

Ethics and Algorithms Toolkit (Beta)

(Section) Worksheet for Part 1

Overview of toolkit beta release

Welcome to the beta release of our Ethics and Algorithms Toolkit! This toolkit is designed to help governments (and others) use algorithms responsibly.

Who is this toolkit for?	If you are building or acquiring algorithms in the government sector this toolkit is for you. Though we expect others will find it useful.
What is the toolkit?	The toolkit is really a process. It walks you through a series of questions to help you 1) understand the ethical risks posed by your use of an algorithm and then 2) identify what you can do to minimize those ethical risks.
What are the parts of the toolkit?	The toolkit comes in several parts: <ol style="list-style-type: none">1. The introduction to the toolkit (this document)2. Part 1: Assess Algorithm Risk3. Part 2: Manage Algorithm Risk4. Appendices (including a handy worksheet)
Who made this toolkit?	The beta release was a collaboration between The Center for Government Excellence (GovEx) at Johns Hopkins, the City and County of San Francisco, Harvard DataSmart, and Data Community DC.
How can I give feedback?	Send feedback on our beta release at http://labs.centerforgov.org/toolkit/ .

Part 1: Worksheet to Assess Algorithm Risk

Step 1: Understand and assess impact																									
Step 1.1 Describe the impact	Step 1.1.1 Identify who or what will be impacted:																								
	Step 1.1.2 Identify the types of impact <ul style="list-style-type: none"> <input type="checkbox"/> Access to goods, benefits or services <input type="checkbox"/> Financial <input type="checkbox"/> Property or equipment <input type="checkbox"/> Reputation <input type="checkbox"/> Emotional <input type="checkbox"/> Life / safety <input type="checkbox"/> Privacy <input type="checkbox"/> Liberty / freedom <input type="checkbox"/> Rights / intellectual Property 																								
Step 1.2 Assess scope of impact	Step 1.2.1 Rate the degree of impact <ul style="list-style-type: none"> <input type="checkbox"/> No discernable <input type="checkbox"/> Minor <input type="checkbox"/> Moderate <input type="checkbox"/> Major 																								
	Step 1.2.2 Estimate the scale of impact <ul style="list-style-type: none"> <input type="checkbox"/> Small <input type="checkbox"/> Medium <input type="checkbox"/> Large 																								
	Step 1.2.3 Assign scope estimate <ul style="list-style-type: none"> <input type="checkbox"/> Very narrow <input type="checkbox"/> Limited/Narrow <input type="checkbox"/> Substantial <input type="checkbox"/> Broad/wide ranging <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th colspan="2" rowspan="2">Scope Estimate</th> <th colspan="3">Scale of Impact</th> </tr> <tr> <th>Small</th> <th>Medium</th> <th>Large</th> </tr> </thead> <tbody> <tr> <td rowspan="4">Degree of Impact</td> <td>No discernable</td> <td>Very narrow</td> <td>Very narrow</td> <td>Limited/Narrow</td> </tr> <tr> <td>Minor</td> <td>Very narrow</td> <td>Limited/Narrow</td> <td>Substantial</td> </tr> <tr> <td>Moderate</td> <td>Limited/Narrow</td> <td>Substantial</td> <td>Broad/wide ranging</td> </tr> <tr> <td>Major</td> <td>Substantial</td> <td>Broad/wide ranging</td> <td>Broad/wide ranging</td> </tr> </tbody> </table>	Scope Estimate		Scale of Impact			Small	Medium	Large	Degree of Impact	No discernable	Very narrow	Very narrow	Limited/Narrow	Minor	Very narrow	Limited/Narrow	Substantial	Moderate	Limited/Narrow	Substantial	Broad/wide ranging	Major	Substantial	Broad/wide ranging
Scope Estimate				Scale of Impact																					
		Small	Medium	Large																					
Degree of Impact	No discernable	Very narrow	Very narrow	Limited/Narrow																					
	Minor	Very narrow	Limited/Narrow	Substantial																					
	Moderate	Limited/Narrow	Substantial	Broad/wide ranging																					
	Major	Substantial	Broad/wide ranging	Broad/wide ranging																					
Step 1.3 Estimate the overall direction of impact	<ul style="list-style-type: none"> <input type="checkbox"/> Positive <input type="checkbox"/> Mostly Positive <input type="checkbox"/> Mostly Negative <input type="checkbox"/> Negative 																								
Step 1.4 Assign overall impact risk	Combine Step 1.2 and 1.3 to assign overall impact risk <ul style="list-style-type: none"> <input type="checkbox"/> Very low <input type="checkbox"/> Low <input type="checkbox"/> Substantial 																								

		<input type="checkbox"/> Significant <input type="checkbox"/> High <input type="checkbox"/> Extreme																															
		<table border="1"> <thead> <tr> <th colspan="2" rowspan="2">Overall Impact Risk</th> <th colspan="4">Overall Direction</th> </tr> <tr> <th>Positive</th> <th>Mostly Positive</th> <th>Mostly Negative</th> <th>Negative</th> </tr> </thead> <tbody> <tr> <th rowspan="4">Scope</th> <td>Very Narrow</td> <td>Very low</td> <td>Very low</td> <td>Low</td> <td>Moderate</td> </tr> <tr> <td>Limited/Narrow</td> <td>Very low</td> <td>Low</td> <td>Moderate</td> <td>Significant</td> </tr> <tr> <td>Substantial</td> <td>Low</td> <td>Moderate</td> <td>Significant</td> <td>High</td> </tr> <tr> <td>Broad/wide ranging</td> <td>Moderate</td> <td>Significant</td> <td>High</td> <td>Extreme</td> </tr> </tbody> </table>	Overall Impact Risk		Overall Direction				Positive	Mostly Positive	Mostly Negative	Negative	Scope	Very Narrow	Very low	Very low	Low	Moderate	Limited/Narrow	Very low	Low	Moderate	Significant	Substantial	Low	Moderate	Significant	High	Broad/wide ranging	Moderate	Significant	High	Extreme
Overall Impact Risk		Overall Direction																															
		Positive	Mostly Positive	Mostly Negative	Negative																												
Scope	Very Narrow	Very low	Very low	Low	Moderate																												
	Limited/Narrow	Very low	Low	Moderate	Significant																												
	Substantial	Low	Moderate	Significant	High																												
	Broad/wide ranging	Moderate	Significant	High	Extreme																												

Step 2: Assess appropriate data use risk

Step 2.1 Rate consistency and compatibility of use	<input type="checkbox"/> Yes <input type="checkbox"/> Somewhat <input type="checkbox"/> Unknown <input type="checkbox"/> No																									
Step 2.2 Rate reputation and perception from use	<input type="checkbox"/> Supportive <input type="checkbox"/> Mixed <input type="checkbox"/> Not supportive																									
Step 2.3 Assign appropriate use risk score	<p>Combine Step 2.1 and 2.2 to assign appropriate use risk score</p> <input type="checkbox"/> Low <input type="checkbox"/> Medium <input type="checkbox"/> High																									
<table border="1"> <thead> <tr> <th colspan="2" rowspan="2">Appropriate Use Risk Score</th> <th colspan="3">Reputation and Perception</th> </tr> <tr> <th>Supportive</th> <th>Mixed</th> <th>Not Supportive</th> </tr> </thead> <tbody> <tr> <th rowspan="4">Consistency and Compatibility</th> <td>Yes</td> <td>Low</td> <td>Low</td> <td>Medium</td> </tr> <tr> <td>Somewhat</td> <td>Low</td> <td>Medium</td> <td>High</td> </tr> <tr> <td>Unknown</td> <td>Medium</td> <td>Medium</td> <td>High</td> </tr> <tr> <td>No</td> <td>Medium</td> <td>High</td> <td>High</td> </tr> </tbody> </table>		Appropriate Use Risk Score		Reputation and Perception			Supportive	Mixed	Not Supportive	Consistency and Compatibility	Yes	Low	Low	Medium	Somewhat	Low	Medium	High	Unknown	Medium	Medium	High	No	Medium	High	High
Appropriate Use Risk Score				Reputation and Perception																						
		Supportive	Mixed	Not Supportive																						
Consistency and Compatibility	Yes	Low	Low	Medium																						
	Somewhat	Low	Medium	High																						
	Unknown	Medium	Medium	High																						
	No	Medium	High	High																						

Step 3: Assess accountability risk

Step 3.1 Determine automation score	<input type="checkbox"/> Low - human mediated <input type="checkbox"/> Medium - algorithm mediated <input type="checkbox"/> High - algorithmically determined
Step 3.2 Determine accessibility score	<p>Step 3.2.1 Determine explainability score</p> <input type="checkbox"/> Easy <input type="checkbox"/> Medium <input type="checkbox"/> Hard
	<p>Step 3.2.2 Determine auditability score</p> <input type="checkbox"/> Easy <input type="checkbox"/> Medium <input type="checkbox"/> Hard
	<p>Step 3.2.3 Assign accessibility score: Combine 3.2.1 and 3.2.1 to assign</p> <input type="checkbox"/> Accessible <input type="checkbox"/> Some concerns <input type="checkbox"/> Major concerns

Accessibility Score		Explainability		
		Easy	Medium	Hard
Auditability	Easy	Accessible	Accessible	Some concerns
	Medium	Accessible	Some concerns	Major concerns
	Hard	Some concerns	Major concerns	Major concerns

Step 3.3 Assign accountability risk	Combine Step 3.1 and 3.2 to assign accountability risk			
	<input type="checkbox"/> Low <input type="checkbox"/> Medium <input type="checkbox"/> High			
Accountability Risk		Automation Score		
		Low - human mediated	Medium - algorithm mediated	High - algorithmically determined
Accessibility Score	Accessible	Low	Low	Medium
	Some concerns	Low	Medium	High
	Major concerns	Medium	High	High

Step 4: Assess third party methodology risk

Step 4.1 Answer third party methodology questions	Enter number of points:
---	-------------------------

Step 4.2 Assign third party methodology risk level	<input type="checkbox"/> Low <input type="checkbox"/> Medium <input type="checkbox"/> High								
	<table border="1"> <thead> <tr> <th>Total Points</th> <th>Third Party Risk</th> </tr> </thead> <tbody> <tr> <td>6-8</td> <td>Low</td> </tr> <tr> <td>3-5</td> <td>Medium</td> </tr> <tr> <td>0-2</td> <td>High</td> </tr> </tbody> </table>	Total Points	Third Party Risk	6-8	Low	3-5	Medium	0-2	High
Total Points	Third Party Risk								
6-8	Low								
3-5	Medium								
0-2	High								

Step 5: Assess risk of historic bias

Step 5: Assess risk of historic bias	<input type="checkbox"/> Low <input type="checkbox"/> Medium <input type="checkbox"/> High
--------------------------------------	--

Step 6: Assess risk of technical bias

Step 6.1 Assess representativeness (sample) and inaccuracy risk	Step 6.1.1 Assess representativeness risk
	Step 6.1.2 Assess inaccuracy risk
	Step 6.1.3 Assign representativeness and inaccuracy risk score

Low
 Medium
 High

	<table border="1"> <tr> <th colspan="2" rowspan="2">Representativeness and Inaccuracy Risk Score</th> <th colspan="3">Representativeness Risk</th> </tr> <tr> <th>Low</th> <th>Medium</th> <th>High</th> </tr> <tr> <th rowspan="3">Inaccuracy Risk</th> <th>Low</th> <td>Low</td> <td>Low</td> <td>Medium</td> </tr> <tr> <th>Medium</th> <td>Low</td> <td>Medium</td> <td>High</td> </tr> <tr> <th>High</th> <td>Medium</td> <td>High</td> <td>High</td> </tr> </table>	Representativeness and Inaccuracy Risk Score		Representativeness Risk			Low	Medium	High	Inaccuracy Risk	Low	Low	Low	Medium	Medium	Low	Medium	High	High	Medium	High	High
Representativeness and Inaccuracy Risk Score				Representativeness Risk																		
		Low	Medium	High																		
Inaccuracy Risk	Low	Low	Low	Medium																		
	Medium	Low	Medium	High																		
	High	Medium	High	High																		
Step 6.2 Assign risk from scope of training data	<p>Step 6.2.1 Determine the actual source of training data</p> <ul style="list-style-type: none"> <input type="checkbox"/> Local <input type="checkbox"/> Non-local <p>Step 6.2.2 Determine the desired source of training data</p> <ul style="list-style-type: none"> <input type="checkbox"/> Local <input type="checkbox"/> Non-local <p>Step 6.2.3 Assign training risk score</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Low <input type="checkbox"/> High <table border="1"> <tr> <th colspan="2" rowspan="2">Training Risk</th> <th colspan="2">Desired Source</th> </tr> <tr> <th>Local</th> <th>Non-local</th> </tr> <tr> <th rowspan="2">Actual Source</th> <th>Local</th> <td>Low</td> <td>High</td> </tr> <tr> <th>Non-local</th> <td>High</td> <td>Low</td> </tr> </table>	Training Risk		Desired Source		Local	Non-local	Actual Source	Local	Low	High	Non-local	High	Low								
Training Risk				Desired Source																		
		Local	Non-local																			
Actual Source	Local	Low	High																			
	Non-local	High	Low																			
Step 6.3 Assign methodology risk	<p>Combine Step 4.2 and Step 6.2.3 to arrive at a methodology risk score</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Low <input type="checkbox"/> Medium <input type="checkbox"/> High <table border="1"> <tr> <th colspan="2" rowspan="2">Methodology Risk</th> <th colspan="2">Training Risk Score</th> </tr> <tr> <th>Low</th> <th>High</th> </tr> <tr> <th rowspan="3">Third Party Methodology Risk (step 4.2)</th> <th>Low</th> <td>Low</td> <td>Medium</td> </tr> <tr> <th>Medium</th> <td>Low</td> <td>High</td> </tr> <tr> <th>High</th> <td>Medium</td> <td>High</td> </tr> </table>	Methodology Risk		Training Risk Score		Low	High	Third Party Methodology Risk (step 4.2)	Low	Low	Medium	Medium	Low	High	High	Medium	High					
Methodology Risk				Training Risk Score																		
		Low	High																			
Third Party Methodology Risk (step 4.2)	Low	Low	Medium																			
	Medium	Low	High																			
	High	Medium	High																			
Step 6.4 Assign the overall risk of technical bias	<p>Combine your risk scores from Step 6.1.3 and Step 6.3 to assign overall risk of technical bias</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Low <input type="checkbox"/> Medium <input type="checkbox"/> High <table border="1"> <tr> <th colspan="2" rowspan="2">Overall Technical Bias Risk</th> <th colspan="3">Methodology Risk</th> </tr> <tr> <th>Low</th> <th>Medium</th> <th>High</th> </tr> <tr> <th rowspan="3">Representativeness and Inaccuracy Risk Score</th> <th>Low</th> <td>Low</td> <td>Low</td> <td>Medium</td> </tr> <tr> <th>Medium</th> <td>Low</td> <td>Medium</td> <td>High</td> </tr> <tr> <th>High</th> <td>Medium</td> <td>High</td> <td>High</td> </tr> </table>	Overall Technical Bias Risk		Methodology Risk			Low	Medium	High	Representativeness and Inaccuracy Risk Score	Low	Low	Low	Medium	Medium	Low	Medium	High	High	Medium	High	High
Overall Technical Bias Risk				Methodology Risk																		
		Low	Medium	High																		
Representativeness and Inaccuracy Risk Score	Low	Low	Low	Medium																		
	Medium	Low	Medium	High																		
	High	Medium	High	High																		