INDEPENDENT REVIEW PROCESS

INTERNATIONAL CENTRE FOR DISPUTE RESOLUTION

BOOKING.COM B.V.,	)	ICDR CASE NO. 50 117 T 00247 14
	)	
Claimant,	)	
	)	
and	)	
	)	
INTERNET CORPORATION FOR ASSIGNED NAMES AND NUMBERS,	)	
	)	
	)	
Respondent.	)	
	)	

ICANN'S RESPONSE TO CLAIMANT BOOKING.COM'S REPLY BRIEF

Jeffrey A. LeVee
Kate Wallace
Jones Day
Contact Information
Redacted

Counsel to Respondent
The Internet Corporation
For Assigned Names and Numbers

TABLE OF CONTENTS

INTRODUCTION

1.   The Internet Corporation for Assigned Names and Numbers ("ICANN") hereby submits its Response to the Reply Brief ("Reply") submitted by claimant Booking.com B.V. ("Booking.com") on October 6, 2014.

2.   In its opening memorandum, ICANN established that the Independent Review Process ("IRP") set forth in ICANN's Bylaws focuses solely on the actions of the ICANN Board and is not available as a mechanism to challenge the conduct of ICANN staff or third parties (such as the third-party evaluators in the context of the New gTLD Program). In its Reply, Booking.com concedes that the IRP Panel's role is limited to declaring "whether the contested actions of the <u>ICANN Board</u> are consistent with applicable rules."[1] Nonetheless, the majority of the conduct that Booking.com complains of in its Reply challenges the conduct of ICANN staff or the third party evaluator retained by ICANN to perform the String Similarity Review for each gTLD application submitted within the New gTLD Program.

3.   For example, Booking.com challenges the methodology adopted by the String Similarity Panel ("SSP") in conducting the review, the SSP's decision to maintain the anonymity of its evaluators, and the decision not to publish a detailed report for each of the 1,930 applications reviewed. Because the ICANN Board was not involved in <u>any</u> of these decisions, and because the IRP is not a mechanism to challenge the decisions of ICANN staff or third parties, Booking.com's IRP Request must be denied.

4.   The only even arguable Board action at issue in Booking.com's Reply is the Board's decision to delegate the String Similarity Review to an independent third party evaluator. Booking.com essentially proposes that the ICANN Board itself should have conducted the String

---

[1] Booking.com Reply Brief, ¶ 3 (emphasis added).

1

Similarity Review of each of the 1,930 gTLD applications. However, the Board and the Internet community, in adopting the Guidebook, made clear that the Board would <u>not</u> undertake such a responsibility, and ICANN's Bylaws and Articles contain no mandate to do so.

5.   In short, Booking.com's IRP Request should be summarily denied. There simply is no Board action that this Panel has been asked to review, and no basis for the Panel to determine that the Board violated its Bylaws or Articles by making the decision to have a panel of experts – rather than the Board itself – conduct the String Similarity Review for each of the new gTLD applications.

<center>ARGUMENT</center>

A.   ICANN Appropriately Delegated The String Similarity Review To A Third Party Evaluator.

6.   Booking.com claims that ICANN violated its Bylaws or Articles by "outsourc[ing] the String Similarity Review, and the quality control of that review, to third parties" and by failing to "supervise or investigate the manner in which the SSP made its determinations or exercise its own independent judgment in relation to those determinations."[2] The thrust of Booking.com's assertion is that ICANN should have been prohibited from delegating the String Similarity Review to an independent evaluator, and that the ICANN Board instead should have substantively reviewed each and every application for string similarity.

7.   Within the New gTLD Program, <u>every</u> applied-for string (or gTLD) was subjected to the String Similarity Review set out in Section 2.2.1.1 of the Guidebook. Thus, under Booking.com's proposal, the ICANN Board – and the ICANN Board alone – was obligated to perform the String Similarity Review for the more than 1,900 new gTLD applications submitted. Booking.com's position is untenable and is not supported by ICANN's Bylaws or Articles.

---

[2] *Id.* at ¶¶ 9-10.

<center>2</center>

8.   Notably, string similarity was only one aspect of the evaluation process. The Guidebook defines six evaluations that each application has been required to go through, including string similarity review, DNS stability review, geographic names review, registry services review, technical and operational capability review, and financial capability review.[3] Each of these evaluations was performed by independent experts specifically engaged by ICANN staff to complete these extensive evaluations. Following Booking.com's logic, the fifteen members of the ICANN Board, not the expert panels (which consist of hundreds of individuals), should have performed the thousands of evaluations necessary to complete the new gTLD application process.

9.   Booking.com's preferred approach is illogical, unworkable and would bring ICANN's operations to a screeching halt. Indeed, "[a]ctive involvement by the board in day-to-day affairs of the corporation does not accord with the realities of contemporary business practices, other than perhaps in a relatively closely held corporation. The role of the board in this context is the formulation of major management policies rather than direct involvement in day-to-day management."[4] The fact that the ICANN Board may have the "power" to review applications for string similarity does not mean that the Board should have exercised that power, particularly at the initial evaluation stage. This is particularly appropriate with respect to the New gTLD Program, which, as noted above, involved over 1,900 applications and literally thousands of evaluation decisions that had to be performed with respect to those applications.

10.   There simply is no requirement – under ICANN's governing documents or imposed by law – that would mandate that the ICANN Board inject itself into the day-to-day

---

[3] http://newgtlds.icann.org/en/program-status/evaluation-panels; *see also,* Cl. Ex. RM-5 (Guidebook), at Module 2. All applicants also underwent background screening, which was conducted in two areas: (i) general business diligence and criminal history; and (ii) history of cybersquatting behavior. *Id*.
[4] *See* Cal. Corp. Code § 300, Legislative Comm. Comments.

affairs of the evaluation process in the manner Booking.com proposes.[5] Indeed, a corporate

board must be in the "position of being an effective overseer of corporate operations – not to

micro-manage such operations but to exercise a more in-depth, informed, and objective oversight

role."[6] Consistent with these well-settled legal principles, neither ICANN's Bylaws, nor the

Articles, nor the Guidebook requires the ICANN Board to conduct <u>any</u> analysis of the decisions

of third party experts retained to evaluate string similarity.

11.   Thus, Booking.com is wrong to argue that ICANN's decision to delegate String

Similarity Review to an independent third party was an abdication of the Board's obligation to

reserve "ultimate responsibility for the New gTLD Program,"[7] and its "right to individually

consider an application for a new gTLD …."[8] The Board's conduct was consistent with

applicable legal standards and the relevant governing documents. Simply because the ICANN

Board has the discretion to consider individual applications does not mean it is required to do so

or that it should do so, particularly at an initial evaluation stage.

12.   In any event, the SSP's determination was well supported. Specifically, the SSP

developed its own process, quality control mechanisms, and considerations surrounding non-

exact contention sets for string similarity evaluation.[9] Out of more than 1,900 applications, only

<u>two</u> non-exact match contention sets were found: (1) .hotels/.hoteis – which is at issue here; and

---

[5] BoardSource, The Source: Twelve Principles of Governance That Power Exceptional Boards (2005) ("Nonprofit boards have primary responsibility for governance – the exercise and assignment of power and authority – of their organizations. Boards reserve to themselves organizational oversight and policy setting, and delegate to the chief executive responsibility for managing operations and resources."); Section of Business Law, American Bar Association, Guidebook for Directors of Nonprofit Corporations 24 (George W. Overton, ed.) (1993) ("The board of directors, as such, does not operate the day to day business of the corporation. In delegating that function to others, it must set policies and oversee the corporate agents.")

[6] American Bar Association, Guide to Nonprofit Corporate Governance in the Wake of Sarbanes-Oxley 17, 19 (2005).

[7] Cl. Ex. RM-5 (Guidebook), at § 5.1.

[8] *Id*.

[9] Cl. Ex. Annex-11, at p. 3 (18 December 2013 ICC Memorandum Re: String Similarity Process, Quality Control and Non-Exact Contention Sets).

(2) .unicorn/.unicom.[10] The SSP explained that a string pair was found to be visually confusingly similar when the following features were present:

- Strings of similar visual length on the page;

- Strings within +/- 1 character of each other;

- Strings where the majority of characters are the same and in the same position in each string; and

- The two strings possess letter combinations that visually appear similar to other letters in the same position in each string

    o For example rn~m & l~i[11]

13.   Importantly (and undeniably), .hotels and .hoteis meet <u>every</u> <u>one</u> of these criteria.

- .hotels and .hoteis are "strings of similar visual length on the page";

- .hotels and .hoteis are "strings within +/- 1 character of each other";

- .hotels and .hoteis are "strings where the majority of characters are the same and in the same position in each string"; and

- .hotels and .hoteis "possess letter combinations that visually appear similar to other letters in the same position in each string", namely "l" & "i".

14.   Further, the Guidebook provides that the SSP would be informed by an algorithmic score for the visual similarity between each applied-for string and each of the other existing and applied-for TLDs and reserved names:[12]

> The score will provide one objective measure for consideration by the panel, as part of the process of identifying strings likely to result in user confusion. In general, applicants should expect that a

---

[10] https://www.icann.org/news/announcement-2013-02-26-en.
[11] *Id.*
[12] Cl. Ex. RM-5 (Guidebook), at § 2.2.1.1.2 (Review Methodology).

higher visual similarity score suggests a higher probability that the application will not pass the String Similarity review.[13]

In its Reply, Booking.com ignores the fact that .hotels and .hoteis scored a 99% for visual similarity using this publicly available SWORD algorithm.[14] As the Guidebook makes clear, "a higher visual similarity score suggests a higher probability that the application will not pass the String Similarity review."[15]

15.   Given that the SSP's determination that .hotels and .hoteis are visually confusingly similar was factually and objectively well supported, the fact that the Board did not independently and substantively review the merits of that determination is not evidence that the Board in some way abdicated its supervisory or monitoring role in violation of its Bylaws or Articles.

B.   Booking.com's Invocation of "Due Process" Does Not Support Its IRP Request.

16.   Booking.com argues – without a single citation – that "the ICANN Board was obliged to ensure due process" in its administration of the New gTLD Program.[16] Booking.com repeats the need for "due process rights" frequently and somewhat theatrically in its brief, but Booking.com never explains the source of its "due process" rights, much less what those rights would consist of.

17.   California non-profit public benefit corporations, such as ICANN, are expressly authorized to adopt and amend the corporation's bylaws, and to define the scope and form of the organization.[17] Pursuant to this explicit authority, ICANN's Board approved the New gTLD

---

[13] *Id.*
[14] *See* https://icann.sword-group.com/algorithm/.
[15] Cl. Ex. RM-5 (Guidebook), at § 2.2.1.1.2.
[16] Booking.com Reply Brief, ¶¶ 15, 17.
[17] Cal. Corp. Code § 5150(a) (authorizing the board of a nonprofit public benefit corporation to adopt and amend the corporation's bylaws).

Program in June 2011. ICANN established the New gTLD Program as a means to implement

policy recommendations made by the Generic Names Supporting Organization ("GNSO"),

which is ICANN's policy-development body responsible for developing and recommending to

the ICANN Board substantive policies relating to generic top-level domains. The procedures

governing the New gTLD Program are set forth in the New gTLD Applicant Guidebook

("Guidebook"), an extensive (and exhaustively vetted) document that provides details to new

gTLD applicants and forms the basis for ICANN's evaluation of new gTLD applications.[18]

18.　　　There is no provision in ICANN's Bylaws, Articles, or Guidebook that

specifically affords any gTLD applicant a right to procedural "due process" similar to that which

is afforded in courts of law. Nevertheless, because ICANN does conduct its activities in the

public interest, ICANN does provide more opportunity for parties to be heard and to dispute

actions taken than most private entities, including the extensive accountability mechanisms that

are set forth in the Bylaws (such as the right to file a reconsideration request and the right to file

an Independent Review proceeding such as this one). There is no "due process" right to have

some other form of remedy available when an applicant disagrees with the determination of an

evaluation panel and, rights of "due process" are not normally available in a corporate setting.

The due process clause in the U.S. Constitution "protects individual rights only from <u>government</u>

---

[18] As set forth in ICANN's Response to Booking.com's IRP Request (*see* ¶¶ 33-34), ICANN's decision to proceed with the New gTLD Program followed many years of discussion, debate and deliberation within the ICANN community, including participation from end users, civil society, technical experts, business groups, governments and others. Meaningful input from the community, including prospective gTLD applicants, formed the basis for the New gTLD Program, and ICANN adhered to the substantive and procedural rules derived throughout that process in processing Booking.com's application for .hotels. Booking.com claims that "ICANN would effectively have *carte blanche* to violate its obligations under its Bylaws and Articles of Incorporation, provided it engaged in consultation beforehand." *See* Booking.com Reply Brief, ¶ 16. Booking.com misses the point. ICANN adhered to the policies and procedures articulated in its Bylaws, Articles of Incorporation, and the Guidebook, the latter of which was adopted only after being publicly vetted with ICANN's stakeholders and the broader Internet community. ICANN's adherence to these governing documents cannot provide a basis for an IRP under ICANN's Bylaws.

action, not from <u>private</u> action."[19] ICANN's approval and administration of the New gTLD

Program is a private action, not that of a state actor.[20]

20.   Second, Booking.com has no legal right to attain any particular gTLD. The

Guidebook, along with the application that Booking.com signed, makes clear that the submission

of an application does not constitute a "right" to operate the applied-for TLD, and ICANN has no

obligation to delegate <u>any</u> TLD.[21] Because Booking.com has not been deprived of a legally

protected interest, its invocation of alleged procedural "due process" rights fails as a matter of

substantive law[22] and also fails to identify conduct of ICANN's Board that violated either the

Bylaws or the Articles.

C.   There Was No Board Action In The Selection Of The String Similarity Panel, And Booking.com's Claims Are Factually Incorrect.

20.   Within the New gTLD Program, every applied-for string (or gTLD) has been

subjected to the String Similarity Review set out in Section 2.2.1.1 of the Guidebook. As set

forth in the Guidebook, "[t]his similarity review [was] conducted by an independent String

Similarity Panel," not by ICANN.

21.   Booking.com claims that "[t]he ICANN Board made several errors in the

resulting SSP selection process."[23] Booking.com asserts that: (a) "ICANN failed to make clear

how it would evaluate candidate responses or how it ultimately did so"; (b) "the identities of the

unsuccessful candidates (if any) to perform the String Similarity Review remain unknown"; and

(c) ICANN "has never demonstrated that [certain] required information was provided by the SSP

---

[19] *Single Moms, Inc. v. Mont. Power Co.*, 331 F.3d 743, 746 (9th Cir. 2003).
[20] While ICANN has historically had certain contracts with the U.S. government, ICANN definitely is not a state actor.
[21] gTLD Application Terms and Conditions, ¶ 3, *available at* http://newgtlds.icann.org/en/applicants/agb/terms.
[22] *Board of Regents of State Colleges v. Roth*, 408 U.S. 564, 569-570 (1972) ("The requirements of procedural due process apply only to the deprivation of interests encompassed by the Fourteenth Amendment's protection of liberty and property.").
[23] *See* Booking.com Reply Brief, ¶ 18.

selected by the ICANN Board."[24] However, contrary to Booking.com's assertions, the Board

was not involved in the selection of the String Similarity Panel. ICANN staff selected third-

party providers, including the SSP, via a solicitation for Expressions of Interest.[25] Accordingly,

the selection of the SSP is not a basis for Independent Review under the ICANN Bylaws.

22.   In challenging ICANN's selection of the SSP, Booking.com is challenging an

ICANN staff action. Following a period of review of the Expressions of Interest against the

selection criteria, ICANN staff selected the evaluation panels for the independent evaluations

defined in the Guidebook.[26] The selections were not made by the ICANN Board. The selected

evaluators were then announced by ICANN staff in October 2011 during a New gTLD Program

Update presentation at an ICANN meeting held in Dakar, Senegal.[27]

23.   The sole purpose of an IRP is to focus on the actions of the ICANN Board; it is

not available as a mechanism to challenge the conduct of ICANN staff or third parties such as

the third-party evaluators in the context of the New gTLD Program. Because the selection of the

SSP was the decision of ICANN staff – and not the ICANN Board – Booking.com has failed to

challenge any Board conduct appropriately subject to an IRP.

24.   Moreover, the selection of the SSP was announced on 26 October 2011.[28] If

Booking.com had concerns related to this selection, Booking.com should have pursued them at

that time, not four years later.[29]

---

[24] *See id.* at ¶¶ 20, 21.
[25] https://www.icann.org/news/announcement-2009-02-25-en.
[26] http://newgtlds.icann.org/en/program-status/evaluation-panels.
[27] New gTLD Program Update, 26 October 2011, *available at* http://newgtlds.icann.org/en/program-status/evaluation-panels (stating that ICANN staff was "[c]urrently negotiating Evaluation Panel services contracts with final candidates" and that InterConnect Communications had been selected to serve as the String Similarity Panel); http://dakar42.icann.org/node/26953 (New gTLD Program Update session led by ICANN's then Senior Vice President, Stakeholder Relations).
[28] New gTLD Program Update, 26 October 2011, *available at* http://newgtlds.icann.org/en/program-status/evaluation-panels.

9

25.   Booking.com's claims are also factually incorrect. Booking.com alleges that "ICANN failed to make clear how it would evaluate candidate responses or how it ultimately did so."[30] Not so. On 25 February 2009, ICANN publicly posted its Call for Expressions of Interest for New gTLD String Similarity Examiners, which detailed the criteria against which applicants would be evaluated.[31] On 25 October 2009, ICANN publicly posted the "General Criteria Matrix," which outlined the basis for ICANN's evaluation, including questions relating to an applicant's proposed: (i) approach and scalability; (ii) process to ensure consistency of evaluation results; (iii) skill-set and experience; and (iv) cost and sustainability.[32] ICANN published an entire webpage on the evaluation selection process, which provided detailed information on the phases of ICANN's review and selection process for independent evaluators.[33]

26.   Booking.com's claim that "the identities of the unsuccessful candidates (if any) to perform the String Similarity Review remain unknown" also is false.[34] On the ICANN webpage dedicated to the Evaluation Panels Selection Process, ICANN posted a document entitled "List of Respondents" on 25 October 2009.[35] There, ICANN identified each applicant that submitted a response to ICANN's Call for Expressions of Interest for independent evaluators.

27.   Booking.com's assertion that ICANN should have "demonstrated" that the selected SSP submitted the requisite application materials likewise is baseless. Booking.com

---

(continued…)

[29] *See* http://newgtlds.icann.org/en/about/evaluation-panels-selection-process; *see also,* Cl. Ex. RM-2 (ICANN's Bylaws), at Art. IV, § 3.3 (requiring that IRP Requests be filed within 30 days of the publication of materials demonstrating the alleged violation).
[30] Booking.com Reply Brief, ¶ 20.
[31] *See* https://archive.icann.org/en/topics/new-gtlds/eoi-string-sim-25feb09-en.pdf.
[32] *See* https://www.icann.org/news/announcement-2009-10-25-en.
[33] http://newgtlds.icann.org/en/about/evaluation-panels-selection-process.
[34] Booking.com Reply Brief, ¶ 20.
[35] http://newgtlds.icann.org/en/about/evaluation-panels-selection-process.

does not identify any policy requiring ICANN to publicly post an independent evaluator's

application materials. And no Guidebook or Bylaws provision requires that ICANN do so either.

In essence, Booking.com argues that, because the Guidebook does not include Booking.com's

"wish list" of procedural requirements, independent review is warranted. Independent review,

however, is only warranted where the Board has acted in a manner inconsistent with ICANN's

Bylaws or Articles.[36]

D.   There Was No Board Action In The String Similarity Panel's Decision Not To Publish The String Similarity Panel's Methodology In Advance.

28.   Booking.com claims that "ICANN is unable to show that there was a pre-

established methodology for performing the String Similarity Review."[37] But the SSP – not

ICANN – was responsible for the development of its own process documents and methodology

for performing the String Similarity Review consistent with the provisions of the Guidebook.

The SSP also was responsible for the maintenance of its work papers.[38] Thus, again,

Booking.com challenges the conduct of a third party evaluator and not the actions of the Board.

29.   Moreover, Section 2.2.1.1 of the Guidebook clearly sets forth the standards to be

used in the String Similarity Review process, and the SSP has confirmed that the standard it used

for string similarity evaluation "comes from the AGB [Guidebook]."[39] Nowhere in the

Guidebook is there a requirement that the SSP's methodology for performing String Similarity

Review be published before the SSP commenced its evaluation process.

30.   Booking.com acknowledges that the SSP (not ICANN) published the process

documentation setting forth its methodology for determining whether applied-for strings are

---

[36] Cl. Ex. RM-2 (ICANN's Bylaws), at Art. IV, § 3.4.
[37] Booking.com Reply Brief, ¶ 24.
[38] Cl. Ex. RM-5 (Guidebook), at § 2.2.1.
[39] Cl. Ex. Annex-11.

visually confusingly similar under applicable Guidebook provisions.[40] Thus, Booking.com's

complaints that the factors outlined by the SSP are "arbitrary" and "baseless"[41] are complaints

against the SSP, not against ICANN and certainly not against the ICANN Board. IRPs do not

exist to address the propriety of action or inaction of third parties.

E.   **There Was No Board Action In The String Similarity Panel's Decision To Maintain The Anonymity Of The String Similarity Review Evaluators.**

31.   Booking.com claims that the String Similarity Review process was "unfair" and

"non-transparent" because the evaluators who performed the String Similarity Review remain

anonymous.[42] This claim fails for the same reasons articulated above. Specifically,

Booking.com is challenging the SSP's conduct in performing the String Similarity Review,

namely its decision not to disclose the identity of its evaluators. Booking.com's IRP Request is

therefore misplaced because it challenges an action of an independent, third-party expert

evaluator rather than an action of the ICANN Board.

32.   Further, Booking.com fails to identify any provision in ICANN's Bylaws,

Articles, or Guidebook that would require the disclosure of the panelists' identities.

Booking.com's belief that the String Similarity Review process should have included certain

requirements (such as a requirement to disclose the identity of the panelists) does not constitute a

Bylaws or Articles violation.

F.   **Booking.com's Claim That ICANN Did Not Publish A Report Detailing The String Similarity Panel's Determination Does Not Support An IRP Request.**

33.   Booking.com claims that ICANN "failed to produce the required independent

report or the reasoning for the [SSP's] determination" that .hotels and .hoteis were visually

---

[40] Booking.com Reply Brief, ¶¶ 24-25; *see also*, Cl. Ex. Annex-11.
[41] Booking.com Reply Brief, ¶ 25.
[42] *Id*. at ¶¶ 26-27.

confusingly similar.[43] Yet Booking.com concedes that "there is no explicit requirement in the

Applicant Guidebook" or elsewhere that ICANN publish the rationale of the SSP.[44] This

concession is dispositive. The IRP Panel's charge is limited to "comparing contested actions of

the [ICANN] Board to the [ICANN] Articles of Incorporation and Bylaws, and […] declaring

whether the Board has acted consistently with the provisions of those Articles of Incorporation

and Bylaws."[45] Absent an established policy or process that requires ICANN to act in the

manner preferred by Booking.com – and Booking.com has conceded that there is none –

independent review is not warranted.

34.   In addition, once again, Booking.com's claims are factually incorrect. The SSP

did publish documentation setting forth the factors considered by the SSP when determining

whether applied-for strings are visually confusingly similar under applicable Guidebook

provisions. Specifically, the SSP's process documentation "provide[d] a summary of the process,

quality control mechanisms and some considerations surrounding non-exact contention sets for

the string similarity evaluation as requested by ICANN."[46] ICANN subsequently published that

information on its website.[47] Notably, .hotels and .hoteis satisfy each of the factors that the SSP

found to create visual confusing similarity.[48]

> **G.   Booking.com's Claim That ICANN Breached Its Bylaws By "Blindly
> Accepting" The String Similarity Panel's Determination Is Incorrect
> And Again Involves No Board Action.**

35.   Booking.com argues that ICANN breached its Bylaws by "accepting," "without

providing effective supervision or quality control," the SSP's determination that .hotels

---

[43] *Id.* at ¶¶ 28-29.
[44] *Id.* at ¶ 28.
[45] Cl. Ex. RM-2 (ICANN's Bylaws), at Art. IV, § 3.4.
[46] Cl. Ex. Annex-11.
[47] http://www.icann.org/en/news/correspondence/mcfadden-to-weinstein-18dec13-en.
[48] Cl. Ex. Annex-11.

and .hoteis are visually confusingly similar.[49] However, again, there is no provision in the Bylaws or Articles that requires the ICANN Board to supervise or perform a quality control review of expert panel determinations. To the contrary, the Guidebook states that the string "similarity review will be conducted by an independent String Similarity Panel," not ICANN.[50] And "[i]n performing this review, the String Similarity Panel [not ICANN] will create contention sets that may be used in later stages of evaluation."[51]

36.   In the provisions governing the outcome of the SSP's evaluation, the Guidebook does not provide ICANN discretion to review or otherwise substantively consider the SSP's determination. Instead, the Guidebook provides that following the conclusion of the Panel's evaluation, "[a]n application for a string that is found too similar to another applied-for gTLD string will be placed in a contention set."[52]

37.   The Guidebook could not be more explicit: once the SSP determines two strings to be visually confusingly similar, the only possible outcome is for the two strings to be placed into a contention set. This is not the result of any ICANN Board action, but a straightforward application of the Guidebook provisions applicable to the SSP's determination.

38.   Accordingly, Booking.com's claim that the ICANN Board violated its Bylaws by "accepting" the SSP's determination on .hotels/.hoteis without a substantive review is wrong. The ICANN Board took no action – and was not required to take any action – with respect to evaluating or accepting the SSP's determination because the Guidebook does not call for the Board to take any such action. Nor is any such action required by any Bylaws or Articles provision. Accordingly, it cannot be a violation of ICANN's Bylaws or Articles for the Board

---

[49] Booking.com Reply Brief, ¶ 30.
[50] Cl. Ex. RM-5 (Guidebook), at § 2.2.1.1.
[51] *Id*. at § 2.2.1.1.1.
[52] *Id*. at § 2.2.1.1.3 (emphasis added).

not to conduct a "quality control" review of the SSP's determination. Because there is no Board action or decision associated with the SSP's determination, there is nothing for the IRP Panel to review.[53]

39.   Booking.com's assertion that string similarity determinations were not subjected to any quality assurance mechanism also is baseless.[54] ICANN retained JAS Advisors "to perform the quality assurance function" in October 2011, explaining that JAS Advisors' role was to ensure "accountability of the panelists and of the firms managing those panelists, continual improvement and some transparency into how the project operates."[55] Work performed by the SSP was subject to a procedural inspection on a statistically relevant randomly selected sample of applications.[56]

40.   Booking.com states that it "doubts that [it] is true" that JAS Advisors was retained to perform the quality review of the SSP's determinations. Booking.com cites only select passages from ICANN's New gTLD Program Update, which identified JAS Advisors as the provider selected to perform the "technical, operational and functional evaluations."[57]

---

[53] Booking.com claims that ICANN made a "commitment to review evaluations" of the SSP. Booking.com Reply Brief, ¶ 36. In support, Booking.com cites ICANN's representation that, following the SSP's determinations, ICANN would "review those results and put it in publishable form." *Id*. (citing RM-17, p. 24.) This, however, is a "processing" function that may require ICANN to briefly review (*i.e.* read) the evaluations in order to prepare the results for publication. *See* RM-17, p. 24. Moreover, any "review" required to prepare the reports for publication was undertaken by ICANN staff, not the ICANN Board. The point is that the ICANN Board did not perform any review with respect to the SSP's determination because the Guidebook does not call for the Board to take any such action.
[54] Booking.com Reply Brief, ¶ 31.
[55] New gTLD Program Update, 26 October 2011, *available at* http://newgtlds.icann.org/en/program-status/evaluation-panels; *see also,* Cl. Ex. RM-27 (Transcript_Dakar New gTLD Program Update), pgs. 14-15, *also available at* http://ccnso.icann.org/de/node/28083.
[56] JAS Global Advisors, gTLD Application Processing: Initial Evaluation Quality Program Report, at p. 12, attached hereto as Resp. Ex. 4.
[57] Booking.com Reply Brief, ¶¶ 31, 33.

15

Booking.com neglects to quote the very next sentence in the Update, which states: "And then underlying this all, JAS has been obtained to perform the quality assurance function."[58]

41.   Booking.com also complains that it is "not aware that any selection process was put in place in relation to the appointment of JAS Advisors to perform the String Similarity Review quality control."[59] Booking.com fails to recognize, as is clear on ICANN's website, that JAS Advisors submitted a response to ICANN's call for Expressions of Interest and was evaluated against the criteria set forth in that document, as well as the criteria published on ICANN's website dedicated to the evaluator selection process.[60] Moreover, just like all of the evaluation panels, JAS Advisors was selected by ICANN staff, and thus the process by which it was selected is not subject to independent review.

42.   Booking.com also challenges the propriety of the quality control review process established and implemented by JAS Advisors.[61] Once again, Booking.com's substantive disagreement with the methodology established by a third party evaluation panel is not properly the subject of an IRP.

H.   ICANN's Bylaws And Articles Of Incorporation Do Not Require That ICANN Provide For Review Mechanisms Beyond Those Already Provided.

43.   Booking.com argues that "[t]he ICANN Board should have corrected the mistakes of the String Similarity Review process on its own motion."[62] In essence, Booking.com asserts

---

[58] *Id*. at ¶ 33 (citing RM-27, p. 14); *see also, id.* at p. 13.
[59] *Id*. at ¶ 32.
[60] https://www.icann.org/news/announcement-2009-10-25-en (identifying selection criteria and also listing JAS Advisors as a respondent to ICANN's call for Expressions of Interest).
[61] Booking.com Reply Brief, ¶ 34 (stating that, in Booking.com's view, "[a] quality check on a random sample of applications cannot provide effective quality control of the String Similarity Review").
[62] *Id*. at ¶ 39. Booking.com also claims that "[t]he ICANN Board clearly knew, or at least should have known, that the process and the implementation of [the string similarity review] process raised serious concerns." *Id*. at ¶ 42. First, Booking.com asserts that "several Board members expressed concerns or abstained during the consideration of Booking.com's [Reconsideration] Request." *Id*. However, as explained in ICANN's opening memorandum, these Board members specifically "agreed that the process was followed," that "the BGC [Board Governance Committee]

that the fact that the Guidebook provides no mechanism for the substantive review of the SSP's

determinations is itself a violation of ICANN's Bylaws and Articles, or that the Board should

have realized that it had some form of "obligation" to conduct an "appellate review" of the String

Similarity Review process.[63] However, neither the Bylaws nor the Articles require ICANN to

reconsider the outcome of the Guidebook development process, a process that was started many

years ago and completed over two years ago (and one for which the time for challenge has long

since past).

44.   In any event, the absence of a review mechanism by which the ICANN Board

would be compelled to review each determination rendered by the SSP does not form the basis

for an IRP because there is nothing in ICANN's Bylaws or Articles requiring ICANN to provide

such a review mechanism. The decision not to have a formal review process was made

following years of consideration and community involvement and is not a violation of any "law"

(or Bylaw or Article).[64]

(continued…)

has done an appropriate job of applying a limited review standard to the application for reconsideration," and "that
the BGC did the right thing" under ICANN's Bylaws governing reconsideration Requests. ICANN's Opening
Memorandum, at ¶ 44 (citing Cl. Ex. Annex-16 at p. 3). Second, Booking.com notes that "ICANN reserved almost
one third of the Application fees (almost 118M USD) for risk costs, including legal defense costs." Booking.com
Reply Brief ¶ 42. This is hardly a concession that the string similarity review process was flawed in some way.
Indeed, the decision on how the application fees would be allocated was made long before any application was
submitted or the String Similarity Review commenced.
[63] Booking.com Reply Brief, ¶¶ 39-42.
[64] ICANN has recently authorized a limited and targeted review of two alleged inconsistent expert determinations
rendered in the context Module 3 of the Applicant Guidebook, which governs formal objections, including string
confusion objections that third parties may lodge against a gTLD applicant. The basis for that decision, however,
was that the applied-for gTLDs in each of the two relevant objection proceedings were the same applied-for gTLD
at issue in other objection proceedings that resulted in opposite conclusions by the respective expert panels. ICANN
has approved this limited further review only in two instances out of more than 270 objection proceedings. *See* 12
Oct 2014 NGPC Resolutions, 2(b), *available at* https://www.icann.org/resources/board-material/resolutions-new-
gtld-2014-10-12-en#2.b. Further, the Board will not conduct that review; an independent dispute resolution provider
is tasked with establishing a three-member expert panel to do so. In any event, that this limited review was granted
in the context of string confusion objections renders it inapplicable here, as string confusion objections are very
different from the String Similarity Review challenged by Booking.com. Specifically, each application within the
New gTLD Program was subject to String Similarity Review, as compared with a string confusion objection, which

17

I.   Booking.com's Request For Affirmative Relief Goes Beyond This IRP Panel's Authority.

45.   Booking.com requests that, in addition to various declarations that ICANN's conduct was inconsistent with ICANN's Bylaws and Articles, the Panel issue a declaration "[r]equiring that ICANN reject the determination that .hotels and .hoteis are confusingly similar and disregard the resulting contention set."[65]   In other words, Booking.com requests that "ICANN should therefore delegate both .hotels and .hoteis."[66]

46.   Booking.com's request for affirmative relief goes well beyond the IRP Panel's authority. An IRP Panel is limited to issuing its opinion as to "whether an action or inaction of the Board was inconsistent with the Articles of Incorporation or Bylaws" and "recommend[ing]" that the Board stay any action or decision or take any interim action until such time as the Board reviews and acts upon the opinion of the IRP Panel.[67] This IRP Panel does not have the authority to grant affirmative relief or to require ICANN to undertake specific conduct.

47.   In its Procedural Order No. 1, this IRP Panel asked Booking.com to address in its Reply "the nature of the relief sought by Claimant."[68] In its Reply, Booking.com does not refute the narrow scope of this IRP Panel's remit. Instead, Booking.com simply reiterates its claim for relief, asserting that "[a]ny other outcome would result in Booking.com continuing to suffer discrimination."[69]

(continued…)

may be lodged by a third-party against an applicant. Further, String Similarity Review is limited to assessing whether applicants are visually confusingly similar, whereas string confusion objections are adjudicated in light of the arguments and evidence submitted by the parties.

[65] Booking.com IRP Request, ¶ 78. Booking.com also requests that "ICANN be required to overturn the string similarity determination in relation to hotels and .hoteis and allow Booking.com's application to proceed on its own merits without reference to the application for .hoteis." *Id*. at ¶ 9.

[66] Booking.com Reply Brief, ¶ 45.

[67] Cl. Ex. RM-2 (ICANN's Bylaws), at Art. IV, § 3.4 and § 3.11(c)-(d).

[68] Procedural Order No. 1, 22 August 2014.

[69] Booking.com Reply Brief, ¶ 45.

48. In so doing, Booking.com urges the IRP Panel to act beyond its authority as established in ICANN's Bylaws. There is no way to harmonize Booking.com's position with the language of the Bylaws because the Bylaws make clear that this Panel does not have the discretion to modify or ignore the Bylaws.

49. Booking.com also is wrong as a factual matter. The SSP determined that the strings .hotels and .hoteis were confusingly similar, such that only one of those strings will be permitted to proceed to delegation. Thus, Booking.com's application for .hotels has not been denied, and it could very well be the applicant that is ultimately awarded a registry agreement.[70]
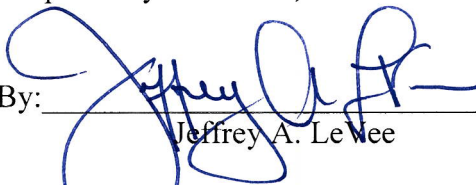
CONCLUSION

50. ICANN's conduct with respect to Booking.com's application for .hotels was fully consistent with ICANN's Articles of Incorporation and Bylaws. The fact that Booking.com disagrees with the SSP's determination to put .hotels and .hoteis in a contention set does not give rise to an IRP. Booking.com's IRP Request should be denied.

Respectfully submitted,

Dated: November 20, 2014   By: _____

Jeffrey A. LeVee

Jones Day
Counsel for Respondent ICANN

LAI-383225448

---

[70] Booking.com also incorrectly views ICANN's Bylaws as granting the IRP Panel authority to issue a declaration that would be "binding" on ICANN. Booking.com Reply Brief, ¶ 11. Booking.com is wrong here too, as IRP Panel declarations are not binding on ICANN. The plain language of the IRP provisions set forth in Article IV, section 3 of ICANN's Bylaws, as well as the drafting history of the development of the IRP provisions, make clear that IRP Panel declarations are not binding on ICANN. There is no ambiguity on this issue. In fact, the first ever IRP Panel constituted under ICANN's Bylaws recognized the proper scope and authority of an IRP panel when it correctly stated: "The holdings of the Independent Review Panel are advisory in nature; they do not constitute a binding arbitral award." *ICM* IRP Panel Declaration (19 Feb. 2010), ¶ 152; *see also, id.* at ¶ 134 (the IRP Panel's declaration "is not binding, but rather advisory in effect"), *available at* http://www.icann.org/en/news/irp/icm-v-icann/news/irp/-panel-declaration-19feb10-en.pdf.

RESP. EX. 4

JASGLOBAL ADVISORS

gTLD APPLICATION PROCESSING: INITIAL EVALUATION

QUALITY PROGRAM REPORT

ICANN

26 August 2014

FOR PUBLIC RELEASE

TABLE OF CONTENTS

1   Summary

New gTLD application evaluation was a labor-intensive business process performed by multiple vendors and hundreds of individuals on a global basis. Initial Evaluation (IE) included seven distinct evaluation types: applicant background, financial capability, technical/operational capability, registry services, geographic names, DNS stability, and string similarity. For commercial and practical reasons, including application volume and handling conflicts of interest between an applicant and evaluator, multiple evaluator firms were contracted. Application evaluation was performed against detailed criteria as published in the *New gTLD Applicant Guidebook* (AGB).[1]   Quality and consistency of evaluation across all applications and all evaluator firms was a key business requirement for ICANN. Given the importance of demonstrable quality, 50% of the applications were subject to quality sampling in some capacity and 100% of the applications were reviewed using analytical techniques. All application data was subject to a suite of manual and automated data consistency checks performed by ICANN staff and JAS.

At a high level, the new gTLD application evaluation training and quality program was designed to both *improve* and *measure*:

- **Consistency/Precision:** a measure of the degree of agreement between independent assessments of a particular sample. Precision is expressed in terms of the standard deviation of the consistency rating among primary and independent half-blind *de novo* assessments (calculation of the consistency rating is described in Section 5.2). Precision is important because multiple evaluator firms should produce similar results given similar applications. Situations where precision was not as expected triggered additional training, documentation, and may inform future process revisions.
- **Accuracy:** a measure of the degree of agreement of a sample with an accepted reference. In the case of application evaluation, the accepted reference is the result of "work-out" conferences between the primary evaluator firm, the quality firm, and ICANN when discrepancies occur. Accuracy is expressed in terms of percent of the samples reflecting the expected value. Situations where accuracy was not as expected triggered additional training, documentation, and may inform future process revisions.
- **Process Fidelity:** a measure of the alignment between the expected process per the vendor's contract and the actual process performed for a given application. Process fidelity is expressed in terms of a percent of the samples where a post-evaluation Procedural Inspection indicated that proper procedures were followed.

As quality measurement and improvement are typically somewhat competing goals (performing quality improvement on a process while measurement is occurring leads to a degree of Heisenberg uncertainty), the overall quality program was designed primarily to monitor, incent, and improve quality during evaluation with a secondary objective of providing analysis and a quantitative baseline to assess the process in arrears and inform future rounds.

---

[1] *New gTLD Applicant Guidebook*, ICANN, 4 June 2012,
http://newgtlds.icann.org/en/applicants/agb

The training and quality program is comprised of six functions:

Unified Training

A unified, cross-firm approach to training was developed and implemented prior to the commencement of production evaluation. Unified training was essential in bringing together the evaluation operations of all evaluator firms – particularly the large-scale operations of the three technical/operational and financial firms – and maintaining ongoing alignment in a challenging and dynamic environment.

For technical/operational and financial panels – the most complex evaluations – all three evaluator firms shared training materials and conducted joint training sessions. For other panels, standardized training templates were utilized.

Content Reviews

Content Reviews were discussions between two or more evaluator firms that had completed a full or partial review of the same application. Content Reviews were designed to improve consistency/precision and accuracy among the three technical/operational and financial evaluator firms. Content Reviews of selected applications were performed as a part of the comprehensive training program prior to commencement of production evaluation and additionally throughout Initial Evaluation to maintain communication and alignment between all three evaluator firms. One special case of content reviews was the applicant-facing Clarifying Question (CQ) pilot that provided immense value. Of the 1917 application IDs receiving Prioritization Draw results, 107 applications were involved in a complete or partial content review at some point.

Blind Content Inspections

Content Inspections were half-blind independent evaluation and scoring of a randomly selected set of applications. The Content Inspection included review of the primary evaluator firm's Clarifying Questions (CQs) prior to issuance, and independently generated final scoring by the quality evaluator firm. Blind Content Inspections were designed to measure and improve consistency/precision and accuracy among the three technical/operational and financial panel firms. The inspections were half-blind in that the primary panel firm did not know in advance which applications were selected for inspection and the quality firm was not aware of the primary firm's scores in advance. Content Inspections were conducted on a randomly selected 15% of the 1917 application IDs receiving Prioritization Draw results.

Blind Procedural Inspections

Procedural Inspections were half-blind reviews of the primary firm's records to gain confidence that the agreed-upon processes and procedures were performed as expected. Procedural Inspections were designed to measure the process fidelity of the panel firms. The inspections were blind in that the primary panel firm did not know in advance which applications were selected for inspection. Procedural Inspections were conducted on a randomly selected 35% of the 1917 application IDs receiving Prioritization Draw results.

Analytics

ICANN received in excess of 1900 applications, largely comprised of unstructured text and attachments. Many latent similarities existed between the applications due to common applicants, consultants, and service providers. Analytical tools were developed to highlight these latent similarities and improve confidence that applications with similar content received a similar final disposition. Moreover, in excess of 5000 Clarifying Questions (CQs) were generated as a part of evaluation; as CQ generation is labor-intensive and subject to a range of error modalities, analytical systems provided automated quality and content checks of CQs prior to issuance.

Data Consistency Checks

Application evaluation was a large-scale global operation with a number of dynamic components. Ensuring that ICANN's systems of record were both internally consistent and accurately reflective of the authoritative evaluation results as documented in numerous vendor reports was critical. Automated systems provided routine data validation and crosschecking spanning numerous systems and record types to reduce likelihood of consistency errors.

1.1   Program Coverage

While designing training and quality programs, the process of application evaluation was divided into *content* and *process* components. The process components covered each vendor's obligation to perform their contracted duties and interact with the broader system and ICANN as specified, and the general requirement to maintain data consistency across several systems given emergent and fast-moving processes. The content components covered each vendor's obligation to evaluate the application pursuant to the Applicant Guidebook and all relevant guidance. The training and quality program recognized and provided coverage to both of these at multiple points in time during application processing.

Content-oriented aspects of the training and quality program were focused on the technical/operational and financial panel types due to the nature of these evaluations and the complexity and scale of the combined evaluation operations of all three evaluator firms. For all panel types, the process-oriented aspects of the quality program were focused on ensuring that all evaluator panels followed procedures agreed upon with ICANN.

Panel Type	Prior to CQ Release	Final Scoring (IE)	
	Content	Content	Process
Financial	Training Content Review Blind Content Inspection Analytics	Ongoing Training & Communication Content Review Blind Content Inspection Analytics	Training Blind Procedural Inspection Data Consistency Checks
Technical/Operational	Training Content Review Blind Content Inspection Analytics	Ongoing Training & Communication Content Review Blind Content Inspection Analytics	Training Blind Procedural Inspection Data Consistency Checks
Registry Services	Training	Analytics	Training Blind Procedural Inspection Data Consistency Checks
DNS Stability	Training		Training Blind Procedural Inspection Data Consistency Checks
String Contention	Training		Training Blind Procedural Inspection Data Consistency Checks
Geographic	Training		Training Blind Procedural Inspection Data Consistency Checks

Table 1: Training and Quality Program Coverage

1.2 Program Scope

The training and quality programs were operational prior to the commencement of production evaluation and continued through the completion of Initial Evaluation. Extended Evaluation was not included in the scope of the quality program.

1.3 Roles and Responsibilities

JAS Global Advisors LLC ("JAS") was responsible for designing the overall training and quality programs based on requirements developed with ICANN. JAS was responsible for administering the quality program during execution, coordinating content reviews, performing Content Inspections, performing Procedural Inspections, implementing analytical and consistency checking systems, and reporting results. JAS was the primary technical/operational and financial reviewer for fewer than 50 applications and only in situations where no other technical/operational and financial firms were available due to a conflict of interest with the applicant. Related to the training and quality programs, all evaluator firms had obligations to provide data, participate in training activities, produce documentation, and generally cooperate with training and quality activities.

2   Program Objectives

The training and quality program was designed to achieve multiple objectives. The most important objective was to provide confidence that applications with similar content received a similar final pass/fail disposition. It's important to note that with respect to scoring, the quality program viewed Initial Evaluation as a pass/fail exercise consistent with the description in the Applicant Guidebook. No meaning is or should be imparted to numerical differences in score between two passing (or two failing) applications.

To achieve this objective, training and quality programs focused on:

- Upfront "calibration" among evaluator firms via unified training, discussion, scoring exercises, and pilots;
- Encouraging and maintaining ongoing communication among evaluator firms throughout the process via training, scoring exercises, and comparison of evaluation results;
- Leveraging analytics to identify latent similarities and determine potential scoring inconsistencies; and
- Providing visibility and early notification to ICANN in the event inconsistencies were discovered.

Clearly, *communication* and *visibility* are the central themes. Given the scale and nature of evaluation, absent active mechanisms to maintain communication between firms and with ICANN, there was a risk that evaluator firms would become isolated and produce increasingly divergent results over time. A central objective was to maintain open communication among all participants during the entire evaluation process.

A second central objective was to provide ICANN visibility into evaluation quality throughout the evaluation time period. Absent active mechanisms to assess quality during evaluation, it would be hard to quickly determine if quality was acceptable or unacceptable, converging or diverging, or if process improvements or additional training was required, leading to a sort of unmanaged Markov process.

By creating active communication and visibility mechanisms, ICANN was able to successfully keep the evaluation process under control.

Additionally, the program had the following secondary objectives:

- Improve quality of issued CQs
- Reduce data and clerical errors
- Provide quantitative baseline for future rounds

3 Content Reviews

Content Reviews were discussions between two or more firms that had completed a full or partial review of the same application. Content Reviews were designed to improve consistency/precision and accuracy among the three technical/operational and financial evaluator firms.

Content Reviews were performed early in the process – during training and early in Initial Evaluation – in order to add maximum value to the calibration process; subsequent and less frequent Content Reviews were performed throughout Initial Evaluation to encourage continued communication and alignment, particularly around emergent issues. Content Reviews were performed on technical/operational and financial panel results.

One special case of content reviews was the applicant-facing Clarifying Question (CQ) pilot that provided immense value; multiple pilots that were not applicant-facing were also conducted.

3.1 Process and Sampling

Content Reviews leveraged approximately 107 applications that both a primary reviewer and a secondary reviewer had evaluated (in part or in full) in some capacity. An effort was made to select applications for Content Review that represented a wide range of applicants and service providers to maximize the value of the exercise. Applications utilized for Content Reviews were not eligible for selection for Content Inspection.

3.2 Roles and Responsibilities

JAS coordinated Content Review activities among the three technical/operational and financial evaluator firms. Prior to the availability of actual applicant data, JAS developed several mock applications as a part of the training materials.

3.3 Exceptions

Differences in scoring were discussed and remediated between the evaluator firms with input from ICANN requested on an as-needed basis.

3.4 Metrics and Reporting

The primary objective was to facilitate calibration and maintain communication; the Content Review program did not generate metrics.

4   Blind Content Inspections

A statistically relevant number of technical/operational and financial evaluations were subject to half-blind Content Inspection reviews performed on a *de novo* basis. A *de novo* review is a complete and independent review performed "from the beginning" by the quality firm simultaneously with – but independently from – the primary evaluator firm. The review is also half-blind; the primary evaluator firm did not know in advance which applications were selected for Content Inspection. The intent of the review was to measure CQ and scoring consistency and accuracy against scoring guidance and training, and to provide an opportunity to quickly detect quality and consistency issues.

4.1   Process and Sampling

Blind Content Inspections were selected via random ordering of the 1917 application IDs receiving Prioritization Draw results. JAS performed the random ordering via computer on 20 Dec 2012. Note that withdrawals reduced the size of the population, requiring limited selection of additional samples to compensate for the aforementioned issues. The first 15% (288) applications in the random ordering were selected for Content Inspection. As additional samples were needed due to withdrawals or other factors requiring de-sampling, applications starting at 289 in the random ordering were selected.

Final metrics for the quality control program were taken on 28 August 2013 at the conclusion of Initial Evaluation work and are as follows:

Total Active Applications (28 Aug 2013)	1768
Applications Sampled	274
Sampled Proportion	15.50%

Table 2: Content Inspection Sampling

4.2   Metrics

The blind Content Inspections produced the following quantitative metrics:

- **Consistency Rating (per question).** This is the simple numeric pairwise comparison between the primary and QC review final scores on a per question basis. A pairwise comparison of 0 indicates that the primary and QC review final scores are identical whereas a pairwise comparison of +1 or -1 indicates the final scores differ. Instances of non-objection were de-sampled (see below).

  For the purpose of QC, no distinction is made between passing scores with score = 1 and score > 1. Any score greater than or equal to 1 will be considered a 1 for the purpose of QC – for both the primary firm score and the QC firm score. For example, a score of 2 is equal to a score of 1 and to a score of 3 – all were transformed to a score of 1 prior to calculation of the consistency rating. This transformation is necessary to align the QC program with the pass/fail design of Initial Evaluation as described in the Applicant Guidebook.

- **Consistency Rating (per application).** This is a proportional measure of consistency of final (pass/fail) dispositions for a given application. The quality evaluator firm maintained the option to deem an application "non-objection" meaning that for reasons related to maintaining the

integrity of the half-blind selection, not enough information was available to score the application but the quality evaluator firm did not find sufficient cause to disagree with the primary firm's pass/fail disposition.

4.3 Roles and Responsibilities

JAS was the quality evaluator firm. If an application was selected for Content Inspection where JAS was the Primary Review Firm (due to conflict with both primary evaluator firms), the application was de-sampled for quality control purposes and the next application in the random ordering that had not already been released was selected.

JAS' small number of primary evaluations were therefore ineligible for Content Inspection; however, as JAS was a party to each and every consistency rating metric, evaluation of JAS' performance as compared to the other firms was evident and obvious.

4.4 Exceptions

Differences in scoring appear in the consistency rating; exceptions were brought to ICANN's attention as soon as they were discovered for discussion with the evaluator firms as necessary.

4.5 Results

Content Inspections generated metrics on a horizontal basis (per question across applications) and on a per-application basis. Content Inspection samples were taken before and after the Outreach phase. Outreach was an ICANN process that in limited situations allowed the applicant to provide missing information that may have stemmed from an oversight.

Shown below are statistics describing the Content Inspection samples taken prior to Outreach; following Outreach, all primary and Content Inspection evaluations were in agreement (consistency rating = 0). Small variances in the sample size in the table below occurred because in certain limited circumstances the quality firm asserted "non-objection" discrepancies as described above and those individual questions were de-sampled for statistical purposes.

In summary, prior to the Outreach phase there were six individual application question/response instances (1 technical/operational and 5 financial) where a bona-fide scoring discrepancy existed that would have impacted the final disposition of the application (moving an application from a pass to a fail or vice versa). To highlight root causes, for purposes of this analysis and presentation, a single scoring issue that cascaded into multiple scoring discrepancies has been reduced to the single root cause and the cascading discrepancies are not reflected here. For example, a discrepancy in financial cost calculations may cascade into a discrepancy in the question 50 Continuation of Operations (COI) Instrument calculation; the former is indicative of a root cause quality issue whereas the latter is not.

Applications containing a question that received a zero score following the Clarifying Question phase proceeded to the Outreach phase. All of the per-question discrepancies below were resolved during Outreach; following Outreach, all primary and Content Inspection evaluations were in agreement and every question selected for Content Inspection received a passing (non-zero) score.

Question #	n where consistency rating = 0 (Consistent)	n where consistency rating != 0 (Not Consistent)	Standard Deviation of Consistency Rating for the Population
24	261	0	0.000
25	256	0	0.000
26	261	0	0.000
27	260	0	0.000
28	261	0	0.000
29	261	0	0.000
30	261	0	0.000
31	261	0	0.000
32	260	1	0.024
33	260	0	0.000
34	261	0	0.000
35	261	0	0.000
36	261	0	0.000
37	261	0	0.000
38	261	0	0.000
39	261	0	0.000
40	261	0	0.000
41	261	0	0.000
42	261	0	0.000
43	260	0	0.000
44	N/A – Optional	N/A – Optional	N/A – Optional
45	258	2	0.037
46	261	1	0.000
47	261	0	0.000
48	261	0	0.000
49	261	0	0.000
50	256	2	0.041

Table 3: Per-Question Consistency Rating

An application must have no individually failing questions (score=0) and reach a minimum score threshold in both technical/operational and financial questions in order to pass evaluation. As an application with all passing individual questions may still fail due to insufficient total points, consistency was also analyzed on a per-application basis to capture this aspect.

In summary, prior to the Outreach phase there were five (5) applications where a bona-fide scoring discrepancy existed that would have impacted the final disposition of the application (moving an application from a pass to a fail or vice versa).

Note that this analysis is considering an application as a whole whereas the previous analysis is considering all question/response instances. In the former, there were six (6) question/response instances where the consistency rating was not zero; in the later, there were five (5) whole applications where the final disposition was not consistent pre-Outreach. All inconsistencies were resolved Post Outreach.

Application Status	n	%
Consistent Pre-Outreach	261	95.26%
Not Consistent Pre-Outreach	5	1.82%
No Objection	8	2.92%
Consistent Post Outreach	274	100.00%

Table 4: Per-Application Consistency Rating

Analyzing the five (5) instances where there was a scoring discrepancy prior to Outreach on a per-evaluator firm basis revealed balanced data (note that aliases are used to identify evaluator firms):

Status	n
Evaluator Firm Alpha consistency rating as compared to quality firm is > 0 (Evaluator Firm Alpha scored **higher** than quality firm)	1
Evaluator Firm Alpha consistency rating as compared to quality firm is < 0 (Evaluator Firm Alpha scored **lower** than quality firm)	2
Evaluator Firm Bravo consistency rating as compared to quality firm is > 0 (Evaluator Firm Bravo scored **higher** than quality firm)	0
Evaluator Firm Bravo consistency rating as compared to quality firm is < 0 (Evaluator Firm Bravo scored **lower** than quality firm)	2

Table 5: Per Evaluator Firm Analysis of Application Discrepancies

4.6 Analysis and Discussion

Given the overall scale, scope, and challenge of Initial Evaluation, evaluation was remarkably consistent. Several points are worth noting:

- Evaluator firms spent considerable effort in training and calibration, and clearly it proved effective. The Applicant Guidebook describes Initial Evaluation as a pass/fail exercise (as long as the minimum point requirements are met, there is no benefit in receiving additional points and no penalty in receiving fewer points). As such, during initial training and calibration, evaluator firms focused on "zero/non-zero" issues/scoring to gain confidence that pass/fail alignment

would be high. As a result, pass/fail consistency was very high but raw numeric scoring – which included the additional points – was less consistent. Analysis of the additional point system beyond the minimum pass/fail thresholds was not a part of the design of the quality program.

- Consistency of CQs was desirable but not always possible. Variance in internal firm processes and other factors reduced the overall consistency of CQs. However, pass/fail application disposition remained high despite variance in CQs. A contributing factor is that a significant proportion of CQ inconsistencies were related to additional points components of questions (criteria required to receive a score of two (2) or three (3) on a question).

- Consistency issues are highly concentrated in very few questions, particularly financial questions 45 and 50. Anyone familiar with the application process will recognize these questions and not be at all surprised with this finding. The fact that these questions were the subject of the majority of post-AGB ICANN guidance – both to applicants and evaluators – underscores the localized difficulties present in these two questions. Discrepancies that surfaced in questions 45 and 50 tended to be systemic issues (symptoms of unanticipated scenarios and/or broader lack of clarity) whereas the discrepancies that surfaced in other questions tended to be isolated and unusual corner cases.

- Numerous subjective terms (such as "adequate," "commensurate," "comprehensive," "highly developed," and similar terms) appear frequently in the Applicant Guidebook. Evaluator firms and ICANN spent significant effort defining these terms crisply and calibrating for the purpose of consistent evaluation. While the results show that this effort was largely successful, additional definition of subjective terms in future revisions of the Applicant Guidebook would be of value.

- The Applicant Guidebook did not recognize the concept of a Registry Service Provider nor did it contemplate an applicant describing a registry being run as a cost center with limited or no revenue. Ambiguity surrounding these concepts was the root cause of several calibration discussions and scoring discrepancies. Overt recognition of these concepts in future revisions of the Applicant Guidebook would be of value.

5 Blind Procedural Inspections

Work performed by technical/operational, financial, string similarity, and geographic name panels/providers was subject to a Procedural Inspection on a statistically relevant randomly selected sample of applications. The intent of the Procedural Inspection was to provide assurance that the application was fully processed, and that all panel providers completed (and provided evidence of completing) all the steps required of them as documented in the Applicant Guidebook and individual SOWs. A team of JAS personnel conducted the Procedural Inspections.

Each of the five panel types had a "procedural checklist" which was developed by ICANN and the panel providers in advance. Multiple firms performing the same function (e.g. financial review) used the same procedural checklist. The procedural checklist was the basis on which the Procedural Inspections were conducted.

5.1 Process and Sampling

Blind Procedural Inspections were selected via random ordering of the 1917 application IDs receiving Prioritization Draw results. The first 35% (671) applications in the random ordering were selected for Procedural Inspection; if additional samples were needed due to withdrawals, selection of an application where the applicant is conflicted with both primary evaluator firms, or other factor requiring de-sampling, applications starting at 672 in the random ordering were selected. Each selected application was subjected to a Procedural Inspection for <u>all</u> panel types. Note that the random ordering generated for Procedural Inspections was different – and independent – from the random ordering generated for Content Inspections.

Procedural Inspections were conducted on final work products after final scoring was submitted to ICANN.

Final metrics for the quality control program were taken on 28 August 2013 and are as follows:

Total Active Applications (28 Aug 2013)	1768
Applications Sampled	639
Sampled Proportion	36.14%
Compliance Rate	99.84%

Table 6: Procedural Inspection Sampling

As the String Similarity panel operated on unique strings, a separate random ordering and selection were performed for these Procedural Inspections. Content Inspection metrics for String Similarity are as follows:

Unique Strings (28 Aug 2013)	1388
Applications Sampled	490
Sampled Proportion	35.30%
Compliance Rate	100.00%

Table 7: String Similarity Procedural Inspection Sampling

5.2   Metrics

Each Procedural Inspection reviewed the primary evaluation as a whole and generated one metric per application. The resulting metric is an assessment of the fidelity with which the primary evaluation followed the agreed-upon Procedural Checklist for the specific application. The metric is one of: *Compliant (C); Minor Discrepancy (MD); Significant Discrepancy (SD).*

5.3   Roles and Responsibilities

JAS was the quality evaluator firm. If an application was selected for Procedural Inspection where JAS was the Primary Review Firm (due to conflict with both primary evaluator firms), the application was de-sampled for quality control purposes and the next application in the random ordering that had not already been released was selected.

5.4   Exceptions

Exceptions were brought to ICANN's attention as soon as they were discovered for discussion with the evaluator firms as necessary.

5.5   Results

Procedural Inspections generated metrics on a per-evaluator firm basis for each evaluation type. One sample was taken after the primary evaluator firm submitted final results for an application that was selected for Procedural Inspection.

Evaluation Type	Evaluator Firm (alias)	n Compliant	n Minor Discrepancy	n Significant Discrepancy
Technical/Operational	Charlie	329	1	0
Technical/Operational	Delta	309	0	0
Financial	Charlie	329	1	0
Financial	Delta	309	0	0
Geographic	Echo	399	0	0
Geographic	Foxtrot	240	0	0
DNS Stability	Golf	639	0	0
Registry Services	Lima	639	0	0
String Similarity[2]	Oscar	490	0	0

Table 8: Per Evaluator Firm Analysis of Procedural Inspections

5.6   Analysis and Discussion

Each evaluation vendor's adherence to agreed-upon evaluation procedures was a critical success factor for the program. Procedural Inspection results show that this adherence did indeed occur.

---

[2] Note that String Similarity Procedural Inspections were performed on 490 evaluations based on applications for 1388 unique strings.

6 Analytical System Review

ICANN received in excess of 1900 applications, largely comprised of unstructured text and attachments. Many latent similarities existed between the applications due to common applicants, consultants, and service providers. Analytical tools were developed to achieve three objectives:

- Provide confidence that all similar applications received similar final (pass/fail) dispositions;
- Help identify potential CQ inconsistencies that could lead to a discrepancy in final disposition;
- Improve the quality of CQs by programmatically checking application and Applicant Guidebook citations.

While the previously described quality procedures applied to a sample of applications, analytical techniques were performed on all applications and CQs.

The analytical system allowed the evaluator firms, quality firm, and ICANN to visually review connections between similar applications, the CQs generated for those applications, the responses to those CQs from applicants, and the final score on an ongoing basis. While complete and absolute consistency through all of those steps would be a desirable – albeit Quixotic – outcome, in reality, analytics allowed discrepancies to be identified and reviewed for impact. Potentially problematic discrepancies were identified and rectified.

6.1 Process

Financial and technical/operator evaluator firms interacted with the analytical system at three points in time:

1. Following submission of CQs to ICANN's application management system (but prior to their transmission to the applicant);
2. Prior to submitting final scores to ICANN; and
3. Following submission of final scores to ICANN.

Following submission of CQs to ICANN's application management system, the analytical system programmatically matched quotes and citations appearing in the CQs to the relevant application and the Applicant Guidebook. Matches were confirmed and potential mismatches were flagged for manual verification. This step reduced the occurrence of misquotes and copy/paste errors given that thousands of similar CQs were generated. This was an especially important error mode to control, given that oft-quoted portions of the applications were confidential. Additionally, the analytical system compared the CQs for the submitted application to the CQs generated for similar applications and flagged discrepancies for manual verification.

Following submission of final scores to ICANN's application management system, the analytical system compared the scores of the submitted application to the scores of similar applications previously submitted. Potential discrepancies were flagged for manual verification.

Finally, at the completion of Initial Evaluation, JAS performed an analytical review of all applications that completed Initial Evaluation successfully vs. those that were referred to Extended Evaluation.

6.2 Analysis and Discussion

The sheer volume and unstructured nature of the application data necessitated an analytical approach. During each weekly application processing cycle, reports were delivered to evaluator firms and ICANN containing the results of the analytical reviews described above. As manual verification confirmed or refuted analytical results, false positives were identified and tuned out to improve future efficacy of the system. Noting that analytical reviews were a backstop measure designed to catch issues that remained undetected relatively late in the application cycle, a low and decreasing number of analytical system exceptions were indicative of high quality work by the evaluator firms. While there was an initial burst of analytical system exceptions, by the end of Initial Evaluation, very few valid analytical exceptions were being identified. This was an indication that the evaluation system was performing adequately and that the internal quality procedures being performed by each firm were effective. This was the desired behavior.

Following the completion of Initial Evaluation, JAS performed an analytical comparison of all applications that completed Initial Evaluation successfully vs. those that were referred to Extended Evaluation and found that the applications that were referred to Extended Evaluation were materially different than the applications that passed Initial Evaluation successfully. As this analysis took the entire population of applications into consideration, this step served as a valuable system-wide double-check on all of the previous sample-oriented quality programs.

Despite acknowledged inconsistencies in CQs and numeric scores (above and beyond the passing thresholds), this last analysis provided a strong indication that – when the process reached completion – all similar applications received passing scores and the applications referred to Extended Evaluation correctly were individual special cases requiring additional clarification.

7 Overall Analysis, Discussion, and Recommendations

The ICANN New gTLD evaluation program resulted in the successful evaluation of over 1900 applications from a full range of global applicants, delivering a demonstrably high level of evaluation consistency while providing ICANN with the practical and commercial benefits of evaluator depth and diversity. Some additional overall comments in closing:

1. The extensive advanced preparation, training, synchronization, and evaluation exercises (pilots) undertaken by the technical/operational and financial evaluator firms were essential and probably the single largest critical success factor. As verified by the positive quality program results, a unified approach to these activities coalesced the team and substantially mitigated the risk of isolation and inconsistent or divergent evaluations.

2. As quality practitioners well know, one value of a proactive quality program is that the mere (visible) existence of such a program helps incent the desired behaviors. In this case, it is highly probable that the existence of a visible and well-publicized proactive quality program properly incented all evaluation panel vendors to be appropriately cognizant of evaluation consistency, accuracy, and process fidelity, and perform accordingly.

3. Although the questions were provided in advance and there was an expectation that applicants would be clear on the material, it was apparent that many applicants, including sophisticated applicants, were confused as to how to respond to the questions. This resulted in two undesirable effects: (a) applicants tended to "over-respond" to the application, adding unnecessary volume and complexity; and (b) there was more effort put into clarification communications (including CQs) than was probably intended in the original vision. While not "providing the answers" there is an opportunity to make the application process more objective and deterministic for both applicants and evaluators. Reducing subjectivity of evaluation will enable improved quality and consistency and reduce costs associated with extensive synchronization activities.

4. The lack of structured application data was an impediment during evaluation; future application rounds should capture data in a more structured format, greatly facilitating evaluation, quality reviews, and subsequent processes like contracting.

5. Several questions, particularly technical/operational questions, have overlapping remits complicating evaluation, quality processes, and unnecessarily creating the appearance of inconsistency. Some topics, such as the use of IDNs, often have material spread throughout several questions. This makes it harder for applicants to "know what to put where" and for evaluators to find the information they're looking for. A highly structured application will help address this issue.

6. Releasing results incrementally opened the opportunity for difficult-to-manage inconsistencies. Future rounds designed for one release of results at the end will make comprehensive consistency and quality checking more effective.

7. The publication of detailed numeric scores confused and undermined the AGB-driven premise that evaluation was pass/fail. Inconsistencies in numeric scores incorrectly sent a message that evaluation was much more inconsistent than the final results and the quality programs assert.

Future application rounds should either publish results as pass/fail only, or re-calibrate the entire process to produce numerically consistent scores.

8. Financial evaluation of questions 45 and 50 exhibited systemic issues that made consistent evaluation difficult. Recognizing applicants that choose to run their registry as a cost center and revising the approach to the problematic question 50 regarding the Continuity of Operations Instrument will go a long way to increase the evaluation consistency of these questions.