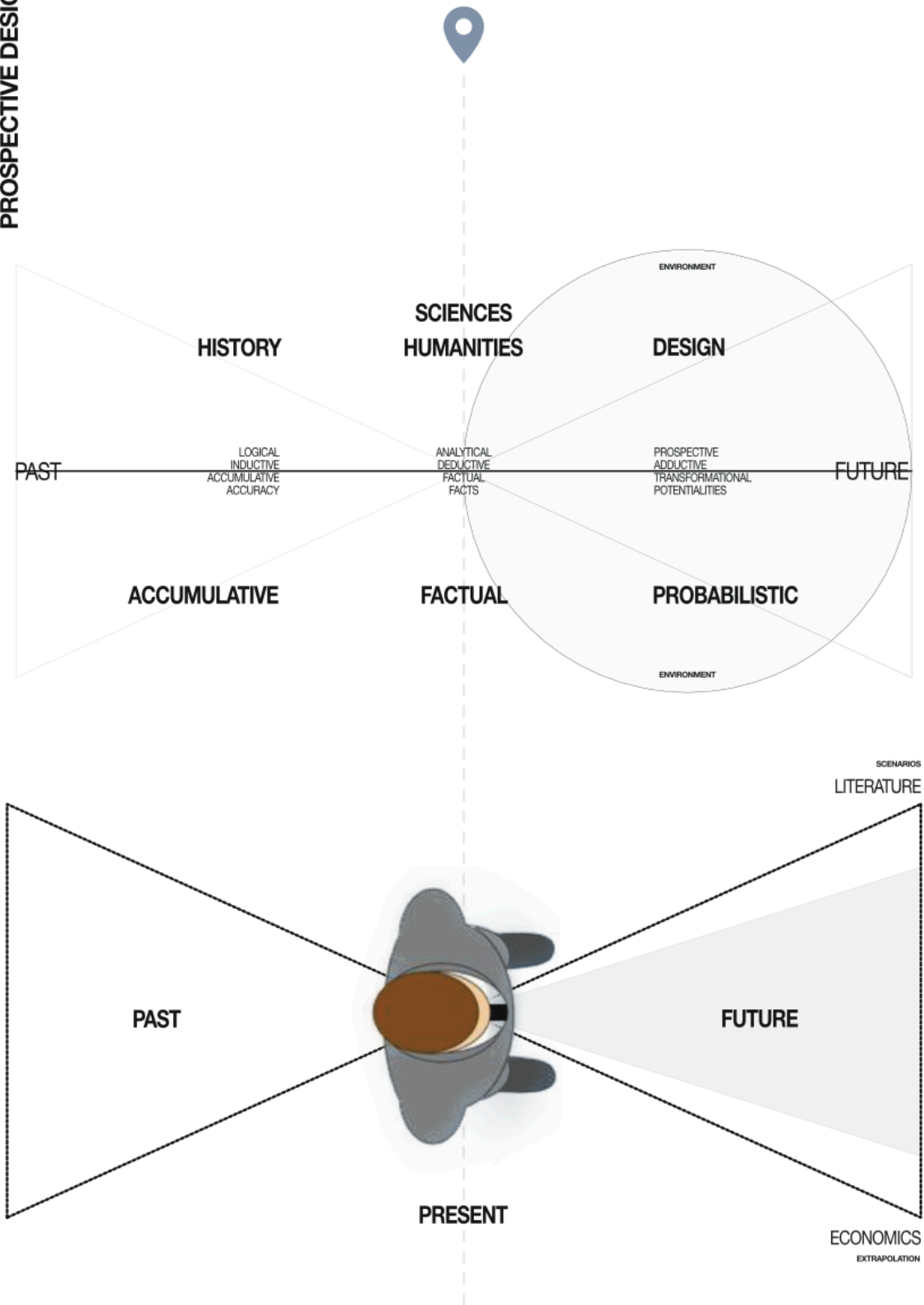


Sometimes, a specific innovation or technology simply lacks a killer 'use case'. Perhaps the field is simply perceived as an ugly duckling, confined to one of the two 'troughs' of [Gartner's hype cycle](#).

Which technologies or scientific ideas are overdue for their moment in the sun?

Type to enter a caption.

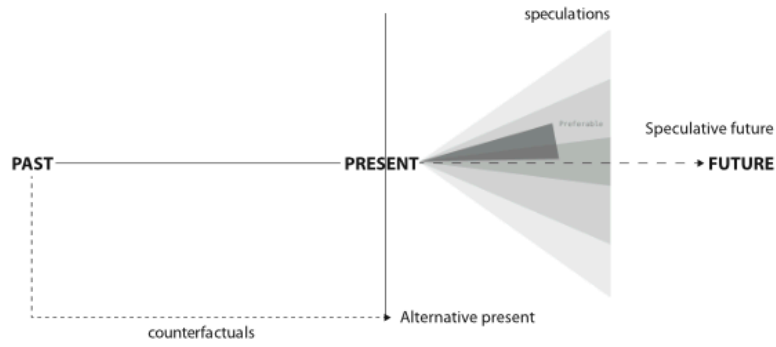
PROSPECTIVE DESIGN



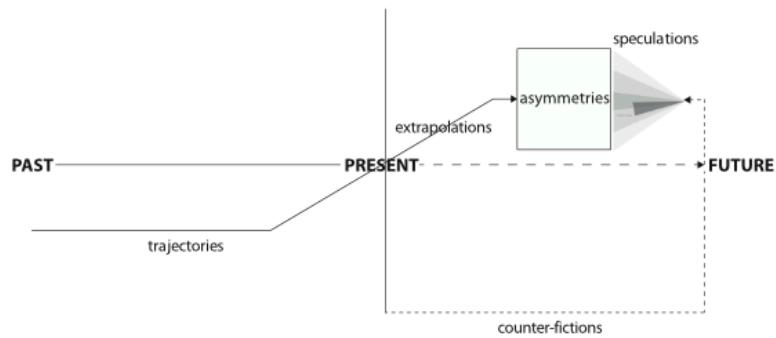
Type to enter a caption.

PROSPECTIVE VS SPECULATIVE

FUTURE METHODS



METHODS	WHAT IF? SCENARIOS (DEFINE INTERVENTION)	Brief Critical analysis
DEVELOPMENT	PLAUSIBILITY	Content analysis
TEST VALIDITY	DEBATE (SOCIAL EXPERIMENT) SOCIAL ENGAGEMENT (AESTHETIC QUALITIES)	Social networks + Exhibitions Content analysis
OUTPUT	INTERVENTION	Physical object



METHODS	TRAJECTORIES ASYMMETRIES (DEFINE INTERVENTION)	Chronological analysis Critical analysis
DEVELOPMENT	DESIGN PRINCIPLES (SAFETY)	Archives + Artefact analysis
TEST VALIDITY	COUNTERFACTUALS (THOUGHT EXPERIMENT) ALTERNATIVES	Focus groups or Workshops Content analysis
OUTPUT	FRAMEWORK FOR PRODUCT DEVELOPMENT	Document - Catalogue



EVIDENCE	DEMO	75% - 100%
EVIDENCE	PROTOTYPE	50% - 75%
EVIDENCE	PATENT	25% - 50%
EVIDENCE	IDEA	0% - 25%

Type to enter a caption.

2012- COUTRE-FICTIONS - Yves Citton

HUMAN-CENTRED METHODS BASED ON SOCIOLOGY

the non-place of the fiction constitutes an unlocking operator become central in the contemporary political conflicts, in that it allows to open a space of rejection of the data, in a world where they are of more and more the "data" that oppress us

VISIONARY DIMENSIONS = ANTY-SYSTEMIC NARRATIVES

INITIATORY CENTER-FICTIONS

- DESIGNERS AS INITIATORS - new worlds
- REALITY = FICTION
- FACT = CONSTRUCTS ISOLATED REALITY (according to certain ends, perspectives, values, strategies, etc.)
- FACT BECOMES FICTION = SIMULATION AND SIMULACRA

DENUNCIATORY COUNTER-FICTIONS - against productivist disorientation
- REACTIONS OF DETACHMENT

DOCUMENTARY COUNTER-FICTIONS - against colonisation of clichés
- TO PORTRAY REALITY BLOCKS (SURPRISE AND DISCOVERY) FROM INTENTIONALITY TO SERENDIPITY

Citton, Y. (2012). Contre-fictions : trois modes de combat. *Multitudes*, 48(1), 72-78. doi:10.3917/mult.048.007 (influenced by Latour)

2015 - COUNTER-FICTION - Ludovic Burel, Ju Hyun Lee

HUMAN-CENTRED METHODS BASED ON PHENOMENOLOGY

Performance and the body

infiltration
piracy

conference/performance/publication

Type to enter a caption.

2012- TRIAD

ARISTOTLE - TRIAD OF KNOWLEDGE

NICOMACHEAN ETHICS, BOOK II, CH. 2 (II39a27-28)
NICOMACHEAN ETHICS, BOOK VI, CH. 2 - 7
METAPHYSICS, BOOK I, CH. 1-3

Builds from Pythagorean tradition

THEORETICAL KNOWLEDGE - episteme

- Metaphysics; mathematics; natural sciences (NE II39a27-28)
- immovable (NE II39a27-28)
- Things universal and necessary (EN II 40b31-32)
- Beyond true object - contemplation
- No practical end, but focused on ends - origin (MET 982b12-21)

PRACTICAL KNOWLEDGE - praxis

- Ethics, politics (NE II39a27-28)
- in motion; has an end (NE II39a27-28)
- concerned with action and human behaviour
- action itself is its end (NE II40b5-7)
- action for the sake of an end (PHY 199a8-18)

PRODUCTIVE KNOWLEDGE - poiesis

- Architecture; medicine; navigation (NE II39a27-28)
- have no end in itself. always implicated/remain in exchange
- resistance to determinate end
- capacity to make involving reasoning
- concern with something coming into being
- things that can be otherwise. always outside of itself.
- it cannot transcend time. depends on time and circumstances (me = future exist)
- ability to size the right moment (timing)
- neither, the user or the producer, are capable of determining productive knowledge. it is defined by an act of exchange
- has no external arbiter, no final judge. Only users and makers who change with exchange
- its transfer always redefine the subjects involved by effecting a shift in power and status
- defined by social exchange
- concerned with competing standards of value rather than securing boundaries of knowledge
- concerned with indeterminate and possible. with alternative possibilities (47;7357a4-5)
- is instrumental and situated
- its value is social and economical
- From passive intellect (reason becoming its object) to active intellect (object being defined)

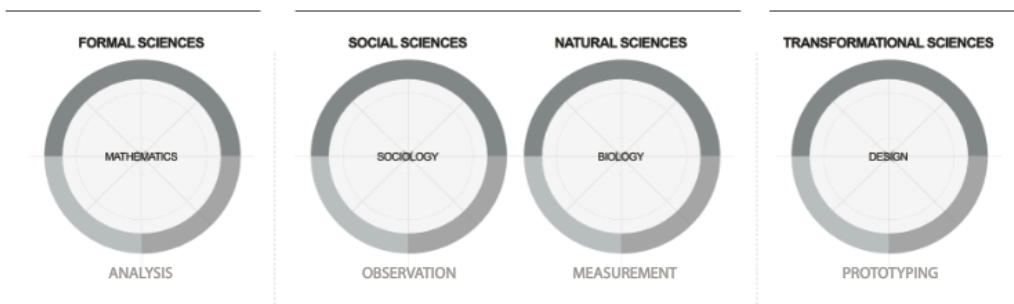
- faculties of doing that which we choose with the material available (Lobkowitz)

Atwill, J. M. (1998). *Rhetoric Reclaimed; Aristotle and the Liberal Arts Tradition*. First Edition. Cornell University Press

Type to enter a caption.

SCIENCE

from the Latin word *scientia*, meaning "knowledge"



A Priori

Theoretical knowledge encompasses abstract subjects. It is concerned with things that are universal and necessary. Yet that cannot be applied. Theoretical knowledge can never be utilitarian.

In situ

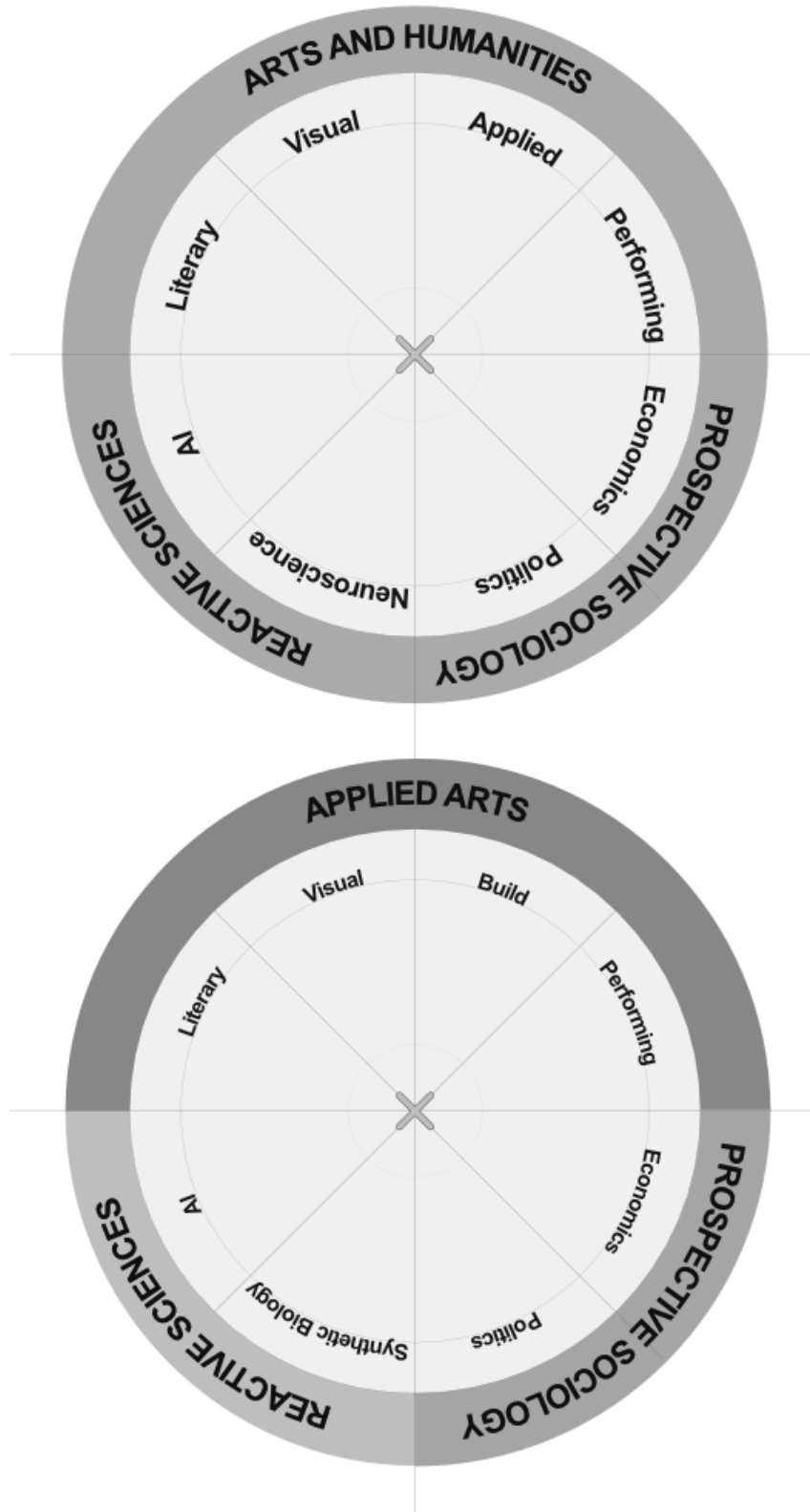
The practical is applied and question based; it has a beginning and an end builds on observation and measurement.

A Posteriori

Based on continuous interaction with the environment. It's transformational and committed to practice (Atwill, 1998). The capacity to make involving reasoning to 'go beyond' of what can be 'otherwise'. It is always in beta.

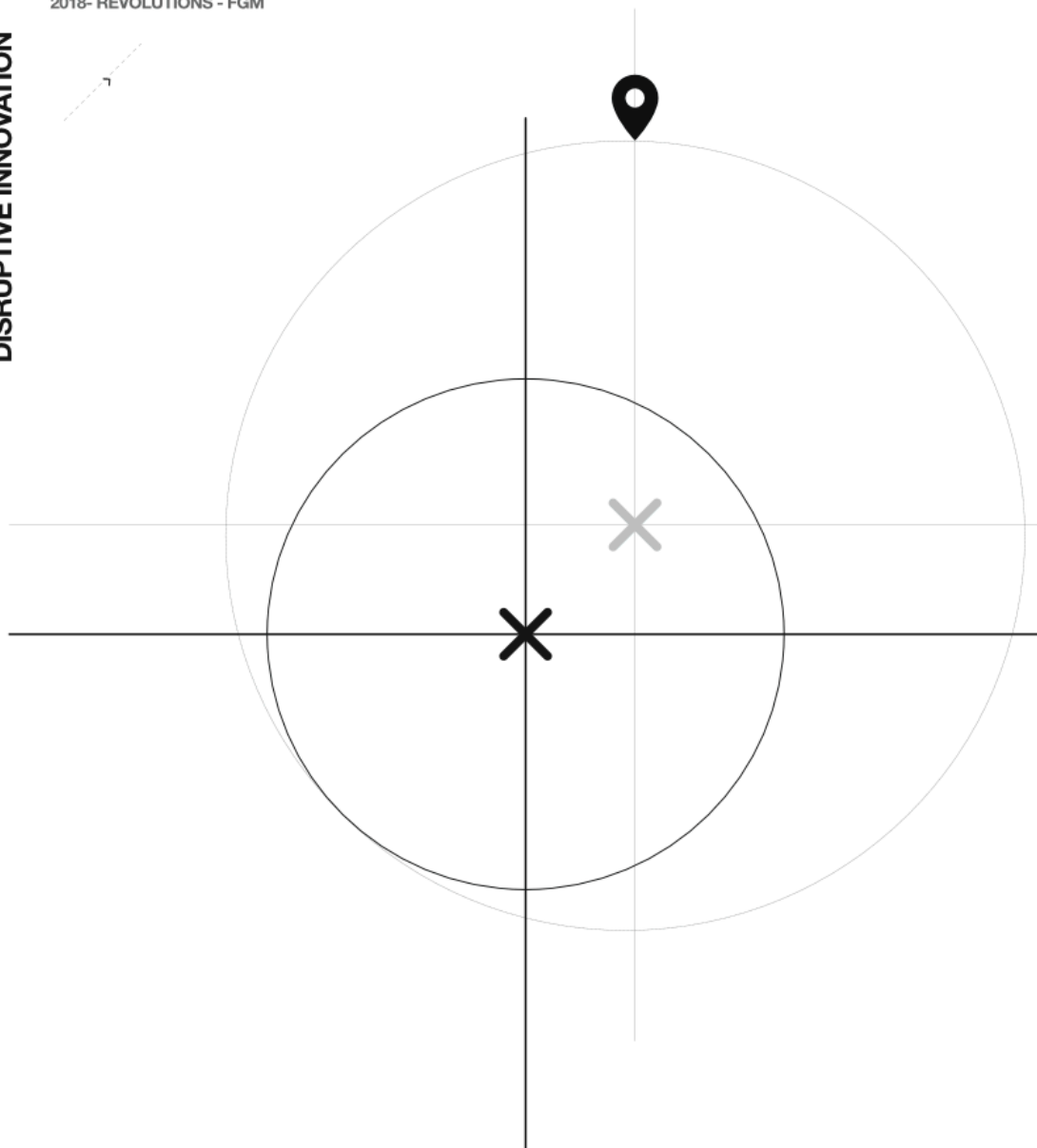
Type to enter a caption.

TRANSFORMATIONAL SCIENCES



Type to enter a caption.

DISRUPTIVE INNOVATION



PROCESS

A revolution is always an intellectual transformation

You need to localise a point to observe the current state from an outside perspective

It demands a passive optic to neutralise an active optic, which only affirms what you are seeing

This points allows you to reflect on what is latent to describe a new centre

- Copernicus - used cosmology as a passive optic to observe and describe a new centre - the sun
- Luther - used justice as a passive optic to observe and describe a new centre - divine justice
- Einstein - used physics as a passive optic to observe and describe a new centre - relativity
- Lenin - used historical materialism as a passive optic to observe and describe a new centre - communism

Type to enter a caption.



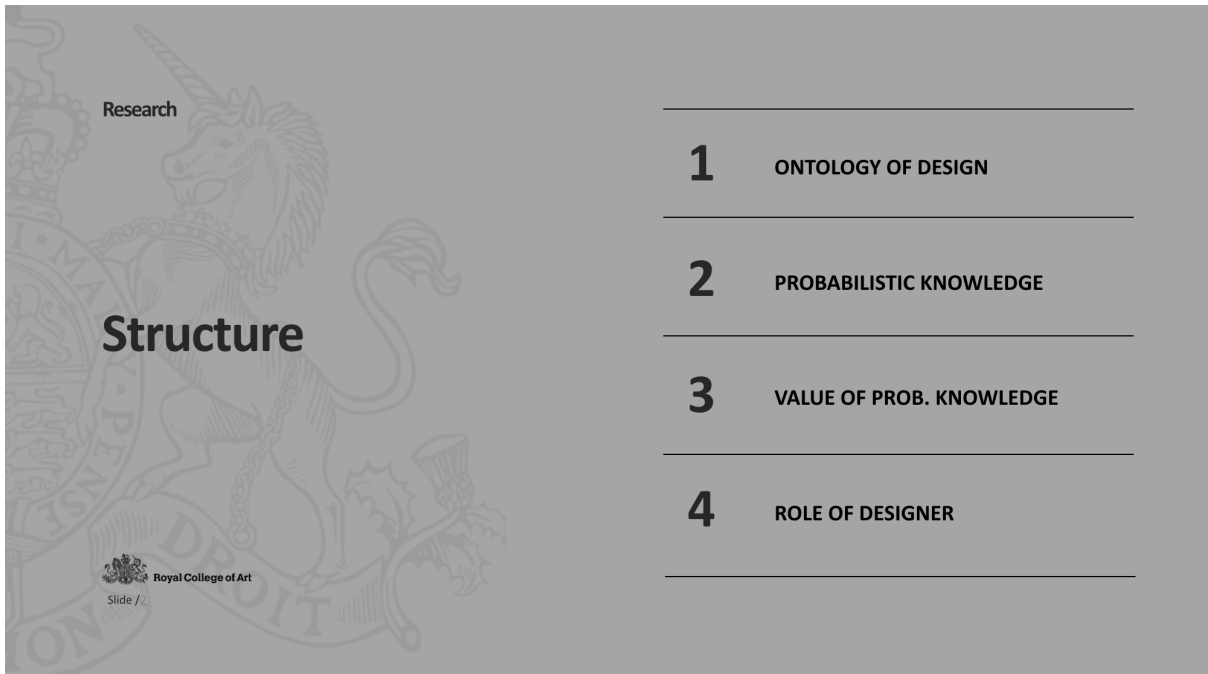
Royal College of Art

5

CHAPTER
DIAGRAMMING
ONTOLOGY



Type to enter a caption.



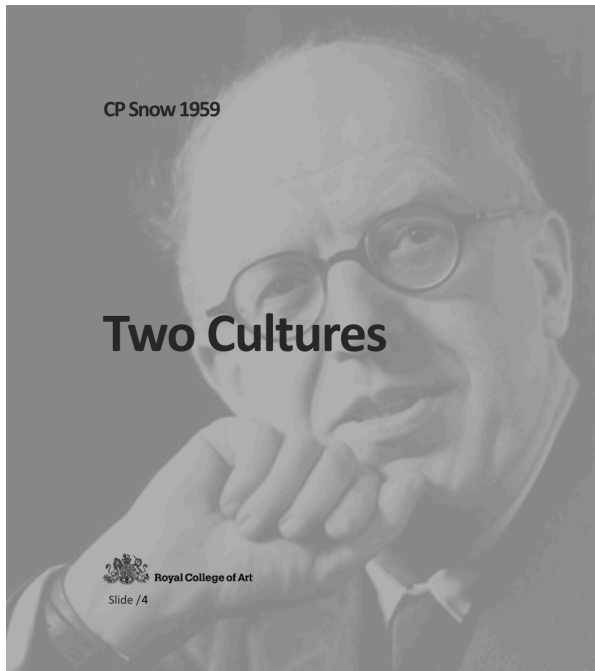
Type to enter a caption.

Research

The ontological nature of design



Type to enter a caption.

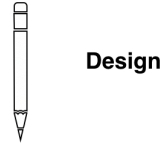
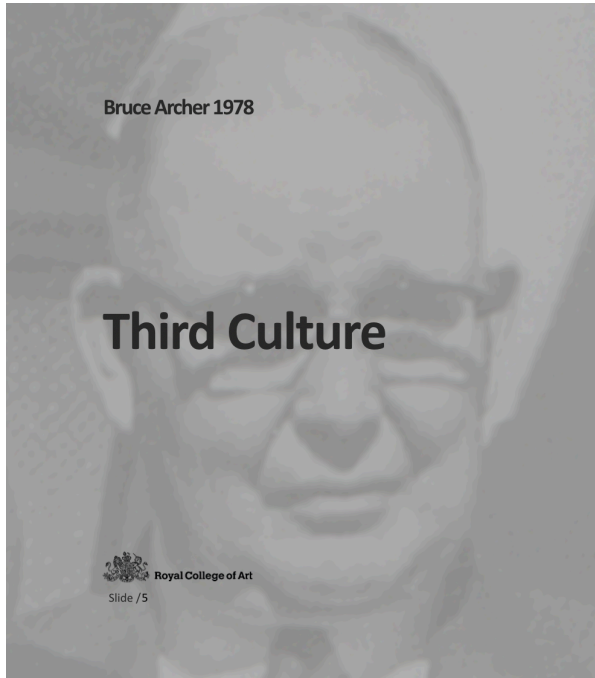


Science

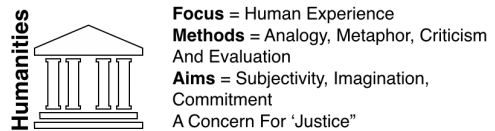
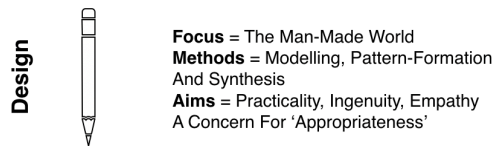
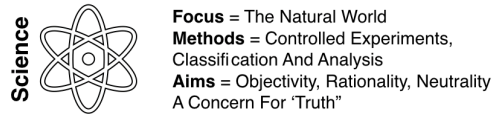
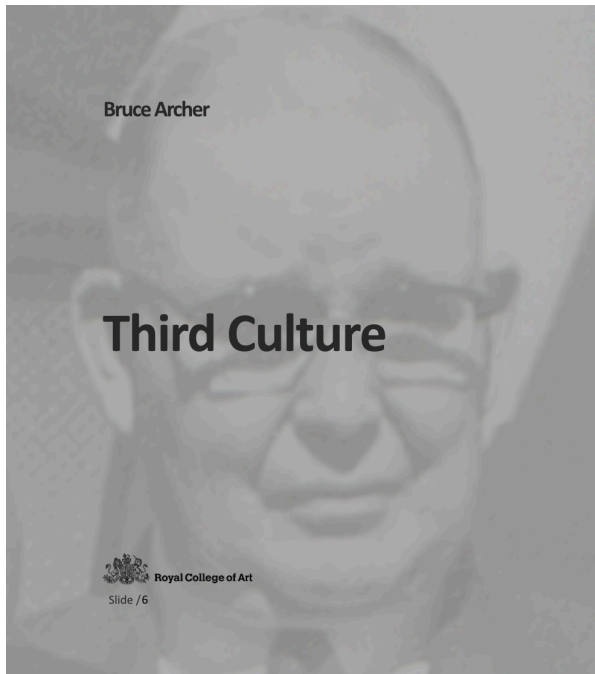


Humanities

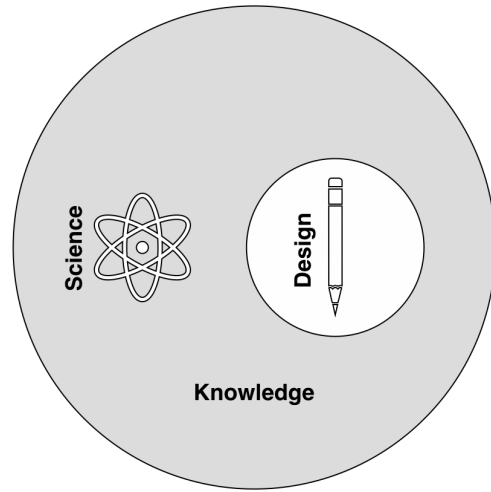
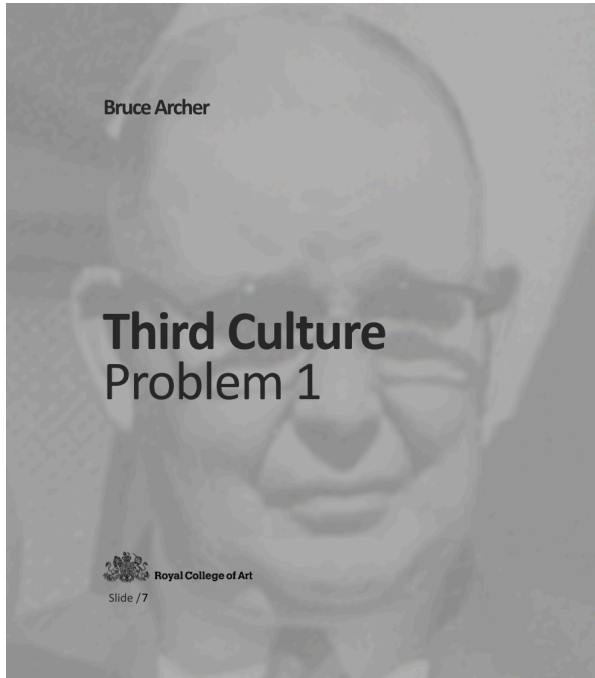
Type to enter a caption.



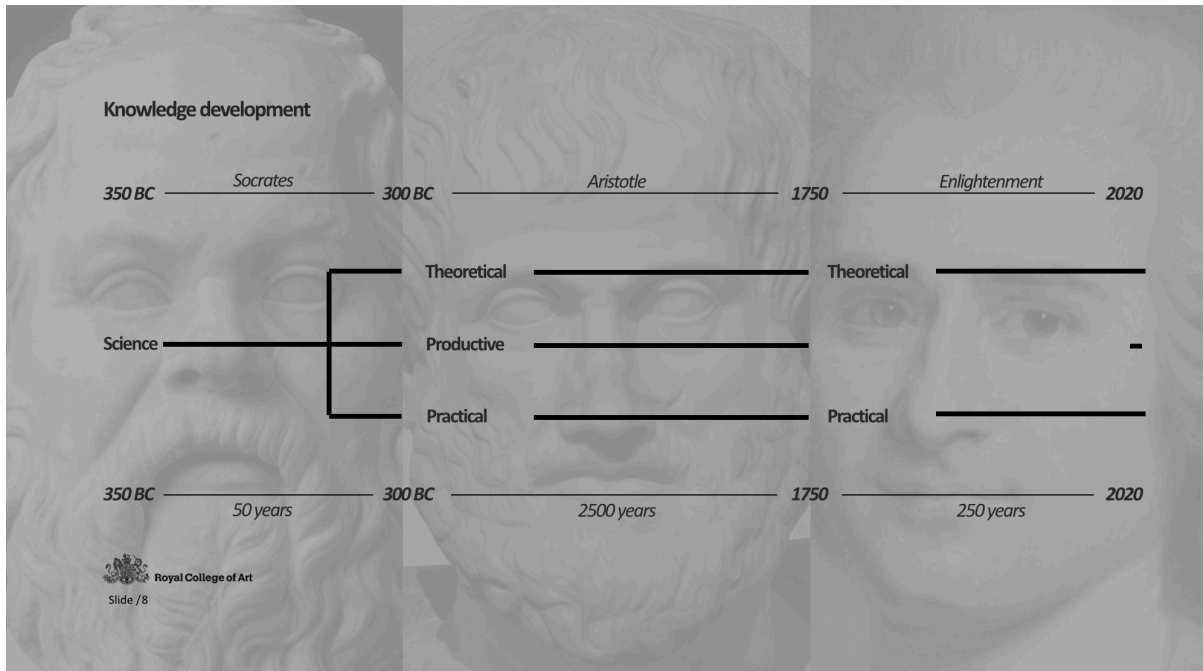
Type to enter a caption.



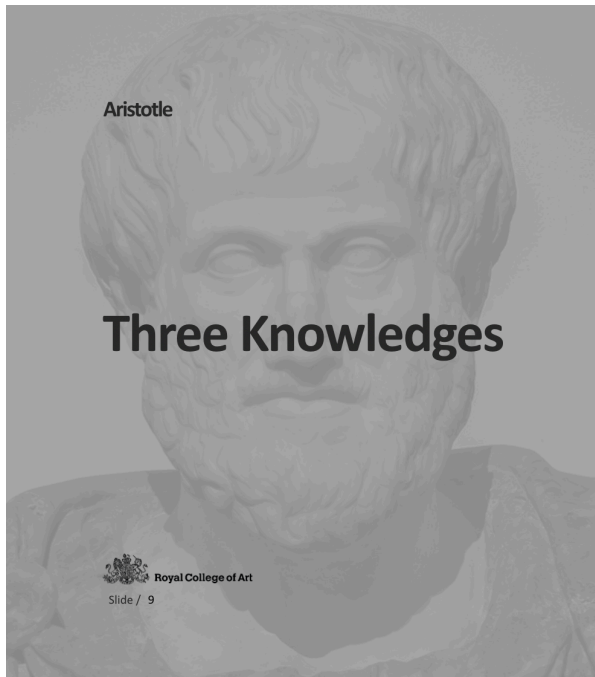
Type to enter a caption.



Type to enter a caption.



Type to enter a caption.



Theoretical



Encompasses Abstract Subjects. It Is Concerned With Things That Are Universal And Necessary. Not Applied.

Productive



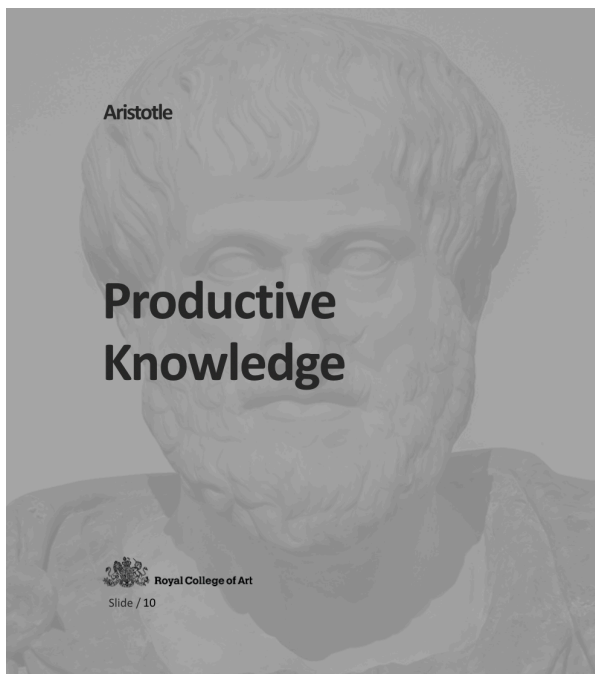
It Is Based On A Continuous Interaction With The Environment. It Is Based On Exchange And Concern With Something Coming Into Being

Practical



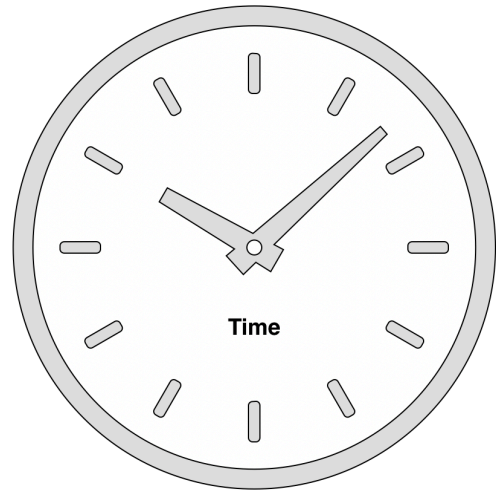
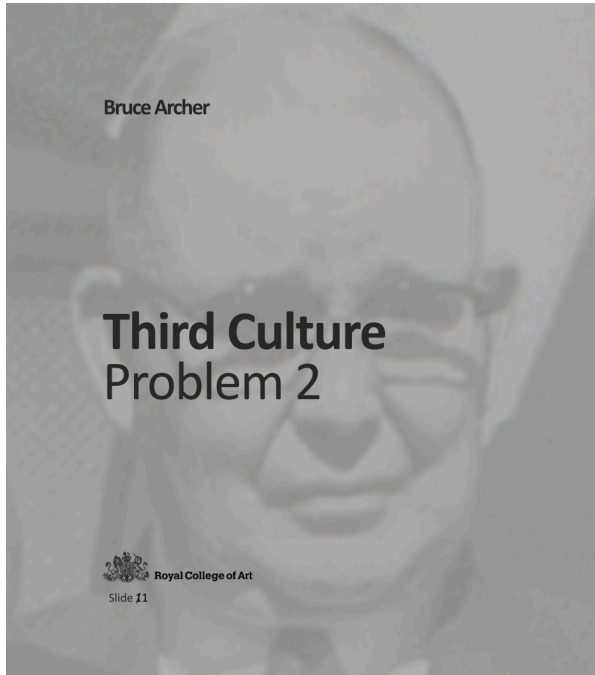
The Practical Is Applied And Question Based; It Has A Beginning And An End. Is Concerned With Action (Praxis) Toward An End

Type to enter a caption.



- 1 Productive practices are directed towards **means**, not ends
- 2 Knowledge is neither in the user, nor the producer, but in the **exchange** between them
- 3 Exchange always **redefines the subjects** involved by effecting a shift in power and status
- 4 It is concerned with competing standards of **value** rather than securing boundaries of knowledge
- 5 Its ontology is **indeterminate** as it is based on **potentialities** or alternative possibilities
- 6 It **cannot transcend time** like mathematics assuming past, present and future timeframes
- 7 Knowledge is always "outside itself" residing in the **use** made by a receiver or audience
- 8 It has no external arbiter and no final judge. Only users and makers who **change** with exchange
- 9 It is **transformational** in nature.

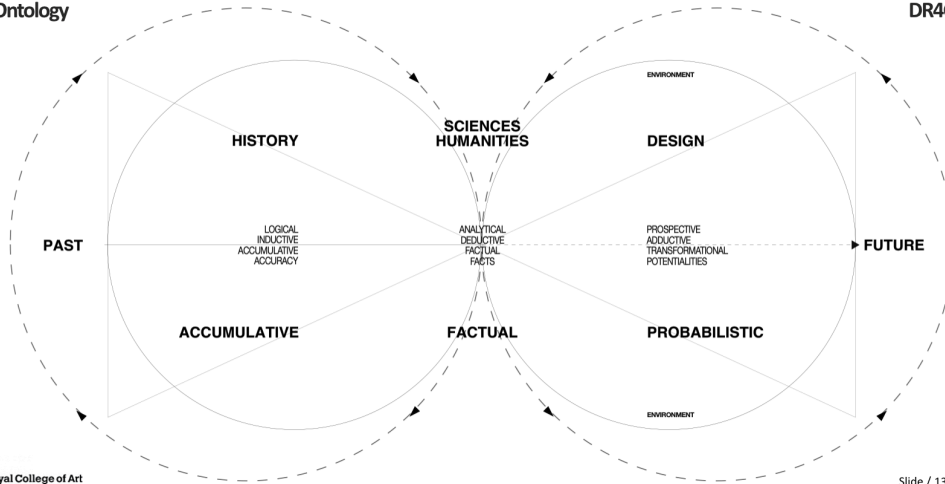
Type to enter a caption.



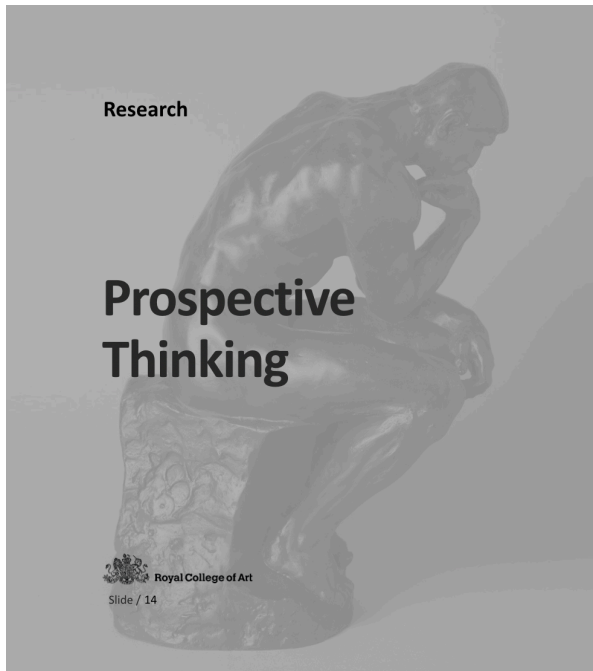
Type to enter a caption.

A presentation slide with a grey background. At the top left, the text "Prospective design" is written in a small, black, sans-serif font. At the top right, the text "DR4C" is written in a small, black, sans-serif font. In the center, there is a quote by John Christopher Jones. The quote is enclosed in large, bold, black quotation marks. The text of the quote is in a smaller, italicized, black, sans-serif font. At the bottom left, there is a small logo of the Royal College of Art, followed by the text "Royal College of Art". At the bottom center, the name "John Christopher Jones" is written in a bold, black, sans-serif font. At the bottom right, the text "Slide / 12" is written in a small, black, sans-serif font. The background features a faint, large-scale portrait of John Christopher Jones.

Type to enter a caption.



Type to enter a caption.



Research

Prospective Thinking

ANTQUITY	1000 BC - 1600 - PROPHECIES AND ALTERNATIVE FUTURES		
Author	Year	Year	Approach
Plato	427-347 BC	380-275 BC	Philosophical
Cicero	106-43 BC	106-43 BC	Philosophical
Da Vinci	1452-1519	1452-1519	Artistic
More	1473-1535	1473-1535	Philosophical
Bacon	1561-1626	1561-1626	Philosophical
Newton	1643-1727	1643-1727	Scientific
Rousseau	1712-1788	1712-1788	Philosophical
Marx	1818-1883	1818-1883	Philosophical
Einstein	1879-1955	1879-1955	Scientific
H. G. Wells	1866-1946	1866-1946	Fictional

Type to enter a caption.

Research

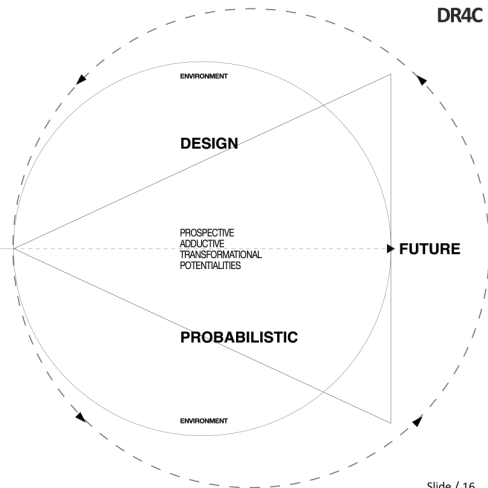
Probabilistic knowledge



Type to enter a caption.

Probabilistic Knowledge


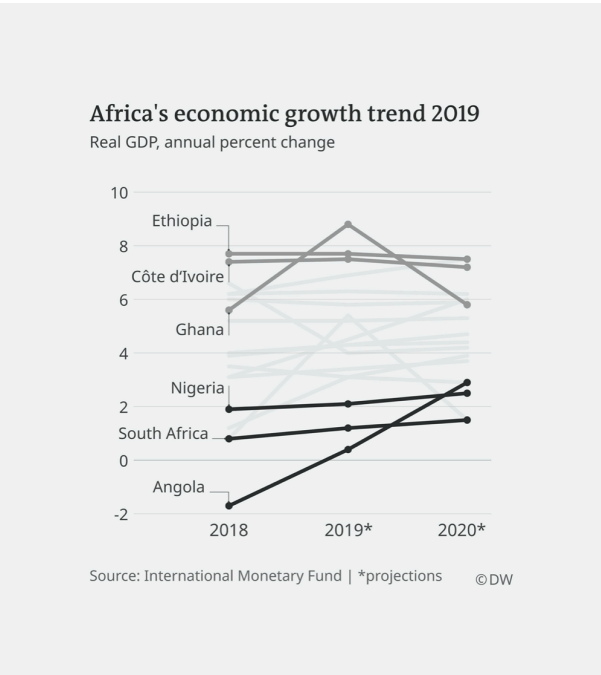
We trade some degrees of accuracy to access areas yet-to-be



Type to enter a caption.

Research

Prob. Knowledge





Type to enter a caption.

Probabilistic knowledge DR4C

Characteristics

- 1** Partial Knowledge - Known a posteriori
- 2** Dependent on environmental factors - (exchange)
- 3** Preliminary in nature - (no conclusive)




Slide / 18

Type to enter a caption.

Research DR4C

Output

- 1** Directional and transformational
- 2** Addressing future risks
- 3** Aims to deliver Impact and value


 Royal College of Art Slide / 19

Type to enter a caption.


Probabilistic knowledge

Actions

Mallet, L., (2018). Creating quality living: the new Swedish town offering innovative solutions to London's housing crisis. Evening standard. <https://www.homesandproperty.co.uk/property-news/the-new-swedish-town-offering-innovative-solutions-to-londons-housing-crisis-a116751.html>

 Royal College of Art

New typologies



Sustainability and Mobility

Type to enter a caption.

Research

The value of Probabilistic knowledge



Type to enter a caption.

Research DR4C

Areas

- 1 Social
- 2 Economical
- 3 Environmental

Royal College of Art Slide / 23


A large, faint watermark of a unicorn crest, which is the crest of the Royal College of Art, set against a grey background.

Type to enter a caption.

Research DR4C

Objectives

- **1** Appropriateness
- **2** Readiness
- **3** Preparedness


 Royal College of Art Slide / 24

Type to enter a caption.

Research DR4C

Outputs

- **1** Translational - knowledge into potential opportunities
- **2** Ethical - Reduce future risks
- **3** Social - Enhance governance to transform society

 Royal College of Art Slide / 25

Type to enter a caption.

Research

The Role of the designer



Type to enter a caption.

Research

DR4C

From a facilitator ...
To an expert ...

Royal College of Art

Slide / 28

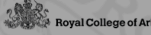
The diagram consists of two parts. On the left, a central white figure in a suit (representing a facilitator) is surrounded by four black figures (representing team members). Arrows point from the central figure to each team member, and curved arrows connect the team members in a circle. On the right, the same setup is shown, but the central figure is now black (representing an expert) and the team members are white. The text "From a facilitator ... To an expert ..." is placed between the two diagrams. The background of the slide is a light gray with a faint watermark of the Royal College of Art crest.

Type to enter a caption.

Research DR4C

Expert role

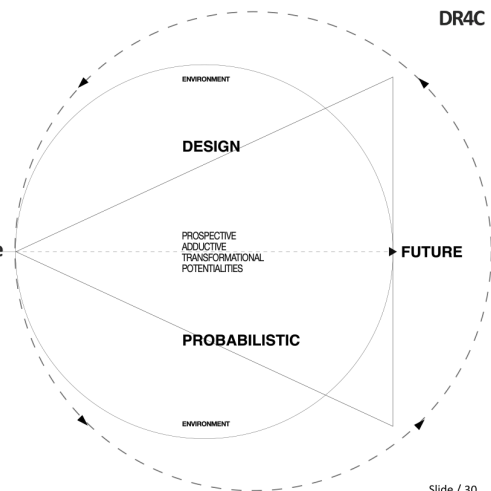
- 1** Prospecting knowledge into potential transformations
- 2** Addressing Ethics to Reduce future risks
- 3** Enhancing social governance

 Royal College of Art Slide / 29

Type to enter a caption.

The Role of the Designer

Shaping/Prospecting new futures through probabilistic knowledge



 Royal College of Art

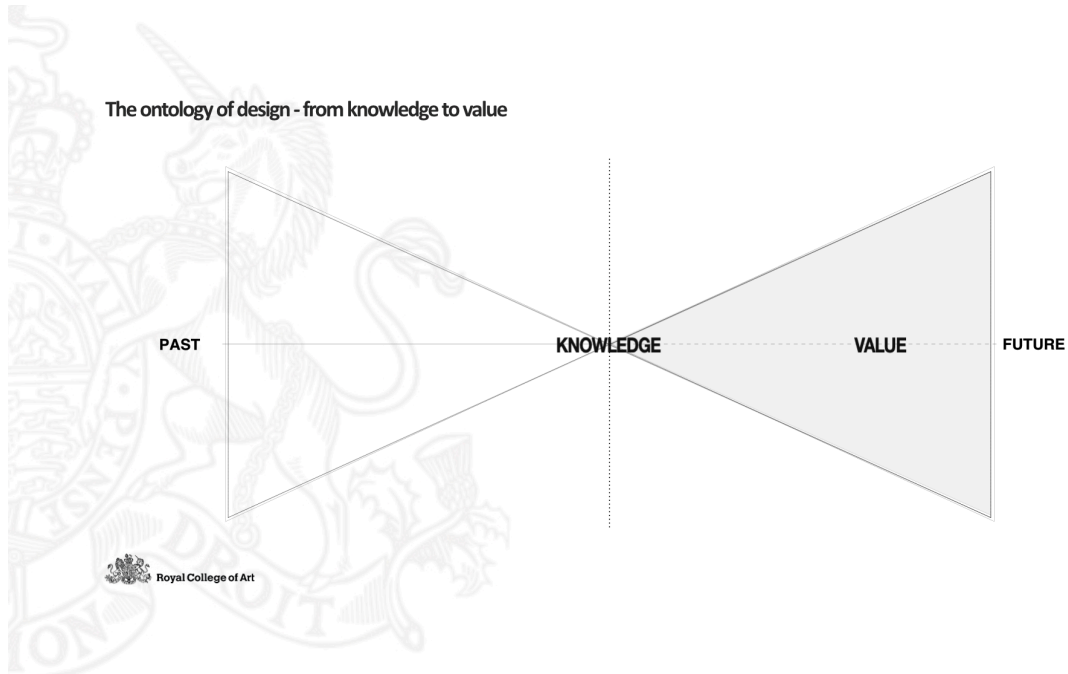
Type to enter a caption.

Research

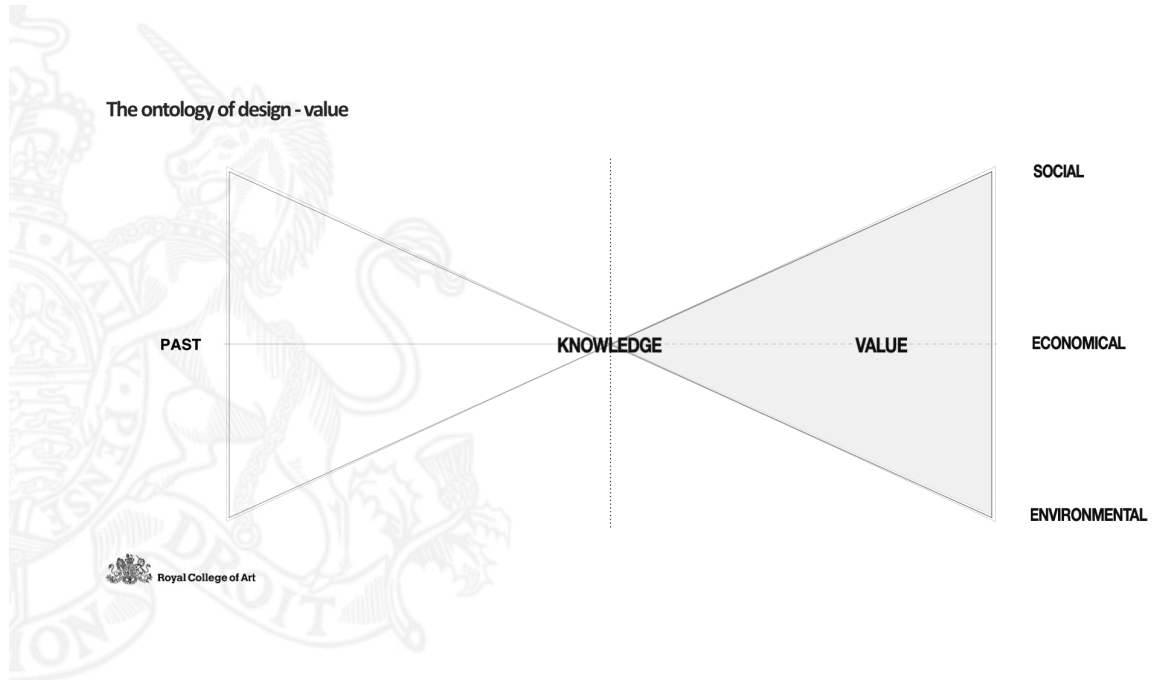
Conclusions



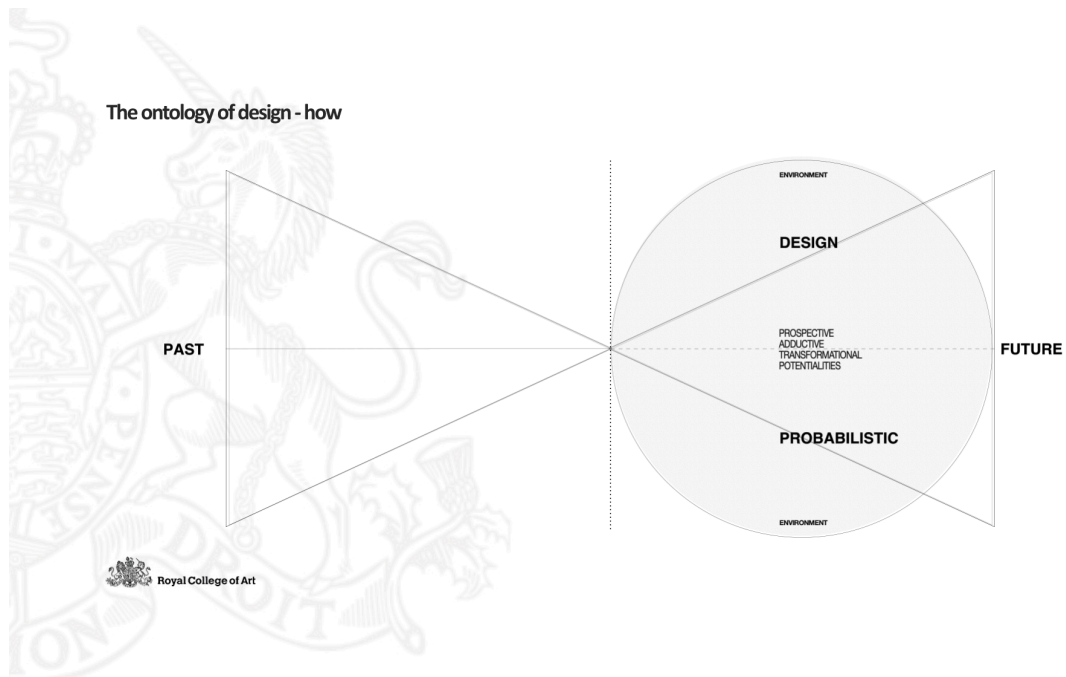
Type to enter a caption.



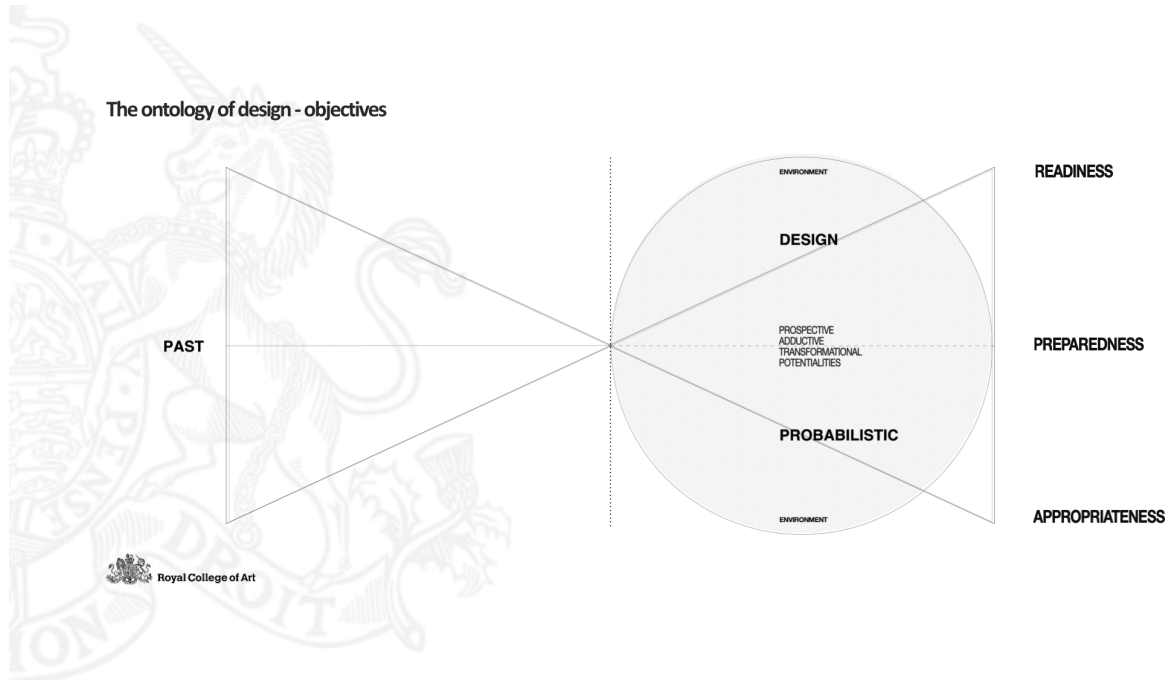
Type to enter a caption.



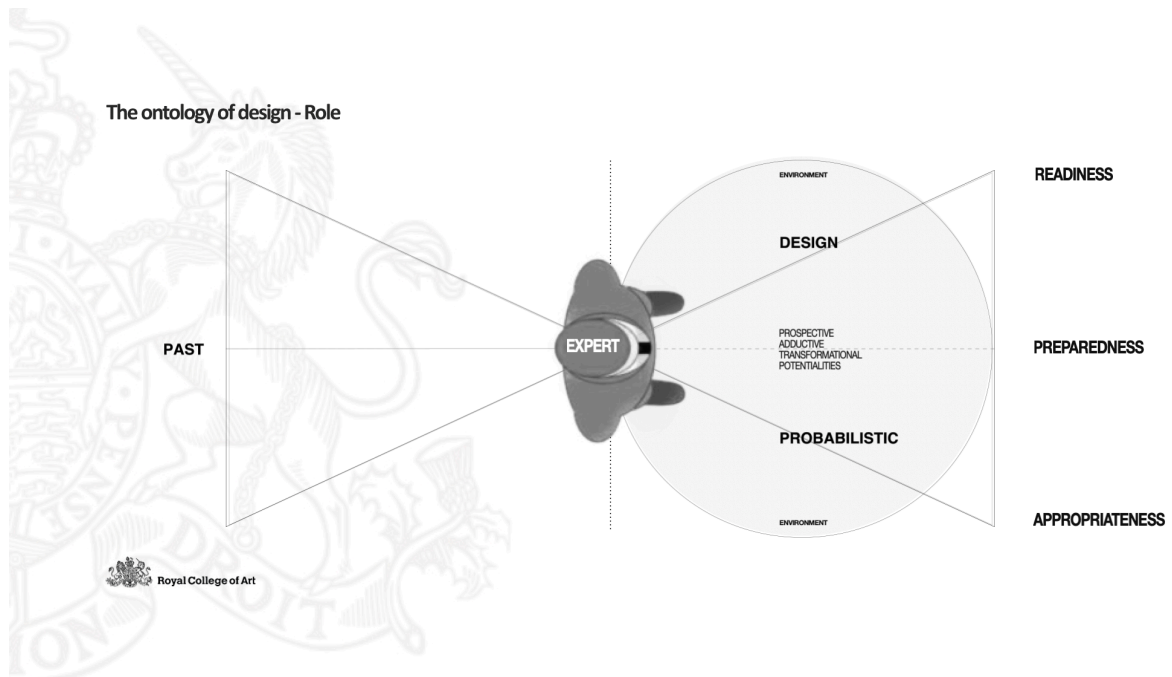
Type to enter a caption.



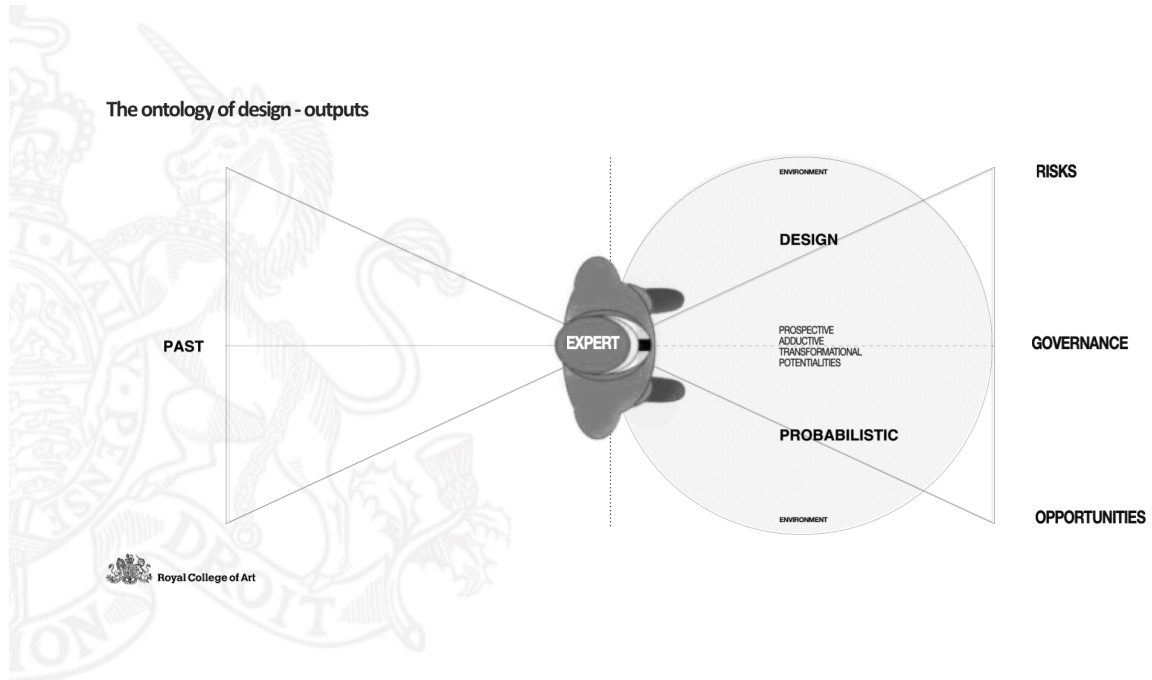
Type to enter a caption.



Type to enter a caption.



Type to enter a caption.



Type to enter a caption.

Research

Conclusions

- 1 Prospective design is defined as the capacity to make involving abductive reasoning to 'go beyond' what exists and propose what can be 'otherwise'.
- 2 Probabilistic knowledge in the context of design could be defined as the potential impact of transformational initiatives.
- 3 The projected potentiality is placed in a society yet-to-be or not-fully-formed. Therefore it cannot be precisely measured or described as it does not fully exist. Therefore, it can only be probabilistic.
- 4 In this context, we trade some degrees of accuracy to access areas yet-to-be or not-fully-formed, therefore research is always preliminary in its nature.
- 5 It is concerned with competing standards of value rather than securing boundaries of knowledge and its practice is based on the capacity to prospect readiness, appropriateness and preparedness.
- 6 It repositions the role of the designer from a facilitator to an expert in prospective future-led transformational developments to enhance transformations and reduce future risks.

Royal College of Art

Slide / 38

Type to enter a caption.

