World Intellectual Property Organization | **Organisation Mondiale de la Propriété Intellectuelle**
Internal Memorandum | Mémorandum Interne

INTERNAL AUDIT AND OVERSIGHT DIVISION

EVALUATION REPORT

Validation of the 2008 PPR

May 25, 2009

TABLE OF CONTENTS

LIST OF ACCRONYMS USED

RBM	Results-Based Management
PPR	Program Performance Report
WIPO	World Intellectual Property Organization
IAOD	Internal Audit and Oversight Division
OECD/DAC	Organization for Economic Co-operation and Development/ Development Assistance Committee
MTS	Medium-Term Strategy
PMPS	Program Management and Performance Section
PID	Performance Indicator Data
PMSDS	Performance Management and Staff Development System

I. EXECUTIVE SUMMARY

A. Introduction

1. In the context of strengthening a Results-Based Management (RBM) approach, a ‘trial’ Validation Exercise took place during March and April 2009. Its aims were to provide independent verification of the reliability and authenticity of information contained in the 2008 Program Performance Report (PPR), and to help raise the quality of reported Performance Indicator Data (PID). The assessment of quality of data was carried out in the context of the main purpose for which this information is designed to be utilized – the annual reporting to Member States.

2. 15 out of the 31 programs of WIPO were sampled, with one result and supporting indicators from each of the programs selected for validation. As this was the first such exercise, the main emphasis was placed on lesson-learning rather than accountability. This report contains an assessment of the quality of data reported against nine criteria: 1. accurate / verifiable; 2. comparable/consistent; 3. timely; 4. efficient / easily accessible; 5. analysis carried out; 6. routinely generated; 7. sufficient / comprehensive; 8. relevant; 9. clear / transparent. It reveals examples of existing good practice in the collection and reporting of data, and identifies areas for improvement, with implications for individual programs as well as the organization as a whole.

3. This exercise was undertaken at a time when WIPO has been experiencing substantial internal changes. These include the introduction of a range of initiatives aimed at improving the effectiveness and efficiency of its services, with an emphasis on longer-term strategic planning and a continuing focus on impact and results. Given the pressures on time, the managers interviewed demonstrated an impressive openness to suggestions and commitment to improving the quality and efficiency of reporting PID. This report is intended to make a practical contribution to continuing efforts to improve the quality of planning, monitoring and evaluation within a results-based framework.

B. Main findings

4. The major strengths of the current practice of reporting PID relate to the criteria of accuracy, comparability, timeliness and accessibility of data.

5. The main limitations of the data relate to the clarity, relevance and sufficiency of the data presented in the PID section of the PPR. There was found to be a general tendency to provide more information than was required and, in several programs, there was too much emphasis placed on the reporting of activities, events and outputs rather than outcomes and results that lead to change and impact.

6. With some notable exceptions, PID are little used for routine or regular monitoring to measure progress and communicate success, either at a program level or by supervising managers.

C. Significant conclusions

7. The data reported in the sample were generally accurate and reliable, although often lacking important elements of clarity and transparency. This may be partly explained by the low level of pressure for implementing managers to routinely report on progress or demonstrate success against the results and indicators, except on an annual basis. Exceptions are evident in a few high profile programs which have set up their own efficient and tight monitoring systems. The utilization of PID for internal purposes is likely to be increased with the forthcoming introduction of the Performance Management and Staff Development System.

8. Data quality is most likely to show improvement as more programs recognize the value of routine monitoring of PID to inform management decision-making in areas such as policy development, performance management and operational efficiency. The increasing integration of sound monitoring practice will, in turn, necessitate a more careful prediction of expected results and targets, together with a selection of indicators which are capable of accurately reflecting what is most valued in relation to progress and success.

9. The commitment of implementing and supervising managers to improving data quality will also be strengthened as program objectives and results are linked more directly with the overarching strategic goals being developed by the senior management team. Additional support is being offered by the Evaluation and the Program Management and Performance Section with the preparation of the Program and Budget document for 2010-2011 and the development of a Medium Term Strategy which encourages planning and reporting within a context of longer-term impact and change.

D. Major recommendations

10. At a time when program managers and teams are reportedly experiencing high levels of organizational change and pressures on time, not least in the requirements of planning for the next Biennium and the Medium Term Strategy, recommendations are likely to be more readily applied if they are aligned as far as practical with existing initiatives. In addition to the specific feedback provided to individual managers during the conduct of this validation exercise, the major recommendations for improving the quality of PID at an organizational lever are:

(a) A review should be carried out to determine the extent to which PID can be more regularly utilized for routine monitoring of progress in programs. Depending on the extent to which this is considered to be a priority for senior management, stronger monitoring systems should be expected and encouraged for the practical integration of the results-based approach into day-to-day management, to complement the existing emphasis of RBM on financial planning and reporting to Member States.

(b) If and when PID becomes increasingly used for internal monitoring purposes, supervising managers should have a more visible role in supporting the development and maintenance of robust monitoring systems. They will also be influential in establishing strong and clear links between program level objectives and overarching organizational strategic goals and objectives.

(c) Specific assistance to supervising and implementing managers and teams should include:

i) Increased technical support for the development of appropriate, computerized data collection, analysis and reporting tools;

ii) A more coordinated collation and analysis of user feedback across the programs, building, possibly, on the proposed Customer Service initiative;

iii) Continued one-to-one training and advice in the understanding and application of good practice standards in performance planning and monitoring systems;

iv) The development of a monitoring tool that is capable of identifying overall progress against key objectives and indicators on a routine basis, e.g. quarterly, for the senior management team.

II. BACKGROUND

A. Program Performance Reports

11. The Results-Based Management (RBM) initiative was introduced into WIPO in 1998. As part of the planning and reporting cycle, each program contributes to a Proposed Program and Budget for a Biennium and an annual Program Performance Report (PPR).

12. A review of the 2006-2007 PPR process was carried out in October 2008. One of its main findings was that, whilst successful efforts had been made within the organization to enhance efficiency, transparency and accountability, the RBM systems were driven too much by the budget systems and too little by the organization's own strategy. A resulting challenge presented to WIPO was 'to move towards a more integrated results and performance-based approach within current arrangements'.[1]

13. This validation exercise focuses on the information being presented in the PPR for 2008, based on the original 2008/09 approved Program and Budget. A revised Program and Budget 2008/09 was approved by Member States in December 2008. This introduced a new strategic framework, incorporating a restructuring of programs the introduction of new programs, as well as changes to some expected results and related indicators.

B. Validation Exercise

14. As noted by the Audit Committee,[2] the role of WIPO's Internal Audit and Oversight Division (IAOD) is 'to validate the authenticity and reliability of the information on program deliveries'. The definition of validation referred to in the Review is 'the process of cross-checking to ensure that the data obtained from one monitoring method are confirmed by the data obtained from another method'. In the OECD/DAC glossary 'validity' is defined as 'the extent to which data collection strategies and instruments measure what they purport to measure'[3]. Independent validation of performance data is increasingly a recognized, standard requirement for any effective RBM process.

15. As will be explained further in the next section on Methodology, this is a preliminary validation exercise with the primary objective of learning how systems and practice can best be improved in the future, when more formal validation of data is likely to be conducted.

16. As preparation for this exercise, the following communication to staff took place:

(a) A memo dated 16th February 2009 was circulated from the Director General entitled 'Preparation of the Program Performance Report (PPR) for the year 2008'. In this it was stated that 'You should be aware that IAOD will conduct a validation exercise on the data used for the PPR 2008'.

---

[1] See p.3 of Internal Review on Program Performance Reporting Process carried out by WIPO's Evaluation Section, dated 10th October, 2008

[2] Document WO/AC/11/2, page 10, Agenda Item 4

[3] OECD/DAC 2002: Glossary of Key Terms in Evaluation and Results-Based Management,

(b) A separate memo was circulated on 23rd February 2009 by the Director of IAOD enclosing:

i) A 'Validation Approach' paper;

ii) A 'Frequently Asked Questions' paper

17. A presentation entitled 'Validation of the 2008 PPR' made by the Evaluation Section of IAOD on 24th February 2009.

18. The original intention, as outlined in the Validation Approach paper, was for the Validation Team to consist of two people – the Senior Evaluator in IAOD and an external consultant. With the secondment of the former to the Program Management Performance Section, the team, along with the number of programs to be covered by the exercise, was reduced by half. 15 programs were included in this Validation Exercise, with at least one expected result selected from each program, providing a sample size of approximately 10% of all expected results.

III. ORGANIZATIONAL CONTEXT

19. This Validation Exercise has been organized to support the 2008 Program Performance Report, which is due to be submitted to the Program and Budget Committee in June, 2009. At the same time planning for the next Biennium is being undertaken, along with the separate development of a Medium Term Strategy (MTS), designed to cover a six-year period.

20. Given the pressure on the time of managers of programs arising from the coincidence of these planning and reporting priorities it was gratifying that all the selected programs were able to put aside time for this exercise and, in almost all cases, have the relevant information available.

21. Following the recent appointment of a new Director-General, there has also been a significant restructuring and realignment of programs. In several of the programs selected for this exercise, this has meant that responsibilities for results or outcomes have migrated to other programs. This has had an influence on which results were selected to maximise the usefulness of the exercise, and on the availability, in some instances, of the managers responsible for delivering the predicted outcomes.

22. This initial Validation Exercise is the responsibility of IAOD and the report is submitted, in the first instance, to its Director. However, in the light of the primary emphasis of this initial exercise on lesson-learning, rather than accountability, it is hoped that the report will make a valued contribution to the quality of planning that is required for the next biennium proposal and the MTS. Responsibility for co-ordinating the 2010/11 Proposed Program and Budget and the MTS lies with the Program Management and Performance Section (PMPS) within the Department of Finance, Budget and Program Management. One of the objectives of this report is, as far as is practical, to align its recommendations with existing priorities and approaches being undertaken by the PMPS.

IV. METHODOLOGY

23. Full details of the design of this Validation Exercise, together with the roles and responsibilities of the Validation Team can be found in detail in the draft Validation Approach paper (see Annex 1) and the Terms of Reference for the independent team member (see Annex 2).

A. Selection of sample programs

24. The intention at the time the 'Validation Approach for Program Performance Report' was finalized was that the validation team would consist of two people and that all programs would participate in the exercise. With the secondment of the Senior Evaluator to the PMPS, the Team was reduced to a single member. As a consequence the decision was taken to reduce the number of participating programs to 15 out of a total of 31.

25. The following criteria were applied to select the participating programs:

a. Strategically most relevant;

b. Highest spend;

c. Substantively in place as a program throughout 2008;

d. Continuation of program from 2009 onwards.

26. The selection process was carried out against these criteria by: the Head of Evaluation, the Head of PMPS, and the Senior Evaluator. All selected programs participated in the exercise. For a full list of participating programs see the list of meetings and interviewees in **Annex 3.**

B. Selection of sample results

27. By the end of the first week of the exercise, 12 of the program managers responsible for reporting results had been visited. A briefing was provided on the scope of the exercise and sample results were agreed, based on the following criteria:

a. Most significant in relation to the contribution to change in the program;

b. Substantially to be retained for the remaining year of the current Biennium and, preferably extending to 2010/11;

c. Challenging or more fruitful for lesson-learning in comparison to others.

28. It was explained to participating managers that the collaborative nature of selection was exceptional for this 'trial' exercise, since the emphasis was on lesson-learning rather than accountability. In future, it would be unlikely that implementing managers would be invited to participate in the selection of either results or specific indicators.

29. A date for the substantive meeting was agreed and an e-mail circulated to each program manager confirming: the meeting arrangements; the selected result and indicators; the type of information that might be required to be produced; and the specific criteria to be applied.

C. Validation criteria

30. Program reports were submitted in advance by all but two of the programs. They followed the template and guidance circulated to all programs by the PMPS. Section A (Analytical Summary) of the template provided an opportunity for programs to identify significant progress, achievements and lessons learned. Section B, which required specific Performance Indicator Data (PID) to be reported, was the main focus for the Validation Exercise.

31. Following the guidance contained in the Validation Approach paper, the criteria used for assessing the authenticity, reliability and quality of data were ultimately identified under four headings, as follows:

Headings	Criteria
(a) **Specification and selection of data**	- Relevance to specific indicators and result - Sufficiency and comprehensiveness
(b) **Collection of data**	- Efficiency of collection and easy accessibility of data - Extent to which it is routinely produced
(c) **Analysis of data**	- Degree to which analysis is carried out - Consistency and comparability
(d) **Reporting of data**	- Accuracy and verifiability - Timeliness of reporting - Clarity and transparency.

32. Definitions of the above terms can be found in Annex 1 and 2 of the Validation Approach paper.

33. In addition, the following, preliminary questions were asked about the objectives, results and indicators:

a. Who developed the objective/results/indicators?

b. Do they still seem appropriate and relevant?

c. Do you make use of this information on a regular basis?

d. Does your line manager make use of this information, as far as you know?

34. Finally, the main learning points were identified, within the following categories:

a. Summary of data limitations;

b. Lessons learned for data quality;

c. Lessons learned for planning and monitoring in the future.

35. An aide-memoir was completed for each program by the Validator. The template can be found in Annex 4. A diagrammatic, non-attributive summary of findings for each program can be found in Annex 5.

D. Timing and conduct of Validation Exercise

36. Preliminary meetings took place in the week beginning 16th March 2009. Validation meetings were conducted over a two week period between 30th March and 3rd April 2009. The final draft report was submitted for approval at the end of April 2009.

37. Meetings were held with each implementing manager and, on occasions, other team members. A full list of meetings and participants is included in Annex 3. During the course of the exercise there were no meetings arranged with supervising managers or members of the senior management team, so their perspective and views have not been directly represented in this report.

E. Variations from Validation Approach Report

38. Whilst substantively keeping within the standards and intentions of the original design, the actual conduct of the exercise differed from the guidance in the following ways:

a. The validation team and sample numbers of programs was reduced by half due to the secondment of the Senior Evaluator of IAOD. The sample size represented approximately 10% of the total number of results and performance indicators.

b. The selection of sample results and indicators were agreed in collaboration with the designated implementing managers, to maximize the usefulness of meetings and immediate lesson-learning for current planning cycles. The selection of results did not use the screening and classification processes as laid out in the Validation Approach paper.

c. Following an early assessment of the quality of reporting in the PPRs and the pressure on availability of participating managers, a decision was taken to focus even more strongly than originally intended on training and lesson-learning as the priority. A consequence of this approach was that less time was available on the cross-checking of supporting data. For similar reasons, a greater emphasis was placed on identifying appropriate systems for the routine gathering of PID to enable monitoring and reporting to be done more efficiently and usefully.

V. FINDINGS

A. Introduction

39. As outlined in the previous section, 15 programs were selected to participate in this validation exercise, with one result and its supporting indicators agreed on for each program. This represents just under half the total number of programs (31) and just over 10% of the total number of results (141 in total) contained in the original Program and Budget for the 2008/09 Biennium.

40. This exercise was set up to validate information at the half-way stage of the Biennium. Although any targets set would not be expected to have been reached, it should be clear from the Performance Indicator Data (PID) how much progress is being made to reach any result or target, and, where appropriate, the likelihood of success at the end of the Biennium. Where indicators, rather than targets, have been selected some analysis or interpretation can be expected to derive and communicate 'meaning' and 'value' from the data.

41. Meetings with managers responsible for implementing the individual programs took place over a two week period, with up to an hour allocated for each meeting. All meetings took place as agreed, with only one designated manager unable to participate, due to absence on sick leave. Despite substantial pressures on the time of managers, not least in relation to the planning demands for the next Biennium and the Medium Term Strategy, all meetings achieved the objectives and, on average, lasted just under 55 minutes. A list of meetings and participants is attached as Annex 3.

42. With one exception, all PPRs were submitted in advance of the meetings. In the case of the exception, general principles of quality reporting were discussed at the substantive meeting, with specific comments subsequently fed back in writing on the PPR and the selected result.

43. The meetings were structured to elicit information in line with an aide-memoir that can be found in Annex 4. A non-attributive summary of findings for each program can be found in Annex 5.

44. The decision to make training and lesson-learning the focus of the validation meetings was informed by the conclusion arrived at early on that some managers revealed too little familiarity with a results-based approach and with the concepts that lie behind it (e.g., outcomes, indicators, baselines, benchmarks).

45. The main findings are presented under the following headings:

---

a. Data quality (in relation to the nine specific criteria)
b. Significant lessons learned:
   - For data quality
   - For planning and monitoring

c. Other significant findings

B. Data quality

i. Background information sought

46. In addition to the nine criteria, four preliminary questions were asked to elicit information about the context in which the PID was generated:

a. Who developed the objective, results and indicators that were selected for this exercise?

b. Do they still seem appropriate, one year later?

c. Is the information used by you on a regular basis?

d. Is the information used by your supervising manager(s) on a regular basis, as far as you aware?

ii. Overall findings

47. For the purposes of providing a simple overview of findings, responses were categorized in three ways: criteria substantially met, partially met and insufficiently met.

48. The following figures summarizes the findings by program, in relation to the nine criteria and two questions:

Figure 1: Findings by program in relation to the nine criteria

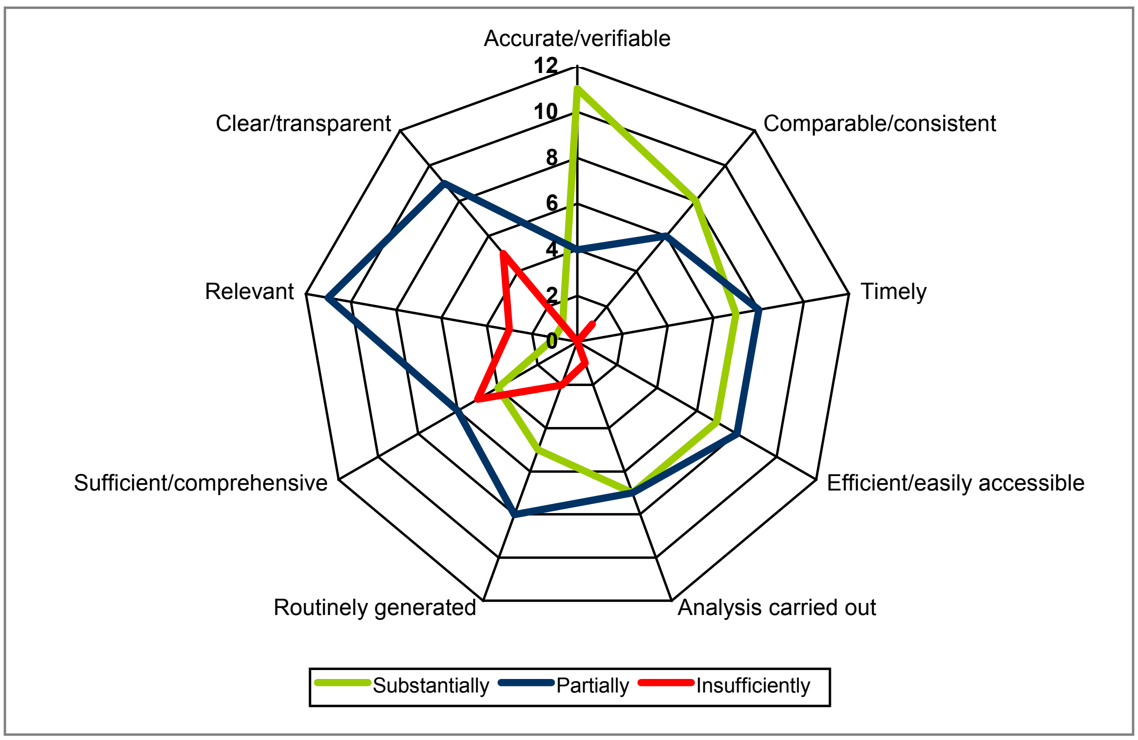
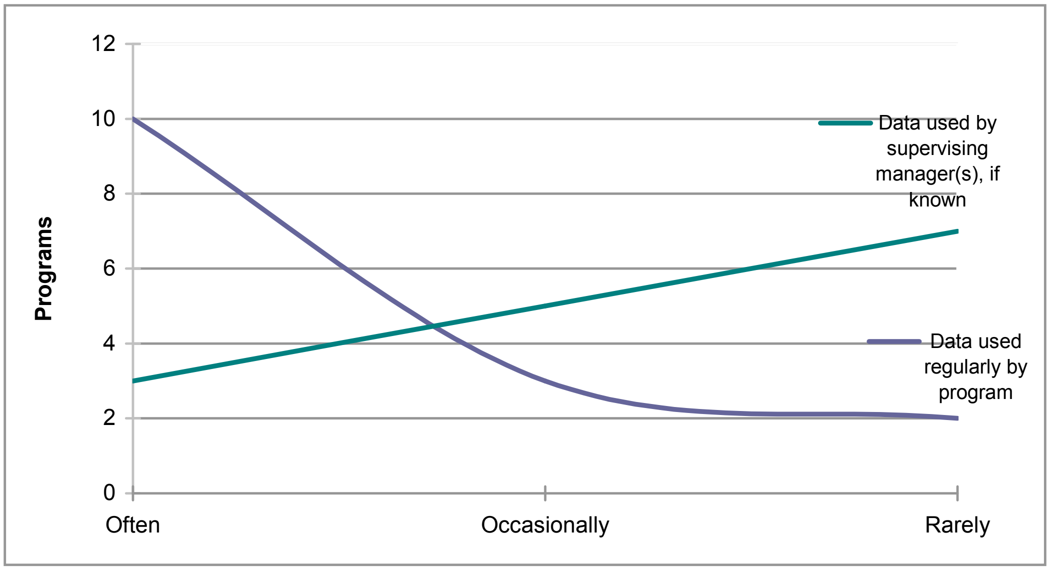


Figure 1: Findings by program in relation to two questions



iii. Data quality and limitations

49. Taken in order of success in substantially meeting the criteria, the following information provides some background to support the assessment. In two programs indicators had been deemed to be no longer valid and have not been included in this assessment.

Accuracy and verifiability (11/4/0):

a. These two criteria were brought together on the grounds that accuracy could only practically be established by cross-checking the information with source documents.

b. For the most part, the data presented appeared to be accurate and verifiable. Little cross-checking was actually carried out because of time constraints and the amount of ground needing to be covered. However, sufficient information was provided to provide assurance that PID was reliable.

c. In those instances where PID only partially met the requirements for quality data, the following limitations were observed:

- The systems for collating the data were not yet fully established;

- A list of activities was provided which were likely to be verifiable but were irrelevant;

- Reporting was too vague to allow full verification;

- One indicator required some element of benchmarking that was not available.

Comparability and consistency (8/6/1)

a. Primarily, the focus of this was the extent to which the data could be compared with previous years or with identified benchmarks. There was little meaningful opportunity to compare progress with similar programs.

b. The main measure used here was the ability to compare performance against previous years. Unsurprisingly, changes in the structure and organization of programs were the main contributor to the difficulties in making comparisons with previous years.

c. The main reasons for the partial or complete lack of comparability and inconsistency were:

- PID consisted of a list of activities;
- Insufficient data was produced;
- Priorities fluctuated too often from year to year;
- Where comparisons were difficult because indicators related to ‘process’ rather than ‘product’;
- Only one of two indicators met the requirements.

Timely reporting (7/8/0)

a. The focus here was on establishing whether the collecting and reporting of data were carried out at a frequency which suited the various purposes for which it was required.

b. Overall, the timeliness of reporting was generally fit for purpose. The assessment was made in relation to the purposes for which the data was used. The results are likely to have been different had more indicators been selected for internal management and monitoring purposes rather than primarily for annual reporting.

c. The reasons behind the partial or inadequate meeting of criteria can be summarized as follows:

- Data are not asked for and used only for reporting for PPR;
- Data are substantively not being collected;
- Reporting is done only for external purposes, and not for any internal monitoring;
- Systems are not yet in place to report more regularly.

Efficiency and easy accessibility (7/8/0)

a. Here the approach taken was looking for evidence that systems for recording and collecting data were not too wasteful or costly in terms of time and resources required.

b. Some programs have developed sophisticated and efficient systems for collecting data, whilst others have basic, manual systems which, nevertheless, may be fit for the purposes required. Several programs are in the process of developing improved systems following recognition of the importance of the information needed and the time wasted in having to collate information manually, even if only annually.

c. Those programs that had not yet developed efficient systems are characterized by the following:

- Those experiencing difficulties in collected information from a number of regions or Bureaux, when systems aren't in place to ensure more organized gathering and collating of PID;

- Those relying on manual calculations when a computerized solution would be preferable or which have yet to develop a database, however simple.

Analysis carried out (7/7/1)

a. Although not a traditional criterion for validation purposes, it was applied here to draw out evidence that some processing of information had taken place to extract some meaning and value from the data, for use in monitoring progress.

b. Only in one instance was there no evidence at all that PID were applied to assess progress against the expected results. This was a program where a list of activities was generated, with little suggestion that the list was used for monitoring progress.

c. In those cases where some analysis appeared to have been conducted, more was not carried out for the following reasons:

- The collection of data was so time-consuming that there was limited opportunity to process it;

- PID were mostly straightforward and self-explanatory;

- The information was mostly generated for external purposes;

- PID were activity-oriented and analysis was not deemed to be important;

- There was insufficient disaggregation of the data to make any meaningful analysis;

- The indicators are no longer as relevant as when they were designed;

- PID have not been used previously for decision-making and there was little anticipation that they would be in future.

Routinely generated (5/8/2)

a. This criterion was applied to find out whether appropriate systems were in place for capturing data in a timely and regular manner.

b. This criterion is linked to the previous one. It is also closely linked with the purposes for which PID are used. If PID are used for regular monitoring of progress or there is an importance placed on the information by senior management, then systems are more likely to be put in place to generate the data routinely. If data are only seen to be useful for reporting externally, and often only annually, then routine systems are unlikely to be put in place.

c. The two programs that lacked a system for routine generation of PID were those that agreed that they would find it useful to have information available more regularly, but as yet did not have a system in place. Both were making it a priority to developing mechanisms for capturing the data in real time.

d. Those programs that had not fully reached the point of developing a routine data collection system, had the following attributes:

- There is little demand for the provision of regular information, other than annually;

- The information is not used for monitoring purposes;

- Systems are in the process of being developed;

- The information is sufficiently straightforward not to require any separate recording system.

Sufficient and comprehensive (4/6/5)

a. Although succinct reporting is looked for, the information provided needs to be sufficiently complete for the reader to be able to assess progress and success against the indicators.

b. The last three criteria are inter-related which explain why they arrive together at the lower end of the table. Compliance depends on an understanding of the purpose of PID and a confidence that specific, pertinent information that directly supports the indicator is what is primarily required.

c. Those that partially or wholly failed to meet this criterion fall within the following:

- There was an absence of baselines or benchmarks when indicators require a comparison;

- The wording of PID was too imprecise to enable the reader to understand adequately what progress was being made. For example, the use of words such as ‘several’ or ‘a number of’ are unhelpful, particularly when specific targets are being supported;

- There was too little information provided to support the indicators, including unsubstantiated statements;

- The information was poorly organized and did not allow the reader to easily draw conclusions.

Relevant (1/11/3)

a. This is self-explanatory in that the information should clearly be supporting either the selected result or, more specifically, the selected indicators. The information should not extend beyond what is required to support these.

b. Almost all programs failed to meet the criterion fully because, whilst some may have provided sufficient information, there was a lot of information provided that was not necessary to support either the indicator or the result. This tendency to provide too much information was generally explained by an uncertainty or lack of confidence about what was required, and a determination to reflect as much success or progress as possible.

c. Those three programs that failed completely to provide relevant information revealed the following:

   - Activities, inputs and outputs were recorded in two programs when outcomes were required by the indicators and result. A justification that outcomes could be assumed from the inputs was not accepted, since some other form of indicators, such as ‘proxy’, more activity-oriented indicators, could have been selected.

   - One set of PID was over-elaborate and confusing, to a point where the relevance to the indicators was lost.

d. Characteristics of those programs that insufficiently met the criterion were:

   - More information than was required was included. In one case the justification was made that Section A did not allow sufficient room to include all achievements and the program did not want to lose the opportunity to record them.

   - The wording was too imprecise. For example, in one program it was not made clear whether the figures given were cumulative over the past few years or confined to 2008 alone.

   - PID were relevant to the result, but not to the specific indicator.

   - Too many details were provided about the methodology behind PID, rather than a more succinct focus just on the results of the approach taken.

   - In one instance, the information needed to be disaggregated to properly inform the result, even if strictly it was sufficient to support the indicator.

Clarity and transparency (1/9/5)

a. This turned out to be the most demanding criterion as it often proved difficult for the reader to identify the precise degree of progress being made from the data presented, and to be helped to make confident decisions from them.

b. Those programs that fell well short of best practice did so on the following grounds:

- Frequently, the language used in the reporting was markedly different from that used by the indicator, sufficient to leave the reader unsure as to whether there was a direct link between the information required by the indicator and PID;

- A list of activities was provided without any evident links to the outcomes reflected in the result or indicators;

- Baselines or benchmarks were not referred to when required;

- In one instance, the 'successful' figures required to support the indicators were available, but were not disclosed in order to avoid the risk of the second year appearing to be a failure in comparison, should progress not appear as good.

c. Those programs that partially achieved compliance against this criteria fell short for similar reasons, although not so markedly. These included:

- Language not linked sufficiently tightly to that used in the indicators. Examples were: the requirement to identify the number of 'centers established' when 'technical assistance was provided' was used; reports of the number of incidents 'closed' rather than 'resolved' as required by the indicator;

- Poor organization of information leading to a difficulty in working out which part of the PID was intended to inform progress against the indicators;

- The need to disaggregate the data to identify what was really of value from what was more 'cosmetic' and insignificant, even if not required strictly by the indicators.

C. Significant lessons learned

i. For Data Quality

50. The results chosen for this exercise were not randomly selected and cannot, therefore, be considered to be a genuine sample. In addition, only half the total number of programs was involved which confines the direct lesson-learning to half the implementing managers.

51. In summary, the following can be extracted usefully from the detailed findings:

a. Generally, the reporting of progress against results and indicators appeared to be accurate and verifiable.

b. Despite the changing political and organizational environment, there is sufficient consistency in the indicators used to be able to make comparisons over the years, and this strength would appear to be maintained even when there is a restructuring of programs as is currently taking place.

c. Reporting frequency inevitably reflects the demand for information. The programs that are a priority for the attention of senior management are thereby encouraged to develop systems that can produce required data in a timely manner. Similarly, those implementing managers that put a priority on monitoring and communicating progress regularly, have similarly designed data recording, collection and reporting systems that meet their needs.

d. Those programs that are required to report to senior management regularly and those that have developed effective monitoring systems as an internal management tool mostly meet the requirements not only of timely reporting, but also of efficient and routine data gathering, together with the need to analyze progress and make comparison either with previous years or with external standards or benchmarks.

e. Almost all programs have some room to improve the organization of their reporting against indicators so that an assessment of progress can be clearly made both by external and internal stakeholders/readers against relevant, sufficient and transparent PID. The quality of reporting would be enhanced particularly by ensuring that:

- Attention is paid to match the wording of the indicators with the wording of the PID;

- Data that is not strictly relevant to the selected indicators and results should be reported elsewhere, for example, in Section A (Analytical Summary) of the PPR reporting template;

- Appropriate, often simple, systems being developed to capture data required for monitoring progress indicators on a regular, routine basis;

- Ensuring that baselines and benchmarks are included in the report when comparisons are required, e.g. when words such as ‘increased’, ‘improved’. ‘maintained’ etc. are used in the results or indicators;

- A better understanding on the part of some implementing managers of the importance of thinking more in terms of ‘outcomes’ and ‘results’ rather than ‘activities’ and ‘outputs’;

- An explicit identification of those indicators that are no longer valid or relevant, with some brief justification for the conclusion.

ii. For Planning and Monitoring Systems

52. Although the scope of the validation exercise is confined to an assessment of the quality of data used in reporting against PID for the PPR, there are important implications arising from the findings for the quality of planning and monitoring within the Results-Based Management initiative.

53. The following are some of the points that arise directly from the findings of the exercise:

a. Targets are useful as more specific and tangible measures of progress and success, rather than indicators that do not require any prediction of quantitative results. However, the use of targets is more relevant where the political or organizational environments are stable and where outcomes are capable of being predicted. It is helpful if there is clarity and agreement about when either indicators or targets are more useful in any individual program's context.

b. Indicators and targets are most valuable and valued when they are required to be supported by information that is also useful for internal monitoring purposes, and not only for reporting externally. The programs that demonstrate better quality of data generally use PID for other purposes, such as policy development, individual or project performance management, or for assessing operational efficiency and effectiveness.

c. There is an acknowledged difficulty in finding good indicators for programs that are primarily responsive and 'demand-led', with an emphasis on 'process' rather than 'product'. Some programs were successful in finding appropriate indicators that focused on the achievement of successful processes, such as the fact of decisions being made, cooperation effected or agreement reached, without having to identify what should be decided or implemented.

d. Some programs have already taken the lead in developing workplans that are linked directly to the objectives, results and indicators that are included in the Biennial Programs. Embedding the habit of thinking about 'impact', 'results' and 'outcomes', rather than activities, inputs and outputs, throughout a program team is clearly a positive development in the integration of a results-based management approach.

D. OTHER SIGNIFICANT FINDINGS

54. In the course of the discussions with implementing managers a number of views and suggestions were presented which have a bearing on the conclusions and recommendations identified in the sections below. These included:

55. Where impact and change are looked for a 'gap' analysis in relation to potential areas for greater focus by individual programs can be usefully carried out before specific results and indicators are selected. Examples given were 'geographical' gaps and gaps in the stages in which groups of countries are at in terms of their development. Similarly, it was suggested that more needs assessments could usefully be carried out to better inform the appropriate choice of objectives.

56. With the recent announcement by the Director General of nine strategic goals and the current development of a Medium Term Strategy, there is are excellent opportunities presented to ensure that individual program objectives and results are more directly and tightly linked to the organization's longer term goals and objectives. Furthermore, the development of the Performance Management and Staff Development System (PMSDS) will also assist in bringing 'alive' the results-based management system and encourage staff to recognize that monitoring progress against results and indicators has value beyond the annual reporting to Member States.

57. Some implementing managers interviewed suggested that further, useful support could be provided in two main ways:

a. Technical assistance in the development of databases and management information systems that contribute to a more efficient recording, collating and organization of data required for planning, monitoring and reporting. In this respect some sort of systematic needs analysis to assess the level of support required for all programs might be appropriate.

b. A more coordinated approach to acquiring user feedback, so that individual programs do not have to go to the time and expense of arranging for separate surveys of users, when the information could be more efficiently acquired on a cross-program basis. The recently announced Customer Service initiative was suggested as providing a significant vehicle for a more coordinated approach.

58. Some managers interviewed revealed that they had an insufficient grasp of some key elements relating to the results-based management initiative, including a clear understanding of concepts behind results-based management, such as 'outcomes' and 'outputs', 'targets and indicators', 'baselines' and 'benchmarks'. Many acknowledged that the culture of the organization was still rooted in the primacy of reporting on the range and quantity of activities, rather than adapting to the requirements of a results-based approach.

59. It was pointed out, on several occasions, that the information required to be generated for reporting against indicators and results was not often sought by supervising managers. The response to the preliminary question about usage of PID by supervising managers clearly supports these views. In this respect there does not appear to be either a strong 'monitoring' culture or visible monitoring systems, save in those programs where individual implementing managers choose to design and implement tailored systems for their own programs. It was felt that more interest shown by supervising managers in finding out about progress being made on a regular basis against selected indicators and results would lead to a greater value and effectiveness resulting from the RBM approach.

VI. CONCLUSIONS

60. This 'trial' validation exercise was part of WIPO's ongoing process of supporting the continuing improvement of good practice in a results-based approach to management. It complemented a number of current initiatives which are seeking to influence the development of the organization towards greater impact and effectiveness. An important contribution towards this development is the improvement of mechanisms for planning, monitoring, evaluating and reporting on progress and success within the context of overarching strategic goals.

61. This exercise has revealed the expected mixture of good and poor practice in the application of quality data for reporting on performance. Within the 15 programs that participated in the exercise there was a wide range of expertise and sophistication in the methods and mechanisms adopted for planning, monitoring and reporting.

62. The factors that appeared to influence a higher quality of data collection, analysis and reporting can be identified as:

a. The interest of senior management in the progress being made in certain programs;

b. The commitment of individual implementing managers to monitoring progress on a regular basis;

c. The selection of objectives, results and indicators that have value to the programs for internal management purposes and not just for the annual reporting to Member States. Policy development, performance assessment and operational efficiency are examples of purposes that stimulate a greater commitment to selecting and utilizing performance data;

d. The grasp that implementing managers have of the concepts and value behind a results-based approach, in contrast to the emphasis placed on reporting on the range and quantity of activities;

e. The development of routine data collection and reporting systems for capturing PID, whether through appropriately simple tables, spreadsheets, and databases or more sophisticated, purpose-designed software such as the one used in the Arbitration and Mediation Center. It was pointed out that these data collection systems are usually much more useful where they are primarily designed and developed internally, rather than contracted out externally;

f. Clear linkages made between individual program objectives and the overarching strategic goals and priorities of the organization as a whole.

63. There is little evidence that programs deliberately report in a misleading or inaccurate way to disguise the amount of progress made. On the contrary, there was, invariably, a refreshingly open and frank assessment of limitations and difficulties, as well as pride in the aspects that were working well. Aside from the reflection on the integrity and commitment of individual implementing managers, there are also organizational influences that can be seen to contribute to the reliability of reported data:

a. There are few observable consequences for failing to meet expected results or incentives for successfully achieving the predicted results and targets. The new PMSDS initiative will undoubtedly make a difference in this respect;

b. The results-based management approach is currently applied more in financial planning cycles and reporting to Member States, rather than in day-to-day monitoring or as a routine management tool.

64. The main limitations in data quality arise mostly from a lack of awareness of what is required as good practice in the reporting of information to support results and indicators, rather than any intention to confuse or mislead. As a result there is a tendency for some programs to report in a lengthy, over-elaborate, and imprecise manner, as well as a reluctance to focus on predicted outcomes rather than activities and events that have been organized.

65. There is scope for the organization to provide greater support to implementing managers and staff, in particular, by:

a. Raising the profile and visibility of senior managers in supporting implementing teams in sound planning, monitoring and reporting habits. At present there appears to be too strong a perception that results based management is driven ‘externally’ by the IAOD or the Performance Management Unit, rather than being a priority with senior management as an integral part of day-to-day management responsibility;

b. Providing more technical assistance for the development of appropriate management information systems and databases for the efficient collection, organization and reporting of data used to monitor performance;

c. Coordinating user feedback at a cross-program and organizational level rather than leaving the organization of surveys and other mechanisms to individual programs. The new customer service initiative is likely to be an important vehicle for introducing a more coordinated approach to user satisfaction levels.

VII. RECOMMENDATIONS

66. Following the main findings and conclusions of this validation exercise, the following are the principal recommendations arising out of this validation exercise:

67. As part of the planning cycle for the next Biennium and the Medium Term Strategy, consideration should be given by senior management to carrying out a review of the extent to which current monitoring systems need to be strengthened in order to integrate the RBM approach more fully as a routine, day-to-day management function.

68. The quality of reporting data will be immeasurably increased with a more highly visible participation of supervising managers in providing monitoring support to implementing managers. Particular attention should be paid to the offering of guidance in the selection of results and indicators and their direct linkages with higher strategic goals, and also in the active monitoring of progress on a regular basis.

69. The organization will do well to maintain its support for implementing managers in the practical benefits of understanding and valuing the results-based management approach, not only for their increased ability to recognize and communicate progress and achievements to various stakeholders, but also in their role of increasing their day-to-day management skills.

70. Additional practical support should also be considered in the following areas:

a. Making more technical assistance available for the development of computerized data collection, analysis and reporting tools;

b. Coordinating the collation of user feedback of relevance to multiple programs;

c. Maintaining the availability of one-to-one training and support in relation to good practice in the identification and framing of appropriate objectives, results, indicators and targets, baselines, benchmarks, milestones, risks and assumptions, etc.;

d. Producing a monitoring tool for capturing data relating to progress against selected indicators, that can be used for reporting to senior management and stakeholders on a regular basis.

ANNEXES

1 Draft Validation Approach Paper
2 Terms of Reference for Validation Exercise
3 List of meetings and participants
4 Aide-memoir template
5 Summary of findings tabular form

ANNEX 1 VALIDATION APPROACH PAPER

WIPO

WORLD INTELLECTUAL PROPERTY ORGANIZATION

GENEVA

Reference: EV/01/2009

INTERNAL AUDIT AND OVERSIGHT DIVISION

Evaluation Section

VALIDATION APPROACH FOR PROGRAM PERFORMANCE REPORT

February, 2009

TABLE OF CONTENT

INTRODUCTION

1. Each year, WIPO assesses its annual performance based on an approved performance framework. For this purpose, the Organization utilizes performance information that is complete and actual (rather than projected) for almost all of its performance measures.

2. Good quality information is crucial if performance measures are to be used effectively to improve service delivery and accountability. Such information should help the Organization to: a) manage delivery against priorities; b) report reliably on its achievements; and c) assess whether WIPO needs to revise policies and programs.

3. Within the Results-Based Management approach adopted, one of the Organization's aims is to provide reliable, timely and accurate information so that it can track performance over time, identify the need for any remedial action to achieve expected results and report clearly to its stakeholders on performance. Therefore a validation approach to all program performance reports has been introduced within the Organization in order to ensure that the information in these reports meets certain quality standards.

4. To facilitate the validation exercise, IAOD's Evaluation Section has prepared the present approach paper which will guide the validation team and inform program staff during the process.

OBJECTIVE

5. The purpose of validation is to have a thorough independent assessment by a Third Party (validation team) of authenticity and reliability of the information on progress on program deliveries. The validation team will assess whether the information used in individual program performance reports is: relevant, timely, accurate, consistent, verifiable and transparent; and if there has been a continuity of information. Further details on the definitions of the mentioned criteria are described in Annex 1 of this approach paper. The validation will finally result in a conclusion by the validation team whether the quality of information used for reporting against defined results meets the above defined validation criteria.

6. This assessment by the validation team will, inter alia:

a) Assess the evidence which supports the contribution reported to specific key performance indicators, expected results and ultimately to the relevant strategic goal(s);

b) Ensure that the approved performance framework (document WO/PBC/13/4) is being applied;

c) Assess information on program delivery by applying the above mentioned criteria. Further details on validation criteria are presented in the Annex 1 of the approach paper.

SCOPE

7. For the validation exercise only one expected result by program will be selected for validation. Further details on the sampling process for expected results has been described on page 7, Step two of this approach paper. The sample size will represent 20 per cent of all expected results. The validation team will validate the authenticity and reliability of the information on program deliveries. The information required by the validation team will vary from program to program and will depend on what has been reported. Some examples of information that might be required include: monitoring data, statistical data, workshop papers, mission reports, reviews, studies, MoUs, records, etc.

8. Program managers shall keep records of the information that is being used for reporting performance within their programs.

9. For validation purposes program managers, upon request, will need to make available all relevant information used for reporting purposes as part of the program performance report. The validation team may request clarifications, including validation of specific documents.

10. Subject to considerations of confidentiality, original documents shall be shown to the validation team to facilitate the validation process. If required, copies of original documents might be requested.

11. It is to note that the validation work will only assess the quality of information underpinning the performance framework used to report to Member States. The validation team will not validate the quality of expected results and indicators, nor will the team provide conclusions about the accuracy of all expected results but rather validation of a representative sample of 20 per cent of all expected results.

IAOD'S ROLE

12. As indicated by the Audit Committee (document WO/AC/11/2, page 10, Agenda Item 4) and following the assignment of responsibility for preparation of the Program Performance Report to the Program Management and Performance Section, Department of Finance, Budget and Program Management (see paragraph 16(b) (i) there), the role of IAOD is to validate the authenticity and reliability of the information on program deliveries.

13. The validation exercise is undertaken prior to the publication of the Program Performance Report which, according to regulation 2.14 of WIPO's Financial Rules and Regulations, the Director General shall prepare. The Program Performance Report is on program performance, based on the program structure, results frameworks, benchmarks and performance indicators, expected results and targets contained in the program and budget, in accordance with the mechanism adopted by the Member States in respect of their involvement in the preparation and follow up of the program and budget of the Organization.

14. Overall, IAOD will:

a) Prepare, in advance an approach paper which will be shared with program managers prior to the validation exercise;

b) Define a sample for the validation exercise (See further details on page 7 of this approach paper);

c) Validate 20 per cent of all expected results defined in the Program and Budget Document;

d) Prepare a validation report which will include findings, conclusions and recommendations.

Additional principles to be applied by the validation team

15. The following principles shall be applied in performing validation and shall also be used as guidance when documents related to validation are prepared.

a) Impartiality and independence: the validation team shall remain independent of the activity being validated and free from bias and any real or potential conflict of interest. The validation team shall maintain objectivity throughout the validation process to ensure that the findings and conclusions are based on objective evidence generated during the validation and are not influenced by other interests or parties;

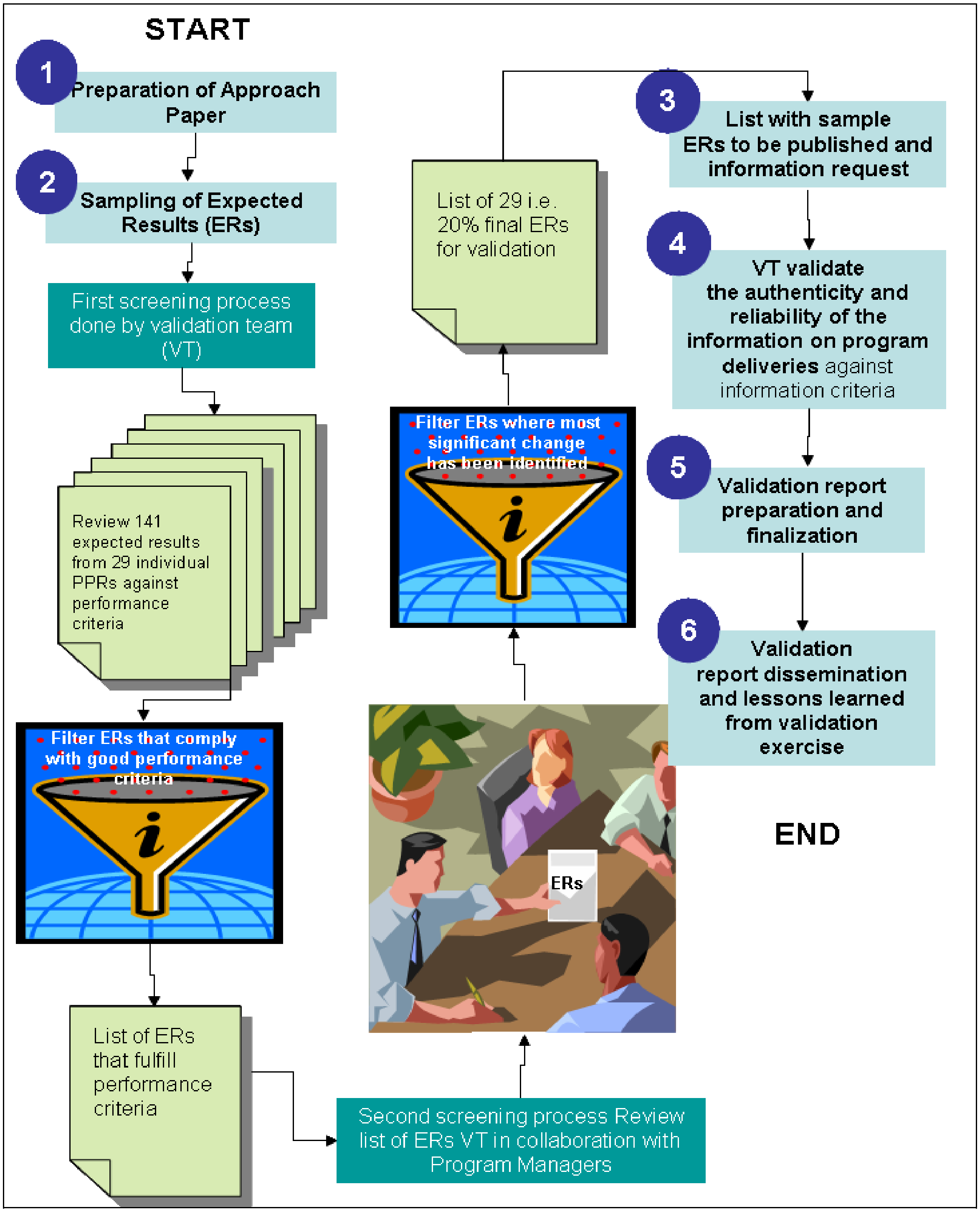
b) Ethical conduct: the validation team shall demonstrate ethical conduct through impartiality, independence, integrity, confidentiality and discretion throughout the validation process;

c) Fair presentation: the validation team shall reflect truthfully and accurately validation activities, findings and conclusions in the report;

d) Confidentiality: the validation team shall ensure that confidential information obtained or created during validation activities is safeguarded and not inappropriately disclosed.

VALIDATION PROCESS

16. Overall the proposed validation process will consist of the six steps that are illustrated in the figure below:



Step one: Preparation of approach paper

17. The present paper has been prepared by IAOD in order to guide the validation team and inform program managers and stakeholders on the validation approach.

18. The paper will serve as a reference and can be modified by IAOD if necessary. The paper also defines validation as well as the criteria that will be applied when validating information.

Validation Definition

19. According to the OECD Glossary of Statistical Terms[4], "Validation" describes methods and processes for routinely assessing source data – including monitoring data, sample surveys, and administrative records – and how the results of the assessments are monitored and made available to guide processes.

20. It also describes how intermediate results are validated against other information where applicable, how discrepancies in intermediate data are assessed and investigated and how discrepancies and other potential indicators or problems in outputs and outcomes are investigated. All controls made in terms of quality of the information to be published or already published are included in the validation process.

21. Validation also includes the results of studies and how they are used to inform processes. In this respect, two dimensions can be distinguished:

(i) validation before publication of individual program performance reports, and

(ii) validation after publication of individual program performance reports.

Step two: Sampling of expected results

22. The validation team will validate the authenticity and reliability of the information on program deliveries. However, taking into consideration the limited amount of time available to produce the validation report and the large amount of information used to report on 257 indicators and expected 141 expected results, a sample for this exercise is required. The sample will consist of a selected number of expected results and indicators from each program with the purpose to provide information on whether the information used for reporting is consistent and accurate.

23. For the sampling exercise all expected results will undergo a two-stage screening process:

First screening process

24. The validation team will first review all expected results based on the UK National Audit Office's criteria for "Good Performance Measures"[5]: relevance, attribution, timeliness,

[4] http://stats.oecd.org/glossary/
[5] *Choosing the Right Fabric.* A Framework for Performance Information, Cabinet Office & HM Treasury, March 2001.

reliability, comparison and verification. Further details on the criteria for "Good Performance Measures" can be found in Annex 2 of this approach paper.

25. For the sampling exercise the validation team will make use of stratified sampling, where all expected results will embrace a number of distinct categories. For instance: a) expected results that fully fulfill the good performance criteria; b) expected results that partially fulfill the "good performance criteria; and c) expected results that do not fulfill the "good performance criteria". The frame can be organized by these categories into separate groups.

26. The two main reasons for using a stratified sampling design are a) to ensure that particular expected results within a program are adequately represented in the sample, and b) to improve efficiency by gaining greater control on the composition of the sample. By reducing the validation sample, IAOD would be able to reduce the cost of the validation exercise including the reporting time.

27. All expected results will be rated against the defined criteria and will be classified into three groups: a) expected results that fulfill the good performance criteria; b) expected results that partially fulfill the good performance criteria; and c) expected results that do not fulfill the good performance criteria.

28. All expected results defined in the Program and Budget Document will be considered during the first screening process but only expected results that partially or fully fulfill the criteria would be considered for the second screening round.

Second screening process

29. As part of the second screening process, a sample will be selected from the two first groups (a) expected results that fulfill the "good performance criteria" and (b) expected results that partially fulfill the "good performance criteria".

30. The validation team, in collaboration with the respective Program Managers and the Program Management Performance Section, will select within their program one expected result for which **"Most Significant Changes"** have been registered.

31. Through this exercise the validation team will identify 29 expected results that will undergo the validation process, i.e. there will be one expected result from each program that will be considered for validation.

Step 3: Publication of sample and information request

32. The final list of expected results for validation will be sent to all program staff in advance. The sample of the list of expected results will be accompanied by a list specifying the information required for validation. **All programs will need to make the requested information available to validation team in a timely manner.**

Step 4: Validation of sample

33. After the sampling exercise has been finalized, the validation team will validate the authenticity and reliability of information used to report against 29 expected results and their respective indicators by measuring the quality of information. For this purpose, the following criteria will be used: relevance, timeliness, accuracy, consistency, continuity, verification and transparency. A definition of each of the above mentioned criteria is provided in Annex 1.

34. The validation team will also apply standard validation techniques to confirm the correctness of the information including *inter alia*:

a) Document Review:

    (i) Review of data and information to confirm the correctness of presented information.

    (ii) Crosschecks between information provided in the program documentation and information from independent background.

b) Follow-up interviews (on site, via telephone and/or via email) using just some key questions (see below) for program staff:

    (i) Interviews shall include relevant stakeholders, staff responsible for program design and implementation, and other stakeholders as applicable.

    (ii) Reasonable crosscheck of information provided by interviewed personnel, i.e. by source check or other interviews, to ensure that no material evidence likely to be available to relevant stakeholders has been omitted from the assessment.

c) If necessary, comparison with programs that have similar or comparable characteristics;

d) Review of the correctness of critical formulas and calculations.

Example of key questions for program staff

35. The key questions that the validation team could pose to program staff in regard to their information systems are as follows:

a) Has the quality of information (relevance, timeliness, consistency, etc.) required to track progress against the target been considered?

b) Are performance measurement definitions clear and easy to understand?

c) Have responsibilities for ensuring data quality been allocated?

d) Are risks to data quality assessed?

e) Are users of information made aware of any weaknesses in reported data?

Step 5: Reporting

36. The validation team will draw up a validation report which describes the purpose, the nature and extent of the validation exercise to provide an appropriate basis for the conclusions, and, where applicable, subsequent recommendations. This approach should allow the readers of the report to understand the work done and to follow the logical links between findings, conclusions and recommendations. In cases where original documents were not made available by program staff, the team will need to consider to what extent this would have an impact on the objective of the validation to report on the adequacy of the system of accountability.

37. The validation report shall give the final conclusions regarding the program conformance with relevant information quality requirements. The validation report may raise issues that need to be subsequently addressed.

38. The validation report shall give an overview of the validation activities carried out by the validation team in order to arrive at the final validation conclusions. Apart from this, the general discussion of details captured by the validation team and conclusions related to project requirements shall be included in the final report.

39. The validation report will provide information at least on:

a) A summary of the validation process and its conclusions;

b) The validation team;

c) Internal quality control;

d) A list of the interviewed person;

e) A list of documentation reviewed

f) All findings and conclusions in regards to quality of data.

Step 6: Validation report dissemination and lessons learned

40. The validation report will be made available to all the stakeholders of the Organization and lessons learned from this process will be summarized.

ANNEX 1: Definition of criteria for validating information on program delivery

1. In order to facilitate the validation process the validation team will apply an adapted version of the "Good practice criteria for data systems" defined by the UK National Audit Office[6]. The information used for reporting on program delivery should be:

a) **Relevant** to what the organization is aiming to achieve. The quantification and reporting shall include information that covers all significant aspects of performance expressed in the expected results and indicators. Data collection methods, criteria and assumptions shall not be misleading. Data and assumptions that do not have an impact on the validation opinion shall not be included;

b) **Timely**, producing information regularly enough to track progress, and quickly enough for the information to still be useful;

c) **Accuracy** - accurate enough for its intended use, and responsive to change. The principle of accuracy requires reduction in bias and uncertainty as far as is practical. Accuracy with reference to the validation is required at two levels.

- The first relates to the accuracy of quantitative data and information;
- The second relates to accuracy of non-quantitative information.

d) **Consistency**, information shall address comparable key indicators that enable meaningful comparisons. The principle of consistency shall not prevent the use of more accurate procedures or methods as they become available. However, any change in procedures and methods shall be transparently documented and justified. Consistency is satisfied by:

- Application of the requirements of the methodology over different periods;

- Similarity of application of available guidance and knowledge among projects and programs with similar characteristics such as application of methodology, use of technology, time period and regional similarities;

- Applying tests and assumptions equally across potential baseline scenario;

- Ensuring equivalent application of principles used for expert judgment, internally and externally, over time and among projects and programs.

e) **Continuity of information** with either past periods or similar programs elsewhere. There are a number of reasons why continuity of measurement is important. Firstly, achieving program performance improvement may involve serious and structural change of the kind that is unlikely to be delivered over the short-term. Such changes will usually take a while to "bed-in" and start affecting results. Secondly, changing how program performance is measured can lead to confusion and lack of focus amongst staff and uncertainty over what they are working towards. Thirdly, in order to make judgments about how the Organization is doing, it is useful to have a good run

[6] see footnote 2

of comparable information. If programs change what is being measured, it will be difficult to make year on year comparisons; and

f) **Transparency** is to disclose information to allow intended users to understand and to make decisions with reasonable confidence. Transparency relates to the degree to which information is seen to as being reported in an open, clear, factual, neutral and coherent manner based on documentary evidence. Information shall be recorded, compiled and analyzed in a way that will enable internal reviewers and external intended users to attest its credibility. Transparency requires, inter alia:

- Clearly and explicitly stating and documenting all assumptions;
- Clearly referencing background material;
- Stating all calculations, methodologies and all information used;
- Clearly identifying all changes in documentation;
- Compiling and documenting information in a manner that enables independent validation;
- Documenting the explanation and/or justification (e.g. choice of procedures, methodologies, parameters, information sources, key factors, sampling criteria);
- Documenting the justification of selected criteria (e.g. for the determination of additionally);
- Documenting assumptions, references and methods such that another party can reproduce reported information;
- Documenting any external factors to the project that may affect the decisions of intended users.

g) **Verifiable**, with clear documentation behind it, so that the processes which produce the measure can be validated.

2. The assessment of quality of information will enable the Organization to produce clear, transparent and comprehensive program performance reports.

ANNEX 2: Criteria for individual performance measures

Criteria for individual performance measures:

A performance measure should be:

a) **Relevant** to what the organization is aiming to achieve;

b) **Attributable -** the activity measured must be capable of being influenced by actions which can be attributed to the organization, and it should be clear where accountability lies;

c) **Well-defined** - with a clear, unambiguous definition so that data will be collected consistently, and the measure is easy to understand and use;

d) **Timely**, producing data regularly enough to track progress and, quickly enough for the data to still be useful;

e) **Reliable** - accurate enough for its intended use, and responsive to change;

f) **Comparable** with either past periods or similar programs elsewhere; and

g) **Verifiable**, with clear documentation behind it, so that the processes which produce the measure can be validated.

ANNEX 2 TERMS OF REFERENCE

VALIDATION OF PROGRAM PERFORMANCE REPORT FOR 2008

Background

41. Each year, WIPO assesses its annual performance based on an approved performance framework. For this purpose, the Organization utilizes performance information that is complete and actual (rather than projected) for almost all of its performance measures.

42. Good quality information is crucial if performance measures are to be used effectively to improve service delivery and accountability. Such information should help the Organization to: a) manage delivery against priorities; b) report reliably on its achievements; and c) assess whether WIPO needs to revise policies and programs.

43. Within the Results-Based Management approach adopted, one of the Organization's aims is to provide reliable, timely and accurate information so that it can track performance over time, identify the need for any remedial action to achieve expected results and report clearly to its stakeholders on performance. Therefore a validation approach to all program performance reports has been introduced within the Organization in order to ensure that the information in these reports meets certain quality standards.

44. To facilitate the validation exercise, IAOD's Evalution Section has prepared the present approach paper which will guide the validation team and inform program staff during the process.

OBJECTIVE OF THE VALIDATION

45. The purpose of validation is to have a thorough independent assessment by a Third Party (validation team) of authenticity and reliability of the information on progress on program deliveries. The validation team will assess that the information used in individual program performance reports meet all the identified and applicable criteria. Validation will finally result in a conclusion by the validation team whether the quality of information used for reporting against defined results meets the defined criteria (see Annex 1)

46. This assessment by the validation team will, inter alia:

d) Assess the evidence which supports the contribution reported to specific key performance indicators, expected results and ultimately to the relevant strategic goal(s);

e) Ensure that the approved performance framework (document WO/PBC/13/4) is being applied;

f) Assess information on program delivery by applying the following criteria: relevance, timeliness, accuracy, consistency, continuity, transparency and verification. Further details on validation criteria are presented in the annex under "**Step one: Preparation of approach paper**".

Scope

47. For the validation exercise only one expected result by program will be selected for validation. Further details on the sampling process for expected results has been described on page 7, Step two of this approach paper. The sample size will represent 20 per cent of all expected results. The validation team will validate the authenticity and reliability of the information on program deliveries. The information required by the validation team will vary from program to program and will depend on what has been reported. Some examples of information that might be required include: monitoring data, statistical data, workshop papers, mission reports, reviews, studies, MoUs, records, etc.

48. Program managers shall keep records of the information that is being used for reporting performance within their programs.

49. For validation purposes program managers, upon request, will need to make available all relevant information used for reporting purposes as part of the program performance report. The validation team may request clarifications, including validation of specific documents.

50. Subject to considerations of confidentiality, original documents shall be shown to the validation team to facilitate the validation process. If required, copies of original documents might be requested.

51. It is to note that the validation work will only assess the quality of information underpinning the performance framework used to report to Member States. The validation team will not validate the quality of expected results and indicators, nor will the team provide conclusions about the accuracy of all expected results but rather validation of a representative sample of 20 per cent of all expected results.

Stakeholder involvement

52. Program managers as well as the Program Management and Performance Section will be consulted before and during the validation process. Final outcomes of the report will be available to all stakeholders of the Organization. The report will also be presented to the Audit Committee.

Background documentation

53. **THE FOLLOWING DOCUMENTS WILL BE AVAILABLE TO THE CONSULTANT:**

- Revised Program and Budget Document 2008/09
- Individual Program Performance Reports for 2008
- Summary Program Performance Report and Individual Program Performance Reports for 2006/07
- Validation approach paper

54. In addition to the above, program managers shall make available all necessary information the consultant might require in order to undertake the validation exercise.

Accountability and responsibilities

55. The consultant will liaise with and report to the Senior Evaluator who will guide him within the organization.

56. The Senior Evaluator is responsible for providing the consultant with all the necessary support and required documentation in order to undertake this assignment.

57. The consultant is required to undertake most part of his/her work from the WIPO Headquarter, in Geneva and is responsible to deliver in collaboration with the Senior Evaluator of WIPO the above mentioned objectives and tasks within the agreed period.

VALIDATION PROCESS

58. Overall the proposed validation process will consist of the six following steps (see attached Validation Approach Paper for further information):

- **Step one: Preparation of approach paper**
- **Step two: Sampling of expected results**
- **Step three: Publication of sample and information request**
- **Step four: Validation of sample**
- **Step five: Reporting**
- **Step six: Validation report dissemination and lessons learned**

Task of external expert

59. The external expert will undertake the following tasks jointly with the Senior Evaluator:

- Finalize validation approach;
- Agree on sampling approach;
- Undertake validation of information jointly, identify findings, conclusions and when necessary provide recommendations for improvement. The consultant and Senior Evaluator will:
    - review documentation used for reporting against expected results and indicators;
    - have follow up interviews with program staff;
    - compare information with programs that have similar or comparable characteristics;
    - review the correctness of critical formulas and calculations.
- Facilitate a smooth undertaking of the validation exercise;
- Prepare validation report in collaboration with Senior Evaluator.

DELIVERABLES

- Format to be used for validation report
- Draft and final validation Report to be done in collaboration with the Senior Evaluator.

Timing

60. It is envisaged that the validation exercise will start on 6th March 2009 and will be finalized on 30th April 2009.

Costs

- Up to 20 man days

DSA and travel allowances will be covered by WIPO and are not considered to be part of the consultant's fee.

50% of the sums approves as part of the ToRs shall be paid after completion of 10 working days.
50% of the sums approved as part of this ToRs shall be paid within 30 days after delivery of the final validation report.

Payments shall be made in Swiss francs.

Skills required

- Extensive evaluation and audit experience
- Extensive international evaluation experience
- Facilitation skills will be an advantage

- English proficiency level is mandatory

Management arrangements

Under the supervision of the Chief Evaluator and the Director of IAOD the exercise will be managed by Julia Flores (Senior Evaluator, IAOD).

IAOD will provide the consultant with temporary office space.

Prepared by:
Julia Flores
Senior Evaluator (IAOD)
WIPO

ANNEX 3 LIST OF MEETINGS FOR VALIDATION EXERCISE

Date	Participants	Designation	Program/Section
Preliminary Meetings			
16.03	Julia Flores	Senior Evaluator	Internal Audit and Oversight Division (IAOD)
	Nicholas Treen	Director	IAOD
	Chitra Narayanaswamy	Senior Adviser	Office of the Director General
	Joe Bradley	Head	Program Management and Performance Section
17.03	David Muls	Director	The PCT System (16)
	Bruno Le Feuvre	Assistant Statistical Officer	The PCT System (16
	Marco Pautasso	Acting-Director Advisor	TA and Capacity Building (6)
	David Chambers	Head	Language Service (28)
	John Tarpey	Director	Public Outreach and Communication (1)
	Hang Gao	Deputy Dean and Head	WIPO Worldwide Academy (11)
	Neila Krifi	Head	Travel and Procurement (30)
	Wend Wendland	Deputy Director and Head	Traditional Knowledge etc. (15)
18.03	Svein Arneberg	Acting Director	HRM Department (25)
	Therese Dayer	Assistant Director and Head	HRM and Social Security Section (25)
	Philippe Baechtold	Director	Patent Division (12)
19.03	Antonios Farassopoulos	Head	Global IP Infrastructure (19)
20.03	Neil Wilson	Director	IT Division (27)
	Gabor Karetka	Deputy Director	IT Program Support Service (27)
	Louise van Greunan Vuagnat	Deputy Director	Enforcement and Special Projects (10)
	Ignacio de Castro Llamas	Deputy Director and Head	Arbitration and Mediation Center (21)
	Francisco Rios	Legal Officer	Arbitration and Mediation Center (21)
	Nicholas Treen	Director	IAOD
	Georg Ladj	Chief Evaluation Officer	IAOD
30.03	Isabelle Boutillon	Director-Advisor	Premises Management (29)
Validation Meetings			
30.03	Wend Wendland	Deputy Director and Head	Traditional Knowledge etc. (15)
31.03	Marco Pautasso	Acting-Director Advisor	TA and Capacity Building (6)
	Nahla Haidar El Addal	Director-Advisor	TA and Capacity Building (6)
	Rowena Paguio	Senior Counsellor	TA and Capacity Building (6)
	Hang Gao	Deputy Dean and Head	WIPO Worldwide Academy (11)
01.04	Neila Krifi	Head	Travel and Procurement (30)
	Neil Wilson	Director	IT Division (27)
	Gabor Karetka	Deputy Director	IT Program Support Service (27)
02.04	Louise van Greunan Vuagnat	Deputy Director	Enforcement and Special Projects (10)
03.04	Svein Arneberg	Acting Director	HRM Department (25)
	Therese Dayer	Assistant Director and Head	HRM and Social Security Section (25)
	Michel Campi	Human Resources Officer	HRM Department (25)
	Antonios Farassopoulos	Head	Global IP Infrastructure (19)
06.04	Isabelle Boutillon	Director-Advisor	Premises Management (29)
	Gregoire Bisson	Head	International Registration Systems (18)
07.04	Erik Wilbers	Director	Arbitration and Mediation Center (21)
	Ignacio de Castro Llamas	Deputy Director and Head	Arbitration and Mediation Center (21)
	Francisco Rios	Legal Officer	Arbitration and Mediation Center (21)
	John Tarpey	Director	Public Outreach and Communication (1)
	Claudine Platret	Administrative Assistant	Communications Division (1)
	Philippe Baechtold	Director	Patent Division (12)
08.04	David Muls	Director	The PCT System (16)
	Bruno Le Feuvre	Assistant Statistical Officer	The PCT System (16)
	William Meredith	Head	Global IP Information Service (16)
	David Chambers	Head	Language Service (28)

ANNEX 4 REPORT FORM FOR VALIDATION EXERCISE

PROGRAM NUMBER: ___ INTERVIEWEE(S): ______________________

PART 1 ***Preliminary Questions***	***Comments***
1. Who developed the objective/ results/indicators?	
2. Do they still seem appropriate?	
3. Do you make use of this information on a regular basis?	
4. Does your line manager make use of this information, as far as you know?	
PART 2 ***Criteria***	
1. ***Data specification/ selection:*** • Relevant	
• Sufficient/ comprehensive	
2. ***Collection:*** • Efficient/easily accessible	
• Routinely acquired	
3. ***Analysis:*** • Carried out (how? by whom?)	
• Consistent/ comparable	
4. ***Reporting:*** • Accurate/verifiable	
• Timely	

• Clear/transparent	
PART 3 **Summary of data limitations** *(including any perverse indicators)*	
Lessons learned for future data quality	
Lessons learned for planning and monitoring *(e.g. re selection of results and indicators)*	
Other comments:	

Time taken: minutes

ANNEX 5 OVERVIEW OF VALIDATION EXERCISE FINDINGS

USED		*PID SELECTION*		*COLLECTION*		*ANALYSIS*		*REPORTING*			*NARRATIVE SUMMARY*
PROGRAM	SUPERVISORS	***RELEVANT***	***SUFFICIENT***	EFFICIENT	ROUTINE	CARRIED OUT	COMPARABLE	VERIFIABLE	TIMELY	CLEAR	
☹	☹	😐	😐	😐	☹	😐	☺	☺	😐	😐	Manager very receptive to ideas for improvement and keen on idea of a basic tool to capture PID systematically; recording of data neat and accessible
😐	☹	☹	☹	😐	😐	☹	☹	😐	😐	☹	PID is list of activities, without links to result or PIs (not selected by current Manager); list of activities reliable but irrelevant
☺	😐	😐	😐	☺	☺	☺	😐	😐	☺	😐	PID straightforward and easily collected; PID could be tighter and clearer to reflect outcomes not just activities; Manager is developing systems to capture PID; time constraints for meeting
😐	😐	😐	☺	😐	😐	😐	☺	😐	☺	😐	Appropriate process indicator which serves the purpose; reporting could be tighter and more transparent
☺	☹	😐	☺	☺	☺	☺	☺	☺	☺	☺	The indicator chosen is well supported by an explanation of the methodology behind it; for PID reporting could be confined to simple figures, without explanation which should be available on request
☺	😐	😐	☺	☺	😐	☺	😐	☺	😐	😐	This program has responsibility for a multitude of different events, making it difficult to find coherent, higher-level indicators; generally, reporting is appropriate; agreed a table of basic information required by indicators could be usefully designed and routinely completed
☺	☹	😐	☹	😐	☹	☺	☺	☺	😐	☹	Information is processed mostly manually, although electronic recording systems are being considered; there is a reluctance to disclose specific figures in PID, although these are available

USED		***PID SELECTION***		***COLLECTION***		***ANALYSIS***		***REPORTING***			***NARRATIVE SUMMARY***
PROGRAM	SUPERVISORS	***RELEVANT***	***SUFFICIENT***	EFFICIENT	ROUTINE	CARRIED OUT	COMPARABLE	VERIFIABLE	TIMELY	CLEAR	
☺	☹	😐	😐	😐	😐	😐	😐	☺	😐	😐	Difficulties acquiring data; PID could be more precise; program revealed strong commitment to utilizing and improving information
☹	😐	😐	☹	☺	😐	☺	😐	☺	😐	☹	Minimalist reporting without baselines; output oriented without explanation of value or relation to result
☺	😐	😐	😐	☺	☺	☺	☺	☺	☺	😐	The only significant limitation in relation to data quality is the imprecision of wording with indicators, and the different language used in reporting
☺	☺	😐	☹	☺	😐	😐	☺	☺	😐	😐	More specific data on baseline figures and more clearly disaggregated data would improve quality
☺	☺	☹	😐	😐	😐	😐	😐	☺	☺	☹	Over-elaborate and confusing reporting; indicator basic and at high level; result and indicator should be reviewed to reveal more value
☺	☺	😐	☺	☺	☺	☺	☺	☺	☺	😐	PIs are carefully worded and data collection is efficient; PID would be clearer if confined strictly to what information is required.
☺	☹	☺	😐	😐	☺	😐	☺	☺	☺	😐	Program operating well within organizational culture and expectations; well presented supporting information; PID would improve with baselines and tighter wording; second result reflects previous rather than current value
😐	☹	☹	☹	😐	😐	😐	😐	😐	😐	☹	Information about improvements is readily available and verifiable, and, if a better indicator had been selected, there would be fewer problems. As it is, the requirements of the indicator could not be met.
10/3/2	3/5/7	1/11/3	4/6/5	7/8/0	5/8/2	7/7/1	8/6/1	11/4/0	7/8/0	1/9/5	