Sweet Summer Child Score

Definitions

Automated decision system	Any techno-social system which automates the classification, prediction or allocative decisions about a cohort of users (could be fixed business logic, linear regression, ML, DL). Note that full automation is not required (for example, your system could make a prediction and an account manager could action that prediction). For our purposes we assume there's a strong causal link between the output of the technical system and the resulting action or intervention.
Customers	the people who pay for your automated decision system
End users	the people interacting with your technology (app, website, etc)
Target cohort	the people that your automated decision system makes predictions about. Note that these may or may not correlate to end-users, customers, or system operators.
Individuals	individual people who are part of your target cohort.

Overview

What is it?

Sweet Summer Child Score (SSCS) is a scoring mechanism for latent risk. It will help you quickly and efficiently scan for the possibility of harm to *people* and *communities* by a socio-technical system. Note that harms to *animals* and the *environment* are not considered.

SSCS takes a step back away from the specifics of your technology and looks at environmental, systems and human factors. It's intended to be used at the ideation part of your process, to critique product ideas or compare competing features. But SSCS can also be used post-hoc to critique an existing system, or propose approaches to harm mitigation.

How long does it take?

If you're familiar with the automated decision system it should take 8-10 minutes to complete. Respond with your gut, don't overthink it too much.

Critiquing proposals/ideas

You can absolutely use the score to critique ideas that aren't developed yet. You need to have sufficiently sketched it out and have the rough shape & feature set of the proposal or idea.

How is it calculated?

You start with a score of 99 (the highest score possible, to acknowledge the impossibility of building and deploying something perfectly fair, or perfectly accurate and bug-free). As you answer questions, points are deducted from your score. The sections are worth roughly 1/3 each. The multiplier score will reduce the overall cost of points deductions in-line with the scope of your system overall.

Who can use it?

This is designed for use by data scientists, product managers, and other technologists. However it's simple enough that end-users, journalists and auditors can use it as a form of socio-technical critique. This score is model and stack agnostic, so you don't need to see 'under-the-hood' to use it.

How should you use it?

This is a qualitative assessment and relies on you to complete it honestly. There's no point if you're just going to give yourself the best score! Without discussing first, have multiple people in your team complete the quiz. How close were your scores? Which areas did you agree and disagree? Discuss and attempt to resolve the differences, especially where there are answers that are 2 or more steps away from each other.

Why don't we scan for minority identities?

We focus on the *disadvantages* usually experienced by these populations rather than the demographics of disadvantage. We don't want to contribute to the conflation of minority identities and harm.

Section #0 - Multiplier

1) How large is the target cohort for your decision system?

The people your system makes decisions, predictions or classifications about.

Note: include the entire live data set, not the subset of people who might experience a bad outcome. If it changes year-on-year, estimate for this year. If you haven't launched yet, give it your best guess.

- A. 1-100
- B. 101-1,000
- C. 1,000-10,000
- D. 10,000-100,000
- E. 100,000-1,000,000
- F. 1,000,000-10,000,000
- G. 10 million +

2) How is your target cohort distributed geographically?

- A. Rural/Remote
- B. Small-Medium Town
- C. Large Town - City
- D. State
- E. National
- F. 2-3 Countries
- G. Global

SECTION #1

Cohort scan

1) Is your target cohort economically vulnerable?
Look at average earning rates of your country if you aren't sure where your demographic sits overall.

    A. Top 10% of earners
    B. Less vulnerable than average
    C. Average
    D. More vulnerable than average
    E. Bottom 10% of earners

2) Is your target cohort psychologically vulnerable?
Are they likely to suffer from mental health issues, self-harm, or to have their capacity to work, study, or pursue their goals impacted by debilitating mental illness?

    A. Very robust
    B. Less vulnerable than average
    C. Average
    D. More vulnerable than average
    E. Highly vulnerable

3) Is your target cohort vulnerable to coercive control?
Think parents controlling LGBTQI children. Jaime Spears. Elder abuse. Intimate partner abuse. Technological coercion or control.

    A. Not vulnerable
    B. Less vulnerable than average
    C. Average
    D. More vulnerable than average
    E. Extremely vulnerable

4) Is your target cohort over-policed or over-surveilled?
Are they more likely than the average population in any aspect of their lives to be targeted by police, social services or surveillance systems?

A. Very unlikely
B. Less likely than average
C. Average
D. More likely than average
E. Very likely

5) Is your target cohort technologically vulnerable?

Will they have more difficulty navigating technological systems than average? Poor digital literacy? Do they have a higher rate of visual impairments or other disabilities which would compound this?

A. Top 10% of computer users
B. Less vulnerable than average
C. Average
D. More vulnerable than average
E. Bottom 10% of computer users

6) Internet access?

Is your target cohort in a remote area with poor infrastructure? Exposed to download or over-limit fees? Access on 1 or more devices? Internet sharing or have own service? Physically at-risk if they try to access public internet? Are they accessing the internet from cellular networks, NBN, ADSL, Broadband, other?

A. Top 10%
B. Better access than average
C. Average
D. Poorer access than average
E. Bottom 10%

SECTION #2

System scan

7) Does your decision system operate as part of a competitive marketplace, or are you the only provider of the specific service, platform, resource or goods?
*Looking for **viable** alternatives. Competitors need to be known to your target cohort. Consider network effects, brand visibility, feature parity, etc. If competitors exist but consumers can't find them, or if the cost of adoption is very high, select 'effective monopoly'.*

Note: government departments act as an effective monopoly, as people cannot meaningfully shop-around or take their business elsewhere.

    A. Effective monopoly / Gov't service
    B. 1 competitor
    C. 2 competitors
    D. 3-4 competitors
    E. 5+ competitors

8A) Is your decision system mandatory or optional?
Mandatory systems will be mostly government applications; paying taxes, rent, complying with judicial hearings or prison sentences, etc.

    A. Mandatory
    B. Optional

<!If Optional>

8B) Is your decision system opt-in or opt-out?
Does the target cohort actively choose to participate or add themselves? Or are they included by default and have to actively remove themselves?

    A. Opt-in
    B. Opt-out

<!If Opt-out>

8C) How visible or discoverable is it that the target cohort has been opted-into the system?
Do you notify them, is the operation of your system visible, or do they have to find it in a setting or lodge their own query?

    A. Very hard to find
    B. Somewhat hard to find
    C. Average
    D. Discoverable
    E. Very discoverable

8D) How difficult is it to opt out?

    A. Very difficult
    B. Somewhat difficult
    C. Average
    D. Straight forward
    E. Very easy

8E) How perpetual is opting out?
Does the user preference remain for system upgrades or migrations, normal use, etc?

    A. Lasts forever
    B. Lasts forever *(but periodically asked to opt in)*
    C. Reset periodically

<!EndIf opt-out>
<!EndIf Optional>

9) Can you turn your decision system off *immediately* if something goes wrong?
Can you do this with one click (Y), or do you have to deploy code to take it down (N)?

    A. Y
    B. N

10) Does your system always output a classification, prediction, or decision?

Or does it have an option to notify a system operator to review when confidence is low or it's confused? Eg decision hand-off

    A. Y
    B. N

11) Do you provide a record of the system decision for individuals, including a unique ID that can be traced in your system log?

For their reference and for looking up that decision later on. This could be on-screen, email, PDF, etc.

    A. Y
    B. N

12A) Does your decision system collect meaningful consent, including a clear explanation of the uses of data and ramifications of providing consent?

Note if your consent request is tokenistic, confusing or deliberately misleading you should answer No

    A. Y
    B. N

<!If Yes>

12B) Have you performed usability or other user testing to assess the comprehensibility of your consent section?

    A. Y
    B. N

12C) Do you provide a mechanism to allow individuals to revoke consent to use their data?

    A. Y
    B. N

<!EndIf>

13A) Do you provide any visual or text-based explanation of your automated system decision on the decision interface?

This could be in the form of displaying results for similar candidates, or identifying which feature had a strong impact on their outcome.

    A. Y
    B. N

<!If No>

13B) Can your target cohort request an explanation of the decision they received?

If they don't understand or agree with the automated decision, is there an existing channel or feedback mechanism for them to request explanation? Are explanations planned and available (e.g. you've designed or planned for them before they are requested)? Note this doesn't include generic feedback mechanisms.

    A. Y
    B. N

<!EndIf>

14A) Can your target cohort request a review of the decision they received?

If they think it looks wrong or might be biased or otherwise unfair, can they flag this concern with you and trigger a decision review? This should be a use-case which has been designed and planned, with prepared approaches and business logic. Decision review must be available in the same channel as they received the decision. If they need to use a generic support channel the answer is "No".

    A. Y
    B. N

<!If Yes>

14B) What is the probability that decision review will reverse or change the decision?

    A. Unlikely

B. No different than chance
C. Likely

14C) On-average *how much effort* is required by your target cohort to lodge a request for decision review?
This considers total time spent on all channels to explain, lodge the request, discuss the results, etc.

    A. 1-2 hours
    B. 1 day
    C. Longer

14D) On-average *how long* will your decision review process take?

    A. Instantaneous (machine review)
    B. 1 business day
    C. Longer

<!EndIf>

SECTION #3 - 30 points

Maslov's Hierarchy of needs - harm scan

*Slightly updated for the modern era. For this section, we compare the likelihood of harm with no intervention (if your decision system didn't exist) to the likelihood of this harm occurring as a result of your intervention. So if your target cohort are all drinking toxic water and your system reduces the number of people getting toxic water, you're **reducing** the probability of this harm because of the high base rate.*

15) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Physiological needs*?
(Air, water, food, shelter, sleep, reproduction)

A. 10% Very unlikely
B. 33% Unlikely
C. 50% (no different than chance)
D. 66% Likely
E. 90% Very likely

16) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Safety needs*?

(Personal security, employment, resources, health, property, internet access, information access, information security)

A. 10% Very unlikely
B. 33% Unlikely
C. 50% (no different than chance)
D. 66% Likely
E. 90% Very likely

17) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Love and belonging needs*?

(Friendship, intimacy, family, sense of connection, access to online communities)

A. 10% Very unlikely
B. 33% Unlikely
C. 50% (no different than chance)
D. 66% Likely
E. 90% Very likely

18) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Esteem needs*?

(Respect, self-esteem, status, recognition, strength, freedom)

A. 10% Very unlikely
B. 33% Unlikely
C. 50% (no different than chance)
D. 66% Likely
E. 90% Very likely

19) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's _Self-actualisation needs_?

(Desire to become the most one can be)

    A. 10% Very unlikely
    B. 33% Unlikely
    C. 50% (no different than chance)
    D. 66% Likely
    E. 90% Very likely

Calculate your score

`99 - (Points Sum * multiplier)`

For example, if your total points add up to 69, and your multiplier is 0.85:
`99 - (69 * 0.85) = 40.35`

Section #0 - Multiplier

This section constructs a multiplier value which is used against the sum of all your point scores for the whole quiz. This helps us reduce the assessed risk where, for example, the number of people affected might be very small and distributed.

1) How large is the target cohort for your decision system?

    A. 1-100 (0.52)
    B. 101-1,000 (0.75)
    C. 1,000-10,000 (0.85)
    D. 10,000-100,000 (0.95)
    E. 100,000-1,000,000 (0.96)
    F. 1,000,000-10,000,000 (0.98)
    G. 10 million + (1.00)

2) How is your target cohort distributed geographically?

Note: Go up means go to a higher score, so if you were at 0.75 and you choose A in this section, your final multiplier is 0.85.

    A. Rural/Remote (go up one)
    B. Small-Medium Town (go up one)
    C. Large Town - City (no change)
    D. State (no change)
    E. National (no change)
    F. 2-3 Countries (if 1D-1G go up two)
    G. Global (if 1D-1G go up two)

Section #1 - 33 points

1) Is your target cohort economically vulnerable?
`[A-0, B-1, C-3, D-6, E-8]`

2) Is your target cohort psychologically vulnerable?
`[A-0, B-1, C-2, D-3, E-5]`

3) Is your target cohort vulnerable to coercive control?
`[A-0, B-1, C-2, D-3, E-5]`

4) Is your target cohort over-policed or over-surveilled?
`[A-0, B-1, C-2, D-3, E-5]`

5) Is your target cohort technologically vulnerable?
`[A-0, B-1, C-2, D-3, E-5]`

6) Internet access?
`[A-0, B-1, C-2, D-3, E-5]`

Section #2 - 36 points

7) Does your decision system operate as part of a competitive marketplace, or are you the only provider of the specific service, platform, resource or goods?
`[A-5, B-4, C-3, D-1, E-0]`

8A) Is your decision system mandatory or optional?
`[A-6, B-0]`

8B) Is your decision system opt-in or opt-out?
`[A-0, B-0]`

8C) How visible or discoverable is it that the target cohort has been opted-into the system?
`[A-3, B-2, C-1, D-0, E-0]`

8D) How difficult is it to opt out?
`[A-3, B-2, C-1, D-0, E-0]`

8E) How perpetual is opting out?
`[A-0, B-1, C-2]`

9) Can you turn your decision system off *immediately* if something goes wrong?
`[A-0, B-3]`

10) Does your system always output a classification, prediction, or decision?
`[A-3, B-0]`

11) Do you provide a record of the system decision for individuals, including a unique ID that can be traced in your system log?
`[A-0, B-3]`

12A) Does your decision system collect meaningful consent, including a clear explanation of the uses of data and ramifications of providing consent?
`[A-0, B-5]`

12B) Have you performed usability or other user testing to assess the comprehensibility of your consent section?
`[A-0, B-2]`

12C) Do you provide a mechanism to allow individuals to revoke consent to use their data?
`[A-0, B-2]`

13A) Do you provide any visual or text-based explanation of your automated system decision on the decision interface?
`[A-0, B-2]`

13B) Can your target cohort request an explanation of the decision they received?
`[A-0, B-1]`

14A) Can your target cohort request a review of the decision they received?
`[A-0, B-6]`

14B) What is the probability that decision review will reverse or change the decision?
[A-2, B-1, C-0]

14C) On-average *how much effort* is required by your target cohort to lodge a request for decision review?
[A-0, B-1, C-2]

14D) On-average *how long* will your decision review process take?
[A-0, B-1, C-2]

Section #3 - 30 points

15) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Physiological needs*?
[A-0, B-1, C-3, D-6, E-8]

16) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Safety needs*?
[A-0, B-1, C-2, D-4, E-7]

17) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Love and belonging needs*?
[A-0, B-1, C-2, D-4, E-6]

18) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Esteem needs*?
[A-0, B-1, C-2, D-3, E-5]

19) What's the probability that decisions, classifications, predictions or interventions from your automated decision system will directly, negatively impact your target cohort's *Self-actualisation needs*?
[A-0, B-1, C-2, D-3, E-4]

Score guide

When you play the game of thrones you either win or die

Note: each score will be presented with a summary of scoring in each section, highlighting the poorest sections, offering mitigation suggestions where appropriate.

0-20
The white walkers
This territory belongs to the undead. Turn around! Don't venture here, where packs of spiders roam. You will harm people and almost certainly won't be able unpick those harms.

21-40
Fields of thorns
This path is paved with dangers and misery, and the long night casts it's shadow on this quest. There might be ways to win, but they'll be risky. Can you reinforce your troups, or find another path?

41-60
Winter is coming
What do you know of fear? You may not have felt the bite of winter's sting, but all the more reason to plan for it's hardships.

61-80
The stronghold
You're shoring up your castle walls - but there are still plenty of opportunities to stumble and impale yourself, or someone else.

81-99
Castles in the sky
You've found the highest ground of all! You're working on a problem space with a surprisingly low risk of harm! Don't become complacent, though.

Recommendations

TO DO: these should be tailored based on the answers you provide...

Try it with less people, first

- Trial your system with a small target cohort (Beta release). Communicate extensively with the cohort. Observe what goes wrong. Be prepared to stop or intervene if things are going badly.
- Staged roll-out of your system - break down your release rather than doing it all at once. Provide sufficient time between roll-out milestones to adjust and course-correct if unexpected issue crop up.

Graceful degradation

- Can you add your automation / model behind a feature switch so you can turn it off quickly if something goes wrong?
- Simple business logic fall-back built out in case the above is necessary?

Phone a (human) friend

- Loop in a person where the machine confidence is low
- Loop in a person when production data changes significantly

Increase transparency

- Add decision explanations (if you don't have them already)
- Externalise level of system confidence or uncertainty, especially against high-risk predictions
- Make sure your consent request is clear, interpretable and contextual
- If opt-out, ensure that the function and existence of system is visible and apparent to target cohort

GDPR Compliance

- Allow target cohort to revoke consent to use their data

Improve your feedback loops

- Let your target cohort flag questions or concerns related to the automated decision
- Plan for time to support and engage with them

Plan for failure

- Make a plan for how you will conduct decision review
- Think through your business logic and explanations so you don't have to work them out when dealing with someone who is angry or upset