



Ethical Data Impact Assessments and Oversight Models

January 2019

An Ethical Data Impact Assessment (EDIA) is a process that looks at the full range of rights and interests of all parties in a data processing activity to achieve an outcome when advanced data analytics may impact people in a significant manner and/or when data enabled decisions are being made without the intervention of people. An EDIA assists an organization in looking at the rights and interests impacted by the data collection, use and disclosure in data-driven activities. In order to determine whether an EDIA may be necessary, the organization should consider, before the activity begins and when there are any changes that affect the scope of the activity, whether the data processing activity involves **advanced analytics such as: evaluation or scoring (including profiling and predicting), automated individual decision-making, systemic observation or monitoring, data processed on a large scale, matching or combing data sets, innovative use or applying new technological or organizational solutions (such as AI and ML).** If the data processing activity may have an impact on an individual or on a group of individuals that may not be anticipated or easily known, then an EDIA should be considered either at the concept stage or at the service/product/analytical development stage or at both stages. If the data processing activity does not require an EDIA, then only a PIA may need to be completed.¹

The Model EDIA consists of four sections:

- I. **Purpose of the activity**
- II. **Data – a full understanding of the data, data use and parties involved**
- III. **Impact to parties and in particular individuals**
- IV. **Decision – whether an appropriate balance of benefits and mitigated risks supports the data processing activity.**

The very nature of an ethical and values-based assessment requires a careful consideration of the data activity benefits as well as the risks to individuals and society, considering the interests of all the parties who may be part of the activity. While open, structured questions can help, a way to organize the ultimate decision as to whether to proceed can be evaluated by using a well-established risk modeling process where the outcome of the analysis (significance, likelihood and effectiveness of controls) is depicted in a “net benefit/risk heat map”. This quantitative portion uses a standardized risk assessment process often found in many organizations Enterprise Risk Management (ERM) programs.

Successful implementation of an EDIA assumes and depends on the full implementation of the Enhanced Elements and, in particular, on highly qualified and competent, accountable roles and responsibilities with appropriate separation of duties. For example, EDIAs could be conducted by the privacy group. The structure of the overall Model EDIA and the questions in each section are illustrative, and the Model EDIA should be adapted as appropriate for each organization and/or industry as well as the different data-processing

¹. A PIA Template example can be found at the CNIL. See <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-2-en-templates.pdf>. The PCPD’s PIA information leaflet, https://www.pcpd.org.hk/english/resources_centre/publications/files/InfoLeaflet_PIA_ENG_web.pdf, also contains information on how to conduct a PIA. The Office of the Privacy Commissioner for Personal Data, Hong Kong, conducted a Privacy Compliance Assessment Report on the Smart Identity Card System (SMARTICS), <https://www.immd.gov.hk/pdf/PCARreport.pdf>

contexts. In particular, in the section that determines and describes how the data-processing could potentially impact the rights and freedoms of individuals, the impact should be assessed against a context-based set of issues. The Omidyar Network and Institute for the Future have established a comprehensive Ethical Framework for Tech–Technomy organized around “risk zones”² that could be used as a reference guide.

The EDIA is broader in scope than the typical PIA; however, the EDIA could be used in conjunction with the PIA. For example, all data are considered in an EDIA and not just personal data. However, to the extent the EDIA can be used to consider and appropriately mitigate the impact of a personal data practice, the EDIA process may supplement (or be woven into) the organization’s PIA process. In this regard, the EDIA process may enhance an organization’s privacy management program and compliance with its legal obligations under various regulatory frameworks.

An EDIA does not replace a PIA; it is designed to be used in conjunction with PIAs; it is not a complete PIA. Organizations may incorporate the EDIA in whole or in part into their own unique processes and programs so as to supplement or evolve with their PIA processes.

As a Model EDIA, other relevant authorities and/or regulatory bodies may provide input into its content and format. The goal of the EDIA is to encourage ICT innovation and competition by demonstrating that an organization has considered the interests of all parties before deciding to pursue an advanced data-processing activity.

². <https://ethicalos.org/>

Model Ethical Data Impact Assessment January 2019

EDIA Question	Answer/Notes
Section 1: Purpose of the Activity	
A. Business objective and purpose of the data activity	
<p>1. What is the business need/goal/objective for this data activity?</p> <p><i>If the purpose of the activity is to solve a question/problem, what particular question/problem is the activity trying to solve? Does the activity fit within a larger theme of work that is currently being contemplated or undertaken?</i></p>	
<p>2. Is this activity an expansion of a previous activity? If yes, determine whether a previous assessment has been done. If a previous assessment has been done, what has changed in this data activity and why (refer to previous assessment)?</p> <p><i>Does the activity fit within a larger theme of work that is currently being contemplated or undertaken?</i></p>	
B. Accountability for the data activity	
<p>1. Who has ultimate decision-making authority for the data activity?</p> <p><i>Who else needs to be involved in making the decision regarding the activity?</i></p>	
<p>2. Who is accountable for the various phases of the data activity?</p> <p><i>Who are the leaders that are responsible for the activity?</i></p>	
C. Legal and Other obligations regarding data collection, analysis and use(s)	
<p>1. What laws apply to the collection, analysis and use(s) of data?</p>	
<p>2. Does the data activity comply with all organizational policies and self-regulatory commitments?</p>	

3. Are there other legal, cross-border, policy, contractual, industry or other obligations linked to the collection, analysis and use(s) of data?	
4. How will all these obligations be managed and satisfied? <i>Have appropriate governance and accountability measures and processes been implemented?</i>	
Section 2: Data – A Full Understanding of the Data, Data Use and Parties Involved	
A. The nature of the data	
1. What specific types of data will be collected, tracked, transferred, used, stored or processed?	
2. Is the data identifiable to a person? <i>Determine whether there has been data linking of an identifiable individual's data or the data is reasonably linkable to an individual.</i>	
3. Is the data anonymous? <i>Determine how and what the anonymizing process was? Is it sufficient? Has the data been aggregated such that it is no longer identifiable personal data? Is reidentification possible? What policy, processes and/or technical measures have been used to minimize the reidentification of the data to an individual</i>	
4. Are there data elements that are the product of a probability-based process, such as a score?	
5. Is the data or anticipated use of the data sensitive? <i>Sensitive categories of data and/or use include Information associated with personal data that is used to decide or discriminate based on race, ethnic origin, religion or philosophical belief, sexual orientation, physical or mental health, information or data that could be used to facilitate identity theft, information associated with personal data that is used to permit access to an individual's account, precise location and/or there is a reasonable expectation the use of the data would be embarrassing to the individual whose data it is.</i>	

<p><i>Would any of the data use be considered sensitive to the individual?</i></p>	
<p>B. The sources of the data to be used in the activity</p>	
<p>1. What are all the sources and governance of the data, internal and external?</p> <p><i>Determine how the data was originated from each source and whether each source is a legitimate entity? How reliable is the source for the data activity? Is the source data permissible for the purposes of the activity? Who has custody or control over the source data and what are the governance arrangements?</i></p>	
<p>2. Determine if the data is provided by the individual (originated in direct action taken by the individual) and whether:</p> <ul style="list-style-type: none"> • The data is initiated (the product of individuals taking an action that begins a relationship) • The data is transactional (created when the individual is involved in a transaction) • The data is posted (created when individuals proactively express themselves) 	
<p>3. Determine if the data is observed (created as the result of individuals being observed and recorded), whether:</p> <ul style="list-style-type: none"> • The data is engaged (instances in which individuals are aware of observation at some point in time) • The data is not anticipated (instances in which individuals are aware there are sensors but have little awareness that sensors are creating data pertaining to the individuals) • The data is passive (instances in which it is very difficult for the individuals to be aware they are being observed and data pertaining to observation of them is being created) 	
<p>4. Determine if the data is derived (created in a mechanical fashion from other data and becomes a new data element related to the individual), whether:</p> <ul style="list-style-type: none"> • The data is computational (creation of a new data element through an arithmetic process executed on existing numeric elements) 	

<ul style="list-style-type: none"> The data is notational (creation of a new data element by classifying individuals as being part of a group based on common attributes shown by members of the group) 	
<p>5. Determine if the data is inferred (product of a probability-based analytic process), whether:</p> <ul style="list-style-type: none"> The data is statistical (the product of characterization based on a statistical process) The data is advanced analytical (the product of an advanced analytical process) 	
C. The accuracy of the data	
<p>1. Is the data accurate enough for the purpose of the activity?</p> <p><i>Determine what steps are being taken to determine the accuracy of source data and if the source data will be accurate enough over time? Has consolidation/transformation impacted the data in such a way the accuracy is affected? Are there concerns about the quality of the final data set relative to the purpose of the activity?</i></p>	
<p>1. What preprocessing will be done on the data before the analysis and will this affect the accuracy and appropriateness for the data activity?</p> <p><i>Determine what work will be done to put the source data used in the analysis in a consistent format? How will the data sources be consolidated for analysis? Will errors and redundancy in the data to be used in the analysis be identified and dealt with during preprocessing? If yes, describe how these errors and redundancies will be identified and addressed.</i></p>	
<p>1. Will preprocessing be done with data that is linkable to an individual? Describe how the preprocessing will be done and if there is any impact?</p> <p><i>Determine if there are any sensitivity issues or unique data protection issues with respect to the preparation of the data used in the analysis? What security is appropriate for preprocessing of the data? Will the preparation steps be accurate enough over time?</i></p>	
D. The governance of the data	

<p>1. Outside of individuals, who are all of the possible stakeholders and parties involved or related to the data activity? What are their interests and potential concerns?</p> <p><i>Stakeholders are very broad and apply to any party impacted by the data. A stakeholder for a framework could be a regulator or advocacy organization. Stakeholders for data and data uses include data partners. However, stakeholders can also include those interested in the success of a data use.</i></p>	
<p>2. If the data has been collected by, shared with and/or received from others, do those parties have authority to share?</p> <p><i>Determine whether the authority of those parties can be relied upon to protect impacted parties.</i></p>	
<p>3. Are there restrictions on data that would affect the use of the data?</p>	
<p>Section 3: Impact to Parties and in Particular to Individuals</p>	
<p>A. Identify all the impacted parties and the impacts on those parties</p>	
<p>1. During the activity, how will data be used and are there identifiable expectations of individuals, groups of individuals, and society for each use of the data?</p> <p><i>For example, could there be an impact (real or perceived) to social or reputation status?</i></p>	
<p>2. Could the data be used in a way that may result in a group of individuals being treated differently from other groups of individuals?</p> <p><i>Determine what the goal of the difference in treatment is.</i></p>	
<p>3. What are the benefits to the individual or groups of individuals?</p> <p><i>Determine and describe what the positive impacts on the parties are that are expected to come from the data activity. Consider factors such as: more objective or safer interactions, better product selection and utilization, better access to new products and services, significant</i></p>	

discounts, improved service or ease of use, more convenience or improved health and well-being. Improved financial condition, lower cost alternatives or increased availability.

4. What are the benefits to society?

Determine and describe what the benefits are that could be realized by someone beyond the immediate individual whose data is being processed. The processing of data will be more legitimate if the community or society can benefit from the usefulness of the data. Consider factors such as: better/lower cost health care, greater access to health services, or better health outcomes or an improved ability to track and assess health outcomes; more accurate sensors or devices to detect or diagnose health conditions or to improve general wellness; improved education; environmental enhancements such as water conservation, energy cost reduction; infrastructure enhancements; economic improvement; more accessible/usable technology; increased job opportunities; protection of reasonable expectation of privacy, including anonymity; protection of freedom of religion, thought and speech or protection of prohibition against discrimination.

5. How significant is the benefit? (1-Low; 3-Medium;5-High)

Description For Significance or Impact	Impact Score
The benefits or circumstance is Highly Impactful	High Impact 5
The benefits or circumstance is Moderately High Impact	Moderately High Impact 4
The benefits or circumstance is Moderately Impactful	Moderate Impact 3
The benefits or circumstance is Moderately Low Impact	Moderately low Impact 2
The benefits or circumstance is Minimally Impactful	Low Impact 1

6. Are the benefits likely to occur? How likely? (1-Low; 3-Medium;5-High)

Description For Likelihood	Probability	Likelihood Score
The benefits or circumstance is relatively certain to occur	90-100%	Expected 5
The benefits or circumstance is highly likely to occur	65-90%	Highly Likely 4
The benefits or circumstance is likely to occur	35-65%	Likely 3
The benefits or circumstance is possible but not likely	5-35%	Not Likely 2
The benefits or circumstance is only remotely probable	<5%	Slight 1

7. What are the benefits to the organization?

Consider factors such as increased revenue; lower costs; improved profitability; greater market share; enhanced employee satisfaction; engagement and productivity; enhanced customer relationships; enhanced or maintenance of brand or reputation; assurance of compliance; fraud

<p><i>prevention; enhanced or maintenance of cyber or physical security; new or improved products or services or customer service; improved manner of marketing; improved ability to assess customer preferences; improvements to innovation or enabling greater, faster, more efficient innovation; improved research processes; improved ability to conduct research and find or enroll study subjects or improved efficiency with studies; and innovative ways to conduct research. NOTE: The benefits to the organization are not factored in the numerical assessment of significance and likelihood.</i></p>	
<p>8. Considering all the factors relating to the data, the likely data use, the associated data activity, the identifiability and sensitivity of the data and the data activity objective, what are the risks to the individual, groups of individuals, and society? Determine and describe how the data processing could potentially impact the rights and freedoms to individuals.</p> <p><i>Consider the risks or increase in risks to the individual whose data is being used and those risks that occur because of the processing being considered. Areas to consider include: physical harm; financial harm; reduced health and well-being or reduced ability to move freely in society; damage to reputation or embarrassment; shock or surprise at the processing activity or the results of the processing; inappropriate discrimination, such as where the discrimination is based on a legally protected class such as race, age, religion or politics; the possibility of inappropriate access to or misuse of data by the company, including sensitive or special categories of data and directly identifiable data; manipulation of needs or desires/wants of the individual (i.e. creation of a need where one previously did not exist); a negative impact of data that are the product of a probability-based process, such as a score; data subjects who may be in a more vulnerable position than the organisation processing the data, such as children or elderly or less-educated or impoverished individuals; larger volume processing (versus a small-scale pilot). Who will have access to this technology and who won't? Will people or communities who don't have access to this technology suffer a setback compared to those who do? What does that setback look like? What new differences will there be between the "haves" and "have-nots" of this technology?</i></p>	

9. Is it foreseeable that the potential data analytical insights or the data activity might seem surprising, inappropriate or discriminatory or might be considered offensive causing distress or humiliation?

Would individuals be surprised by the data activity about them? Would the data activity about individuals align with the choices they have provided and the choices they have made? Determine whether there are other sensitivity issues with the potential insights and what aspect of collection/processing/analysis or use of potential insights might be considered unfair to the individual or society. Are all users treated equally? If not—and your algorithms and predictive technologies prioritize certain information or set prices or access differently for different users—how would you handle consumer demands or government regulations that require all users be treated equally, or at least transparently unequally? Are you creating data that could follow users throughout their lifetimes, affect their reputations, and impact their future opportunities? Will the data your tech is generating have long-term consequences for the freedoms and reputation of individuals?

10. Is the accuracy and/or quality of the data appropriate for the data activity?

Determine the impact of inaccurate data.

11. How significant is the risk? (1 – Low; 3 – Medium; 5 – High)

Description For Significance or Impact	Impact Score
The risk or circumstance is Highly Impactful	High Impact 5
The risk or circumstance is Moderately High Impact	Moderately High Impact 4
The risk or circumstance is Moderately Impactful	Moderate Impact 3
The risk or circumstance is Moderately Low Impact	Moderately low Impact 2
The risk or circumstance is Minimally Impactful	Low Impact 1

12. What factors about the activity have the highest impact on the likelihood any of these risks could be realized?

13. How likely is the risk to be realized? (1 – Low; 3 – Medium; 5 – High)

Description For Likelihood	Probability	Likelihood Score
The risk or circumstance is relatively certain to occur	90-100%	Expected 5
The risk or circumstance is highly likely to occur	65-90%	Highly Likely 4
The risk or circumstance is likely to occur	35-65%	Likely 3
The risk or circumstance is possible but not likely	5-35%	Not Likely 2
The risk or circumstance is only remotely probable	<5%	Slight 1

14. Are there technical and procedural safeguards (mitigating controls) that could be implemented to prevent and mitigate risks should they occur (e.g. encryption and delinking of data or increased transparency)?

A mitigating control is a type of control used to discover and prevent mistakes that may lead to uncorrected and/or unrecorded misstatements that would generally be related to control deficiencies. A mitigating control may help to remedy any elevated risk identified in the analyses above. Determine what risks can be mitigated and how these risks can be mitigated

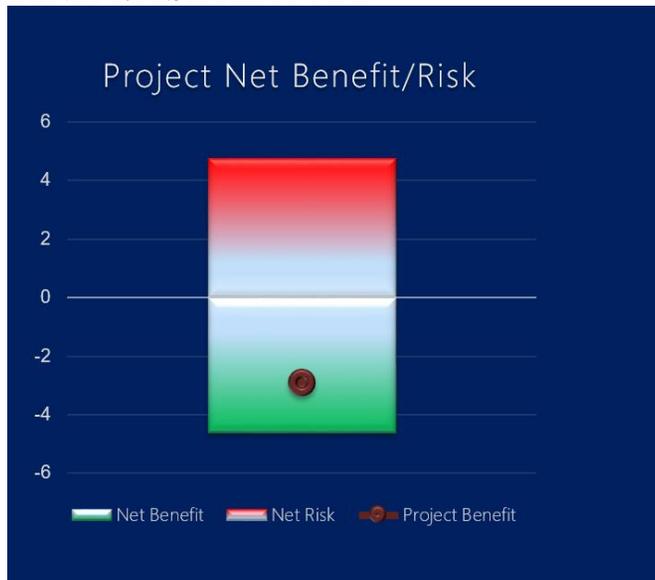
15. In the case of analytical driven models, insights or algorithmic decision-making, what is the useful life of each insight for each user? (Periodic recalibration of the insight might be necessary.) Are there appropriate testing and review mechanisms in place? How has the risk of bias in the data activity been addressed?

Determine how long the potential insight might endure and determine whether potential insights could become less useful or valuable over time. Are potential insights progressive and sustainable (repeatable over time) and for how long are potential insights sustainable? Application of potential insights could impact behavior in a manner that could reduce predictive value of insights over time. Does this technology make use of deep data sets and machine learning? If so, are there gaps or historical biases in the data that might bias the technology or data output? Is the technology reinforcing or amplifying existing bias? Is there a lack of diversity in the people responsible for the design of the technology? Are your algorithms transparent to the people impacted by them? Is there any recourse for people who feel they have been incorrectly or unfairly assessed?

16. Is there a less data-intensive way to achieve the goals of the data activity (including potential insights)?

<p><i>Determine whether the minimum possible amount of data has been used in the data activity or to obtain potential insights.</i></p>													
<p>17. If data is to be shared with any identified stakeholder have appropriate mechanisms to ensure adherence to data obligations been put in place?</p> <p><i>A PIA should be done even in the case of an EDIA, and core third- party sharing controls should be evaluated for effectiveness.</i></p>													
<p>18. Does the data activity include mechanisms that explain how data is used, how benefits and risks to individuals are associated with the processing, and how individuals may participate and object where appropriate?</p> <p><i>Determine what the transparency and individual accountability mechanisms are and whether they are appropriate for the data activity use. Does your technology do anything your users don't know about, or would probably be surprised to find out about? If so, why are you not sharing this information explicitly—and what kind of backlash might you face if users found out? If users object to the idea of their actions being monetized, or data being sold to specific types of groups or organizations, though still want to use the platform, what options do they have? Is it possible to create alternative models that build trust and allows users to opt-in or opt-out of different aspects of your business model moving forward?</i></p>													
<p>19. How effective are these controls and safeguards in reducing risk (1 – Low; 3 – Medium; 5 – High)</p> <table border="1" data-bbox="596 1062 1054 1252"> <thead> <tr> <th>Description For Significance or Impact</th> <th>Impact Score</th> </tr> </thead> <tbody> <tr> <td>The controls are Highly Effective</td> <td>Highly Effective 5</td> </tr> <tr> <td>The effectiveness of controls are Moderately High</td> <td>Moderately High Effectiveness 4</td> </tr> <tr> <td>The controls are Moderately Effective</td> <td>Moderately Effective 3</td> </tr> <tr> <td>The effectiveness of controls are Moderately Low</td> <td>Moderately Low Effectiveness 2</td> </tr> <tr> <td>The effectiveness of controls are Low</td> <td>Low Effectiveness 1</td> </tr> </tbody> </table>	Description For Significance or Impact	Impact Score	The controls are Highly Effective	Highly Effective 5	The effectiveness of controls are Moderately High	Moderately High Effectiveness 4	The controls are Moderately Effective	Moderately Effective 3	The effectiveness of controls are Moderately Low	Moderately Low Effectiveness 2	The effectiveness of controls are Low	Low Effectiveness 1	
Description For Significance or Impact	Impact Score												
The controls are Highly Effective	Highly Effective 5												
The effectiveness of controls are Moderately High	Moderately High Effectiveness 4												
The controls are Moderately Effective	Moderately Effective 3												
The effectiveness of controls are Moderately Low	Moderately Low Effectiveness 2												
The effectiveness of controls are Low	Low Effectiveness 1												

The OUTCOME of the assessment of benefits, risks and controls reflected in a Residual BENEFIT/RISK HEAT MAP ³



Section 4 – Decision: Whether an Appropriate Balance of Benefits and Mitigated Risks Supports the Data-Processing Activity.

A. Outcome

1. Are there any other factors that should be considered?⁴ Determine whether the interests, expectations and rights of individuals have been effectively addressed and what additional contextual based individual participation and choice factors should be considered.

³ Net or Residual Benefit/Risk model is for illustration purposes. Individual organizations can develop and modify consistent with their own Enterprise Risk Management system; the illustrative model consists of a numerical assessment of benefits (Significance and Likelihood) – Risks (Significance and Likelihood) = Inherent Risk – Effectiveness of controls = Residual Risk.

⁴ For additional technology specific solutions additional ethical factors should be considered – see [An Ethical Toolkit for Engineering/Design Practice](#) – Markkuhla Center for Applied Ethics and the [Omidyar Ethical OS Toolkit](#)

<i>Consider if the risks are necessary and proportional to the benefits? Have the risks have been mitigated to the extent possible? Are the mitigated risks sufficiently balanced by the benefits?</i>	
2. Does the purpose of the activity fit within the values of the organization?	
3. Does the purpose of the activity fit within the values of society?	
4. Does the data activity comply with all laws, cross-border, policy, contractual, industry or other obligations organizational policies and self-regulatory commitments?	
5. Have appropriate governance and accountability measures and processes been implemented? How will and how well will these obligations be managed and satisfied?	
6. Have all the stakeholder concerns identified in the Governance of Data section been appropriately addressed?	
7. After considering all the above factors, is the activity a “go,” “no go,” or should some aspect of the activity be recalibrated to reduce the residual risk?	
B. Approvals	
1. Have all the individuals described in I.B.1. through I.B.2. above been involved in the decision?	

IV. The Process Oversight Model

Assessments conducted solely by the parts of a business implementing intensive data activities may raise issues of trustworthiness. Where the oversight of the assessment and accountability process is done by the organization itself (versus the accountability or regulatory agency), then the oversight should be conducted pursuant to a common framework.⁵ Until such an approach is established, the Process Oversight Model looks at how an organization has translated organizational ethical values into principles and policies and into an “ethics by design” program. It considers how well established internal review processes, such as EDIAs and effective individual accountability systems, have been implemented. It presumes the oversight process is independent from the assessment process. It could be a function performed by, for

⁵. See IAF, Report for Comprehensive Assessment Oversight Dialog: Canadian Ethical Data Review Boards Project, March 31, 2018, [18-24](http://informationaccountability.org/wp-content/uploads/Report-for-the-Comprehensive-Assessment-Oversight-Dialog-Canadian-Ethical-Data-Review-Boards-Project.pdf) [IAF Oversight Report].

example, an internal audit group. It may be likened to an assessment of “controls and controls effectiveness” by the internal audit group.

The internal audit group usually is established by the Audit Committee of the Board of Directors or the highest level of the governing body. The Chief Audit Executive reports functionally to the Board, and the internal audit function is independent and objective. The scope of internal audit’s responsibilities encompasses, but is not limited to, the examination and evaluation of the adequacy and effectiveness of the organization’s governance, risk management, and internal controls.⁶ The Process Oversight Model can be thought of as analogous to a set of control definitions against which the capability and effectiveness of the organization’s assessment process is tested. A set of control parameters across functional assessment domains is established and then, through a set of audits, the effectiveness of the relative controls is tested. While this oversight could be performed by internal audit, it could also be accomplished by way of an assessment or test conducted by an external resource (e.g. a consulting firm). This sort of audit and testing work is similar to work already performed by these external firms in other domain areas.

The Oversight Model consists of questions in seven sections:

- I. Accountability for the oversight process**
- II. Translation of organization values into principles and policies**
- III. Translation of organizational values into an “ethics by design” program**
- IV. Utilization of the EDIA**
- V. Internal review process**
- VI. Individual accountability system**
- VII. Transparency of process.**

The questions in each section of the Process Oversight Model are illustrative, and the Process Oversight Model should be adapted as appropriate for each organization to oversee the trustworthiness of its assessment process.⁷ The Process Oversight Mode is designed to address the ethics part of data stewardship and assumes other internal oversight processes exist to address core elements of privacy programs.

Evidence of oversight is important. Oversight provides rigor to the assessment process and demonstrates that oversight of the EDIA process has occurred. Whether this oversight occurs internally, for example by the audit group, or

⁶. “Model Internal Audit Activity Charter,” The Institute of Internal Auditors (rev. 05/2013)

<https://global.theiia.org/standards-guidance/public%20documents/modelcharter.pdf> .

⁷. An assessment of the process is designed to be different than a secondary assessment of a specific data-intensive activity.

externally, for example by a consulting firm, it is necessary that documentation exists that demonstrates how the oversight was conducted and that, in fact, it was conducted. The oversight process should measure whether the EDIA process is being conducted with honesty and recognizes the full range of interests of all parties in order to demonstrate that the interests of the organization were not placed in front of the interests of other parties.⁸ The organization should stand ready to demonstrate its assessment governance process and individual assessments to regulators with appropriate authority.⁹ The Process Oversight Model provides guidance regarding how such oversight should be conducted and documentation that the oversight actually occurred.

The questions in each section of the Process Oversight Model are illustrative and should be adapted as appropriate for each organization to oversee the trustworthiness of its assessment process¹⁰. The Process Oversight Mode is designed to address the ethics part of data stewardship and assumes other internal oversight processes exist to address core elements of privacy programs. As process oversight models evolve, there may be input and guidance from other relevant authorities and/or regulatory bodies. Such input and guidance will increase the trustworthiness of the EDIA process.

Process Oversight Model

Oversight Question	Answer/Notes
I. Accountability for Oversight Process	
1. Are accountability and responsibility for achieving outcomes established through clearly defined roles throughout the organization? <i>Are the accountable and responsible roles carried out by competent and capable individuals? Is there a clear separation of duties between data activity roles?</i>	
II. Translation of Organizational Values into Principles and Policies	
1. Are shared organizational values described and/or articulated and have they been integrated into the organization?	

⁸. [IAF Oversight Report](#) p. 21

⁹. Id. pp. 23-24.

¹⁰. An assessment of the process is designed to be different than a secondary assessment of a specific data intensive activity.

<p><i>Have the values have been condensed to core and guiding principles that are understood by technical staff? Have they been fully translated into organizational policies and processes? Have they been programmed into data and activity objectives?</i></p>	
<p>2. Have the articulated values been aligned to the varied geographic-values across the organization’s reach and footprint?</p> <p><i>Could design choices become international standards or norms?</i></p>	
<p>III. Translation of Organizational Values into an “Ethics by Design” Program</p>	
<p>1. Does the organization have an “ethics by design process” that is part of its products/service development process?</p> <p><i>Determine whether Core or Guiding Principles are understood by staff (in particular by technical staff) and have been programmed into activity objectives and the full product/service development lifecycle.</i></p>	
<p>2. Does the product/service development process ascertain whether there is benefit to individuals and society in addition to the organization?</p>	
<p>IV. Use of an EDIA</p>	
<p>1. Does the organization use an EDIA to achieve a principles-based outcome of data? Is the assessment process effective?</p> <p><i>Does the organization assess all risks and benefits to an individual, group of individuals, and society? Are the risks effectively mitigated? Does the EDIA process effectively evaluate that data use avoids actions that seem inappropriate or discriminatory, might be seen as generating unequal treatment, might be considered offensive or causing distress or humiliation?</i></p>	
<p>2. Does the EDIA process effectively assess the complexity and potential impact of the data and data use?</p> <p><i>Does the EDIA process consider all the factors relating to the data, the likely data use, the associated data activity, the identifiability and sensitivity of the data, as well as the potential impact of the data activity?</i></p>	
<p>3. Does the EDIA process effectively evaluate if the purpose of the activity fits within the values of the organization and society?</p>	

<p>4. Is there an effective triage process to determine what type of assessment is appropriate? Is this process effectively employed?</p> <p><i>A triage process determines the level of review of the process necessary. Where data processing is very similar to processing that has been done in the past and therefore it was concluded no assessment was necessary, only a quick review may be required to confirm those understandings. Where data uses are more complex and/or less obvious to the parties and more rigorous assessments were conducted, a more rigorous review should be required. Where the uses are most complex, an EDIAs that effectively weighs the risks and benefits should be used.</i></p>	
<p>V. The Internal Review Process</p>	
<p>1. What kind of periodic assurance reviews will occur over time?</p> <p><i>Do the periodic reviews appropriately consider the data and data-use objectives? Is the periodic review process established at appropriate timeframes?</i></p>	
<p>2. For intensive data impacting systems, does the review assess that outcomes are as intended with the objectives of the activity and impacts are mitigated as planned, harms are reduced, and unintended consequences are understood?</p> <p><i>Determine whether this analysis includes the likelihood of benefits being achieved and risks effectively mitigated. Does the post review include an assessment of if the anticipated outcomes were achieved?</i></p>	
<p>3. Have analytic models and insights been tested for their accuracy and predictability?</p> <p><i>Is there an ongoing systematic process to ascertain whether analytic models are tested for their consistency with organizational values and principles? Are data-intensive technologies subject to appropriate human direction and control?</i></p>	
<p>4. Does the review process include a risk review by senior-accountable leadership? Are higher risk activities approved by senior-accountable leadership?</p> <p><i>Is there a formalized, risk-ranked review process where higher impacting data activities are reviewed? Where internal reviewers need external expertise? Is this expertise sought?</i></p>	

5. Does the assessment and review process ascertain whether all parties' concerns are assessed and appropriately addressed as part of the data-system lifecycle?	
6. Have systems themselves, and the data that feed those systems, been assessed and protected proportionate to the risks?	
7. Does the review evaluate whether only the minimum data that is needed is used?	
VI. Individual Accountability System	
1. Are there effective systems that provide appropriate opportunities for feedback, relevant explanations, and appeal options for the individuals impacted? <i>Will individuals have some ability to engage in how their data is used? How will individual situations be remediated, if necessary?</i>	
VII. The Transparency of the Process	
1. Does the organization have mechanisms that explain how data is used, how benefits and risks to individuals are associated with the processing, and how individuals may participate and object where appropriate? <i>Is the use of the data transparent and effectively made available for all data activities?</i>	
2. Is the organization ready to demonstrate the soundness of the processes they use so that data and data-use systems are consistent with established values and principles? <i>Can the organization demonstrate its data stewardship accountability processes?</i>	