

# Proposed methodological changes to the 2027 DFI Transparency Index

## Contents

Introduction.....	1
Completeness of project disclosure.....	2
Segmenting analysis of E&S documents according to risk category.....	5
Changing IATI scoring protocols.....	8
Reducing the first-round project sample size.....	10
Proposed use of AI-assisted search and screening.....	11

## Introduction

This paper sets out our proposed changes to Publish What You Fund’s DFI Transparency Index for the 2027 edition. The proposed changes are a result on an in-depth internal review of the Index. We plan on using this paper to inform discussions and seek feedback from stakeholders.

As we conduct our methodology review, we are guided by two principles: to improve the rigour of the DFI Transparency Index, and to maintain a degree of comparability between the 2023, 2025 and 2027 editions.

The changes we propose for 2027 are:

- Clarifying the completeness penalty system
- Risk segmentation of analysis for environmental and social (E&S) disclosures
- Changing International Aid Transparency Initiative (IATI) scoring protocols
- Reducing the first-round project sample size to 10
- Introducing AI-assisted evidence search and screening

We welcome written feedback on the proposed changes. **Please send any written feedback on this consultation paper to [ella.remande-guyard@publishwhatyoufund.org](mailto:ella.remande-guyard@publishwhatyoufund.org) by the 20th April 2026.**

## Completeness of project disclosure

### ***Rationale***

Comprehensive disclosure of investment activity is essential for ensuring that stakeholders, including affected communities, civil society organisations, researchers, and partner governments, can understand the full scope of development finance operations. Where project databases reflect only a partial record of a DFI's activities, it becomes difficult for users to analyse portfolios, track development outcomes, or hold institutions accountable for their investments.

### ***Current approach***

In previous iterations of the DFI Transparency Index, a single test has been used to incentivise both timely and comprehensive disclosure of projects.

This test compares the number of projects disclosed in the most recent period of the project sample with historical disclosure levels. DFIs are expected to have disclosed a near-complete set of projects for the relevant portion of the sample period. Where this threshold is not met, a penalty is applied to project-level indicator scores. For the 2025 Index, to determine completeness of the projects for the first 6-month period of 2024, we checked whether there were at least 80% as many projects as the semi-annual average for the prior three years (2021, 2022, and 2023).

While this approach has been effective in encouraging institutions to publish projects from the most recent period, it has primarily functioned as a proxy for timeliness of publication, rather than a direct assessment of disclosure completeness. The test assumes that projects have not yet been disclosed due to delays in publication, rather than because they are not disclosed at all.

### ***Limitations of the current approach***

The current approach assesses completeness by comparing disclosure levels with those of previous years, implicitly assuming that historical disclosure was broadly complete. However, experience from recent Index cycles has shown that this is not always the case. Where institutions have historically under-disclosed projects, comparisons with previous years may fail to identify systematic gaps in disclosure. This means that an institution that consistently publishes only a portion of its investments may still appear to meet the test, provided that its disclosure levels remain similar to those of previous years, even when its public database does not reflect the full set of approved investments.

To address this limitation, the Index will undertake a more explicit test of completeness of disclosure, which compares what DFIs report that they have done with what they have actually disclosed publicly.

### ***Proposed approach***

The Index will introduce a dedicated completeness test that assesses whether the projects disclosed on a DFI's public database represent a substantially complete record of the institution's investment activity.

This test compares the investment activity reported in a DFI's annual report with the projects disclosed on the DFI's public data portal or project database. The aim is to determine whether the publicly available dataset reflects the scale of activity reported by the institution.

The primary metric used for this comparison will be total investment volume approved or committed, as this is reported more consistently across institutions and avoids definitional differences in how projects are counted. Where annual reports do not clearly state total investment volume, the number of new investments or projects can be used as an alternative metric.

### ***Assessment year***

For the 2027 DFI Transparency Index, the project sample used for indicator assessment will consist of projects approved in 2026.

However, the completeness test will use 2025 data. This is because annual reports for the sample year are not consistently available across institutions at the time the Index analysis is conducted. Using the previous year ensures that annual reporting figures are finalised and comparable across all DFIs, allowing for a fair and robust comparison between reported investment volumes and publicly disclosed data.

Under this approach, the total investment activity reported in the 2025 annual report will be compared with the projects disclosed for that same year in the DFI's public database. Using the most recent year for which reporting is finalised allows the Index to distinguish between genuine under-disclosure and normal publication delays.

### ***Completeness threshold***

DFIs are expected to disclose at least 50% of the investment activity reported in their annual report.

The 50% threshold is intended to identify cases of systematic under-disclosure, while recognising that differences may arise between annual report figures and project-level datasets due to definitional variations, timing differences, or reporting practices across institutions. Setting the threshold at this level allows the Index to distinguish between minor discrepancies and situations where a substantial portion of a DFI's activity is not publicly disclosed.

If the investment volume (or number of projects) disclosed on the DFI's public database represents less than 50% of the activity reported in the annual report, Publish What You Fund will contact the institution to clarify the discrepancy. This allows DFIs to provide explanations

for legitimate differences that may arise due to definitional issues, reporting conventions, or timing of publication.<sup>1</sup>

If no satisfactory explanation is provided, the DFI will receive a 50% penalty applied to all project-level indicator scores.

A 100% penalty will be applied where no projects from the relevant year are disclosed at all.

### ***Improvements in disclosure***

The completeness test is intended to identify cases of systematic under-disclosure, rather than penalising institutions that are actively improving their transparency practices.

Where a DFI demonstrates that it has substantially improved disclosure of recent projects, for example by publishing over 50% of its investment activity for the Index sample year, even if historical disclosures remain incomplete, Publish What You Fund may take this into account when interpreting the results of the completeness test.

In such cases, the assessment may draw on newer annual reports where available, or other officially reported statistics or figures published by the institution (such as audited financial statements or official portfolio statistics), to determine whether recent disclosure levels more accurately reflect the institution's current level of activity.

This ensures that the methodology continues to identify cases of systemic under-disclosure while recognising genuine improvements in transparency practices.

---

<sup>1</sup> DFIs may disclose projects at different stages of the project cycle (such as pipeline, approval, signing, or disbursement). For the purposes of this assessment, the Index expects that projects approved and committed during the relevant year can be disclosed at least at a basic level, regardless of their implementation status. Differences in internal disclosure practices do not therefore affect the expectation that a substantially complete record of approved projects should be publicly available.

## Segmenting analysis of E&S documents according to risk category

### ***Current approach***

We assess the transparency of environmental and social (E&S) document disclosure and the assurance of community disclosure using our general sample. This means that we assess a randomly chosen selection of 15 projects with no intentional sampling of low, medium, or high E&S risk projects. To pass the relevant indicators, an institution has to have sufficient disclosure in at least 80% of the sample (12 projects or more).

In the case of E&S documents, we assess transparency of projects according to their E&S risk profile. For direct investments, low risk projects are expected to include an explanation of why they are low risk, while high risk projects are expected to include an environmental and social impact assessment (ESIA) or equivalent, and a stakeholder engagement plan (SEP), at minimum. Medium risk projects are expected to state whether ESIA's were produced and disclose them if so. For financial intermediary (FI) investments, FI1 investments are expected to disclose an Environmental and Social Management System (ESMS); FI2 investments are expected to have a summary of an ESMS unless it is explicitly stated that an ESMS was not required; and FI3 investments need a justification for risk categorisation including description of E&S risks.<sup>2</sup>

We look for assurance of community disclosure in a similar way - expecting to identify it for high risk projects and for medium risk projects unless it is explicitly stated that community disclosure was not necessary.

### ***Problem with current approach***

As the current approach uses the general sample and owing to the fact that MDBs and DFIs disproportionately classify projects as medium risk, we normally assess mainly medium risk projects, with a smaller number of low and high risk projects. This creates an issue in instances where an MDB or DFI discloses required information for high risk projects but does not do so (or does so in a limited manner that does not satisfy our assessment) for medium risk projects. In these instances, the MDB or DFI can be expected to perform similarly in our assessment to an institution that does not disclose required information for either high risk or medium risk projects.

For example, DFC typically discloses ESIA's and SEPs for high risk projects, but does not do so for medium risk projects. No other bilateral DFI discloses these E&S documents for either high risk or medium risk projects. However, as DFC invests in relatively few high risk projects, it does not reach the threshold to pass the indicators involved. As such, despite a clear

---

<sup>2</sup> For more detail on the current approach, see the DFI Transparency Tool page here: <https://www.publishwhatyoufund.org/dfi-index/dfi-transparency-tool/es-project-plans-assessments/>

difference in their disclosure practices, DFC perform similarly to other bilateral DFIs for the indicators in question.

### ***Proposed new approach***

We propose segmenting our analysis of E&S disclosures according to the risk profile of investments to add a new degree of nuance to our assessment. In practice, this will involve creating specific risk-aligned samples for E&S analysis for certain indicators (identified below). We propose creating two distinct samples - a high risk sample of up to 15 projects, and a low and medium risk sample of up to 15 projects. The high risk sample will include Category A direct investments and Category FI1 financial intermediary investments. The remaining categories will be included in the second sample. We will then analyse the indicators twice - once for each sample. For the purposes of scoring, the indicator weights will be distributed evenly between the high risk sample and the low and medium risk sample. While we recognise that low and medium risk projects are more common, we believe that the significant risks of high risk projects merit equal weighting in the indicator scores.

The new approach requires several caveats. First, in instances where DFIs do not disclose the risk category of projects, we will continue to use the general sample for analysis and to assess all projects as if they were high risk. This approach is consistent with our current methodology and ensures that institutions do not gain an unfair advantage by not disclosing E&S risk categorisations. Second, where institutions invest in less than 15 projects in either risk category, we will construct smaller samples that capture the totality of relevant investments within the assessment period. This will ensure that our assessment is representative of the institution's approach during the assessed period. Finally, in instances where an institution does disclose risk categorisations, but has not made any investments in a particular risk category during the assessed period, we will only assess the other category and will assign all indicator points to that category. For instance, if an institution discloses risk categories, but has not made any high risk investments during the assessed period, we will only assess low and medium risk projects, and this sample will account for 100% of the available points for the indicator. While we recognise that in such cases there may be earlier examples of high risk investments, our assessment is designed to capture disclosure practices in a time bound manner, and to measure recent transparency levels. As such, it would not be appropriate or comparable to use older investments.

### ***Affected indicators***

The indicators and survey questions that will be affected by the segmented approach are:

#### **E&S Project Plans/Assessments**

- Does the DFI disclose the minimum E&S documentation for the risk categorisation?
- Does the DFI disclose what E&S documentation was produced for the activity?

- Does the DFI disclose all identified E&S documents that were produced for the activity?

### **Assurance of Community Disclosure**

- Does the DFI state if disclosure of the activity to projected-affected people was required?
- If yes, then: Does the DFI state the date of community E&S disclosure?
- Does the DFI state the place of community E&S disclosure?
- Does the DFI state the method of community disclosure?
- Does the DFI state what documentation was disclosed?
- Does the DFI identify the language of disclosure to project-affected people?
- Does the DFI state if the presence of a project grievance mechanism was disclosed and the method it was disclosed?
- Does the DFI state if the presence of an independent accountability mechanism was disclosed and the method it was disclosed?

## Changing IATI scoring protocols

We propose changing the scoring protocol relating to IATI publication. The current scoring protocol for indicators that have applicable data fields in the IATI Standard incentivises consistent publication of data to IATI by allowing DFIs to score more for publication in this format. In the current 2025 methodology, to be awarded points for IATI publication, information in IATI must be consistent with that published in other formats including bulk download, via the webpage, or in documents such as pdf files.

Currently, scoring works as follows:

<b>Bulk Download Protocol (IATI Publication)</b>	<b>Open Format Protocol (IATI Publication)</b>
IATI consistent with any other format: 100 points	IATI consistent with other format: 100 points
Bulk download format without consistent IATI publication: 75 points	All formats without consistent IATI publication: 75 points
Data on webpage without consistent IATI publication: 50 points	
Data in non-machine-readable file (pdf or similar) without consistent IATI publication: 25 points	

We suggest removing the requirement for information published to the IATI Standard to be supplemented by, and consistent with, disclosure in other formats. Under the revised approach, the assessment of IATI publication would be fully independent from the assessment of disclosure via bulk download, webpages, or document formats. DFIs would therefore be able to score 25 points for publishing relevant information to the IATI Standard, even if the same information is not available elsewhere.

The rationale for this change is to avoid unfairly penalising institutions that publish relatively good-quality data or specific data points solely to IATI. This change allows us to distinguish between DFIs that publish no information and those that publish to IATI even if other formats remain limited. Recognising IATI publication in its own right ensures that institutions taking initial steps towards transparency receive some credit, while still encouraging improvements across other formats over time.

The new protocol we suggest is as follows:

**Bulk Download Protocol (IATI Publication)**

IATI consistent with any other format: 100 points

IATI with no other format disclosed: 25 points

Bulk download format without consistent IATI publication: 75 points

Data on webpage without consistent IATI publication: 50 points

Data in non-machine-readable file (pdf or similar) without consistent IATI publication: 25 points

**Open Format Protocol (IATI Publication)**

All formats with consistent IATI publication: 100 points

IATI with no other format disclosed: 25 points

All formats without consistent IATI publication: 75 points

## Reducing the first-round project sample size

Under the current methodology, we assess DFIs through two rounds of sampling and analysis. In the first round, a sample of 15 projects is selected at random from investments made over the defined period and this round is explicitly not used to contribute to final index scores; instead, it serves to provide DFIs with initial feedback and an opportunity for engagement and improvement ahead of the main scoring round. The second round, also comprising 15 randomly selected projects from the later timeframe, is the only round that contributes to the final scores in the Index.

We propose reducing the first-round sample size from 15 to 10 projects. This change is justified on the following grounds:

- The first round's primary purpose is to provide an early, indicative assessment that can be discussed with DFIs to highlight initial areas for improvement and foster joint learning. Because this round does not affect final scores and is designed as an engagement exercise, a smaller sample can still achieve these aims effectively.
- Our experience from previous assessments shows a high degree of consistency in disclosure practices across projects within the same institution. As such, reducing the sample size is unlikely to meaningfully alter the insights generated in this early phase but will reduce unnecessary workload.
- Reducing the first-round sample reduces burden on both the assessment team and DFIs during initial engagement, allowing us to allocate time and resources more strategically without compromising the rigour of the second round, on which the sample size will remain unchanged.

## Proposed use of AI-assisted search and screening

The Index assessment process relies on systematic review of large volumes of publicly available information, including project documents, environmental and social (E&S) assessments, policy papers, webpages, and bulk data files. In previous iterations of the Index, manual review proved relatively time and resource intensive.

To strengthen the rigour and reduce the workload of the assessment process, we propose introducing a limited and clearly defined use of artificial intelligence (AI) tools to support document identification and analysis. Under this approach, AI would be used as an evidence-retrieval and screening tool, and not for scoring or evaluative judgement. This change is intended to enhance the efficiency of evidence collection, while maintaining the central role of human review.

Under the proposed approach, AI tools would be applied only in two clearly defined circumstances: (1) to assist initial identification of sources; and (2) to assist the screening of identified sources.

The **initial identification of sources** would be relevant for the following indicators, all of which require checking whether certain types of documents, links, or files are publicly disclosed: 1, 2, 3, 19, 20, 21, 25, 26, 27, 28, 29, 31, 35, 36, and 45.

In practice, this would involve:

- Conducting manual searches of DFI websites and public search engines to identify relevant bulk downloads, webpages, and PDF documents, in line with the current approach
- In cases where relevant sources cannot be readily located through manual searches, using AI-assisted searches to support identification
- Reviewing all identified sources to confirm relevance and compliance with indicator criteria

**Screening identified sources** would be relevant for the following project-level indicators, which rely on open format disclosure: 14, 15, 22, 23, 24, 30, 32, 33, 37, 39, 40, 42, 43, 44, 46, 47, and 48. For these indicators, relevant information is often buried in lengthy, non-machine-readable documents and webpages, meaning that identifying it requires time-consuming manual review and carries a risk of information being missed.

In practice, this would involve:

- Uploading relevant sources (selected text extracted from bulk download files, PDFs, webpage URLs)
- Using standardised prompts to search for predefined disclosure elements and keywords

- Extracting short verbatim quotations and their locations (page number, section heading, or URL)
- Reviewing quotations to confirm compliance with indicator criteria
- Conducting manual review of all sources where AI-assisted screening does not identify relevant information, to confirm whether disclosures are absent

AI tools would *not* be used to:

- Score indicators, assign points, or make evaluative judgements
- Interpret compliance with indicator definitions or determine the absence of disclosure
- Replace manual review, generate narrative assessments, or substitute for quality assurance
- Override human judgement

All AI-identified evidence would be subject to mandatory team review prior to scoring and only verified evidence will be used. To support transparency, the team will maintain documentation on indicator-specific prompt templates, and all AI-assisted searches will be conducted through a centrally managed AI tool account to ensure coordination and consistency across DFI assessments.