Name: Era Sarda

Email id: erasarda2024@gmail.com

The decision-making process involved seven main criteria, each with sub-criteria, making it challenging to assign weights and select scoring methods for five topics. While objective factors like the problem's scale and funding attention were available, it was difficult to generalize these factors globally due to a lack of comprehensive statistical reports. Thus, I often limited the scope of analysis to developing countries like India.

Subjective elements also played a significant role, such as personal interest, potential consequences, and whether the impact would be short- or long-term. Scores were assigned on a 0 to 1 scale, where higher scores indicated a greater need for attention or an easier problem to solve, while negative scores accounted for potential adverse outcomes. These scores were relative, not absolute, as incorporating z-scores was impractical given subjective uncertainties.

The highest score went to “Large-scale workshops on Building Self-awareness and Mental Resilience.” I am particularly motivated to focus on issues such as child abuse, cyber abuse, deepfake prevention, and enhancing women’s safety. I believe these areas require an interdisciplinary approach involving AI alignment, AI governance, promoting mental health (since childhood traumas often correlates with criminal behavior), moral education for children from an early age, and increasing employment opportunities.

Sources:

- Wikipedia
- UNICEF Website
- Resources provide by Thinking Competition

Weighted Factor Model

S.no.				1	2	3	4	5
			Intervention	Research on Aligning AI systems with human intent	Research on Novel Frameworks for International AI Governance	Large-scale workshops on Building Self-awareness and Mental Resilience - Women Safety and Child abuse prevention and mental health	Funding Apprenticeships for underprivileged youth through Conditional Cash Transfers	Teacher Training Programs in Underserved Regions
			Core Objective	Minimising risks from intelligent autonomous systems to ensure they are safe and reliable	Improving global coordination on response and mitigation of damage from intelligent autonomous systems	Improving access to mental health support communities (in-turn this can in longterm contribute to women safety and child abuse prevention)	Universal access to quality career skills training (increase in employment can inturn lead to reduced crimes like murder or rape)	Raising the bar of education delivery standards universally (including moral education which is necessary for long term wellbeing of society)
	Weight of criteria	Criterias for Comparison		Score				
1)	0.5	Personal fit - How easy it is for me to enter the field	Background	0.9	0.8	0.1	0	0.9
			How easily i can educate/train myself	0.9	1	0.9	0.5	0.7
			Personal interest (2.0 weight)	0.65	0.8	1	0.6	0.9
2)	0.9	Scale of the problem		1	1	1	0.8	0.9
3)	0.7	Longterm/Shortterm - How fast we may observe the results	5/10/20/40 years or Entire generation	0.6	0.8	0.6	0.7	0.8
4)	0.8	Neglectedness-How much funds/efforts currently it is receiving		0.9	0.9	1	0.7	0.7
5)	0.7	execution difficulty of the intervention	direct govt intervention	1	0.4	0.5	0.6	0.3
			recipient easyness	0.5	0.7	0.3	0.5	0.7
			mafia interventions	1	0.6	0.3	0.8	0.9
			targeted work/widespread work	1	0.7	0.3	0.3	0.3
6)	0.8	stigma and skepticism around the problem		0.1	0.1	1	0.5	0.8
7)	1	consequences of efforts	positive	0.5	0.4	0.7	0.4	0.5
			negative	-0.5	-0.6	-0.3	-0.6	-0.5
			Total Weight	**6.77**	**6.24**	**6.8**	**4.96**	**6.71**

img.

Name: Era Sarda

email: erasarda2024@gmail.com

WEIGHTED FACTOR MODEL COMPREHENSIVE

Thinking Competition

S.no				1		2		3		4	
			Intervention	Research on Aligning AI systems with human intent		Research on Novel Frameworks for International AI Governance		Large-scale workshops on Building Self-awareness and Mental Resilience		Funding Apprenticeships for underprivileged youth through Conditional Cash Transfers	
			Core Objective	Minimising risks from intelligent autonomous systems to ensure they are safe and reliable		Improving global coordination on response and mitigation of damage from intelligent autonomous systems		Improving access to mental health support communities *(in-turn this can in longterm contribute to women safety and child abuse prevention)*		Universal access to quality career skills training *(increase in employment can inturn lead to reduced crimes like murder or rape)*	
	Weight of criteria	**Criterias for Comparison**		**Description**	**Score**	**Description**	**Score**	**Description**	**Score**	**Description**	**Score**
1)	0.5	Personal fit - How easy it will for me to enter the field	Background	Relevant background in AI and CS	0.9	Relevant background in AI and CS, but not in policy related work	0.8	Need some background in psychology/criminal psychology/management	0.1	Need some background in psychology/public affairs/management/fund-raising	0
			How easily i can educate/train myself		0.9	With some efforts towards policy projects as more inclined towards policy work than technical work	1	By volunteering and exposing myself to do the field	0.9	Need experience	0.5
			Personal interest (2.0 weight)		0.65		0.8		1		0.6
2)	0.9	Scale of the problem			1		1	several households with financial and health related crisis; unemployed people; Abuse suffered women and children; Children in care	1	5 crore in just India	0.8
3)	0.7	Longterm/Shortterm - How fast we observe the results	5/10/20/40 years or Entire generation	Longterm benefit but slow results (it may take 5 or more years to reach that level of research)	0.6	Short as well as longterm	0.8	Longterm benefit but slow results (maybe 10-20 years)	0.6	Short as well as long term (5-10 years)	0.7
4)	0.8	Neglectedness-How much funds/efforts currently it is receiving		Not paid much attention to in developing countries in comparison to developed countries	0.9	Not paid much attention to in developing countries in comparison to developed countries	0.9	Not paid much attention to in developing countries in comparison to developed countries	1	Even if not much neglected, results are not very obvious	0.7
5)	0.7	Execution difficulty of the intervention	direct govt intervention	NA	1	Highly Required	0.4	Required	0.5	Required	0.6
			recipient easyness	will corporate companies incorporate the alignment research	0.5	corporate companies should incorporate rules, without any corruption	0.7	Stigma around the issue makes it difficult, and the results are very subjective to different background and	0.3	Youths even after attending training programs, mayn't utilize their learned skills	0.5
			mafia interventions	NA	1	may or may not	0.6	High chances of their intervention	0.3	may not intervene	0.8
			targeted work/widespread work	Widespread:research can be used around the globe	1	Targeted:policies may vary from region to region	0.7	Targeted: to regions	0.3	Targeted: to regions	0.3
6)	0.8	Stigma/Skepticism around the problem		Emerging research field	0.1	Emerging governance field	0.1	Lot of stigma in developing countries like India	1	Doubts on output of training uneducated youths, and even on how good these programs actually are.	0.5
7)	1	Consequences of putting our efforts	positive	Research may be able to output significant results	0.5	Necesarry important restrictions put	0.4	Mental health, happiness, and overall life satisfaction can increase, which may lead to lowering of crime rates as well; Exposure of abusive people	0.7	Increased employment; reducing (even if not significant amount) of crime rates	0.4
			negative	Research may not be able to output significant results	-0.5	Corruption, Biasedness possible	-0.6	People get more triggered of their traumas; improper and untrained mental health supoort; exposure of abusive but dangerous people	-0.3	Misuse of the cash transfers	-0.6
			Total Weight		**6.77**		**6.24**		**6.8**		**4.96**

5	
Teacher Training Programs in Underserved Regions	
Raising the bar of education delivery standards universally *(including moral education which is necessary for long term wellbeing of society)*	
Description	**Score**
Have technical background	0.9
Need some background in public policy/teacher training programs,	0.7
	0.9
7 crore in just India	0.9
Short as well as longterm benefits (can be seen under 5 years)	0.8
Even if not much neglected, results are not very obvious	0.7
Highly required, specially in government schools	0.3
the teacher to student transfer of knowledge may or may not work even after training	0.7
may not intervene	0.9
Targeted: to regions	0.3
doubts on how receptive teachers are even for the program	0.8
Well-trained teachers educate their students well	0.5
the teacher to student transfer of knowledge may or may not work even after training	-0.5
	6.71

Simplified Wighted Factor Model

S.no				1	2	3	4	5
			Intervention	Research on Aligning AI systems with human intent	Research on Novel Frameworks for International AI Governance	Large-scale workshops on Building Self-awareness and Mental Resilience - Women Safety and Child abuse prevention and mental health	Funding Apprenticeships for underprivileged youth through Conditional Cash Transfers	Teacher Training Programs in Underserved Regions
			Core Objective	Minimising risks from intelligent autonomous systems to ensure they are safe and reliable	Improving global coordination on response and mitigation of damage from intelligent autonomous systems	Improving access to mental health support communities	Universal access to quality career skills training	Raising the bar of education delivery standards universally
	Weight of criteria	**Criterias for Comparison**		Score				
1)	0.5	Personal fit - How easy it is for me to enter the field	Background	0.9	0.8	0.1	0	0.9
			How easily i can educate/train myself	0.9	1	0.9	0.5	0.7
			Personal interest (2.0 weight)	0.65	0.8	1	0.6	0.9
2)	0.9	Scale of the problem		1	1	1	0.8	0.9
3)	0.7	Longterm/Shortterm - How fast we may observe the results	5/10/20/40 years or Entire generation	0.6	0.8	0.6	0.7	0.8
4)	0.8	Neglectedness-How much funds/efforts currently it is receiving		0.9	0.9	1	0.7	0.7
5)	0.7	execution difficulty of the intervention	direct govt intervention	1	0.4	0.5	0.6	0.3
			recipient easyness	0.5	0.7	0.3	0.5	0.7
			mafia interventions	1	0.6	0.3	0.8	0.9
			targeted work/widespread work	1	0.7	0.3	0.3	0.3
6)	0.8	stigma and skepticism around the problem		0.1	0.1	1	0.5	0.8
7)	1	consequences of efforts	positive	0.5	0.4	0.7	0.4	0.5
			negative	-0.5	-0.6	-0.3	-0.6	-0.5
			Total Weight	**6.77**	**6.24**	**6.8**	**4.96**	**6.71**