ISSN:1738-6764, eISSN:2093-7504

Development of a Capability Assessment Model for Korea's National HPC Centers

Myoungju Koh [1,]*, Noori On [2], Hyungwook Shim [3] and Minho Suh [4]

[1] Korea Institute of Science and Technology Information; myju@kisti.re.kr
[2] Korea Institute of Science and Technology Information; claireon@kisti.re.kr
[3] Korea Institute of Science and Technology Information; shw@kisti.re.kr
[4] Korea Institute of Science and Technology Information; mhsuh@kisti.re.kr
* Correspondence

https://doi.org/10.5392/IJoC.2025.21.4.140

Manuscript Received 27 August 2025; Received 4 December 2025; Accepted 4 December 2025

Abstract: *Recognizing the strategic importance of domain-specific high-performance computing (HPC), Korea designated six specialized national centers in 2023, with plans to expand this number to ten by 2030. To ensure effective operation and governance, a capability assessment providing a current and comparable view of these centers is essential. This study proposes a policy-linked framework for assessing the capabilities of Korea's national HPC centers. The assessment model is organized into six categories: Foundation, Input, Process, Output, Outcome, and Policy Alignment, with a set of representative indicators specified for each category. These indicators were developed through an inductive analysis of relevant literature to identify key measures, along with a deductive integration of attributes specific to centers with statutory policy mandates. Each indicator is normalized against field-specific targets, and category-level and within-category weights are determined using the Analytic Hierarchy Process (AHP) to produce a single composite capability score for each center.*

Keywords: High-performance Computing; HPC Capability Assessment; Analytic Hierarchy Process; Capability Assessment Model; National HPC Center

1. Introduction

Recognizing the strategic importance of advancing high-performance computing (HPC), Korea has designated a National Supercomputing Center along with domain-specific specialized centers to drive adoption. The Act on the Utilization and Promotion of National Supercomputing and its Enforcement Decree (2011) [1] established the legal basis for the national center. Building on this foundation, the National Supercomputing Innovation Strategy (2021–2030) [2] announced a plan to designate ten specialized centers by 2030; six were designated in 2023 and are operating under defined mandates.

Collectively, these actions provide policy momentum for strengthening domain capabilities and shared use within Korea's national supercomputing ecosystem, which is slated to expand to ten centers by 2030. To support effective operations, a transparent, comparable, and policy-linked model to assess capabilities across centers is required.

Although related indicator sets exist, they generally do not fully implement (i) field-specific, practice-calibrated normalization, (ii) a policy-alignment lens capturing shared use and standardized reporting, or (iii) a single, auditable AHP-weighted composite score suitable for national governance.

Accordingly, this study develops a conceptual model to assess center capabilities, organized into six categories—Foundation, Input, Process, Output, Outcome, and Policy Alignment—and, for each category, specifies representative indicators. The indicator set was constructed through an inductive analysis of cases in the relevant literature to surface salient measures, complemented by a deductive incorporation of attributes characteristic of centers endowed with statutory policy mandates.

A preliminary version of this work was presented at ICCC(2024) [3]. This study substantially extends that version by: (i) adding a Policy Alignment layer tailored to statutory mandates (ii) standardizing target-based normalization for all indicators, (iii) formalizing AHP weighting at category and indicator levels with consistency checks.

2. Related Works

This chapter reviews related research from two main perspectives. First, it offers an exploratory survey of the key indicators used for assessment. Second, it examines the theoretical foundations that inform the design of assessment models.

2.1 Related Literature

The logic model—formalized and widely disseminated in program-evaluation practice, with the W.K. Kellogg Foundation (2004) serving as a canonical guide—is used to make complex initiatives evaluable and auditable by specifying a transparent causal chain (inputs → activities → outputs → outcomes) [4]. In our context, it helps locate each indicator at the correct stage of value creation: resources such as infrastructure capacity, operating funds, and staffing (inputs); operational work such as allocation, queue management, user support, and training delivery (activities); immediate, countable service products such as active users, supported projects, and allocated compute hours (outputs); and research/innovation effects such as, publications, software releases, patents, and technology transfers (outcomes). By preventing category drift (e.g., treating an output as an outcome) and by attaching appropriate evidence artifacts at each link (logs for activities, program statistics for outputs, bibliometrics or transfer records for outcomes), the logic model supplies the causal architecture on which our evaluation categories and audit trail rest.

The Balanced Scorecard (BSC)—introduced by Kaplan and Norton (1992) to translate strategy into a balanced portfolio of measures beyond financials—is used to ensure that measurement reflects strategy execution rather than a narrow operational slice [5]. Its four perspectives (financial, customer/stakeholder, internal process, learning & growth) act as a coverage test that ties daily operations to strategic intent and guards against siloed

(e.g., capacity-only or publication-only views). Applied to research-infrastructure assessment, this means placing service reliability and efficiency under internal processes (uptime, queue, allocation cycle), user and partner value under customer/stakeholder (active users, supported projects, industry collaboration), resource sustainability and readiness under financial (OPEX per core-hour, electricity-cost share), and capability building under learning & growth (staffing sufficiency, training budget). In short, the BSC compels a multi-perspective lens on performance that complements the logic model's causal staging.

Recent work has explored pushing the Balanced Scorecard (BSC) beyond its four classical perspectives to include governance and compliance objectives. A 2023 bibliometric review maps the field's shift toward themes such as integrated reporting, sustainability/SBSC, and strategic performance management, indicating growing interest in governance-related extensions of the BSC [6]. Panitz, Wiener, and Amberg (2010) propose a Compliance Balanced Scorecard, showing how the standard BSC can be repurposed for regulatory adherence by embedding compliance workflows, remediation tracking, and risk classification into a modular, benchmarkable design [7]. Kaplan and Nagel (2003/2004) outline a three-part scorecard program—enterprise, board, and executive scorecards—to help boards discharge responsibilities such as strategy approval, oversight, and compliance monitoring via a dedicated Board Balanced Scorecard aligned with enterprise and executive views [8]. Together, these studies demonstrate both the feasibility and the academic grounding for extending BSC frameworks to governance and compliance domains. In this spirit, our study introduces Policy Alignment as a deliberate governance lens layered onto the classical BSC, tailored to national HPC evaluation: it captures statutory requirements (e.g., shared utilization, standardized reporting, master-plan compliance) within the same scorecard architecture used for strategy execution, making governance performance visible, auditable, and comparable across centers.

In this study, we deductively design the categories of the evaluation model based on the theories discussed above, and then exploratively derive and map key indicators.

2.2 Indicator Families

The PRACE consortium maintains a KPI portfolio covering infrastructure usage (supported projects, allocated core-hours), training and outreach, and program-level participation (including industry/SME) to promote transparency and continuous improvement across a multi-country research infrastructure [9]. These

KPIs are designed to monitor mission delivery and steer incremental improvement rather than to attribute causal impact, which makes them a useful reporting and governance baseline for national ecosystems seeking comparability without overclaiming impact.

The data-center community (e.g., the Green500 and EE-HPC Working Group) maintains standardized power-measurement methodologies and the Power Usage Effectiveness (PUE) ratio with the explicit aim of transparent, repeatable energy-efficiency benchmarking across heterogeneous HPC facilities [10, 11]. These materials position PUE and multi-level measurement procedures as instruments to monitor operational efficiency and steer incremental improvement, rather than as direct measures of scientific impact—making them a practical reporting and governance baseline for centers that seek comparability of facility performance.

The HPC facility-operations community (e.g., national user facilities and program guidance) maintains an operational portfolio—uptime, utilization, queue wait time, allocation cycle time, training volume, and help-desk responsiveness/SLAs—with the explicit aim of monitoring service reliability and user access/throughput for planning and oversight [12]. These indicators are positioned as instruments to track operational performance and surface bottlenecks rather than as causal measures of scientific impact, which makes them a practical reporting and governance baseline for centers that need comparable views of service quality.

Taken together, the PRACE materials, the PUE/power-measurement corpus, and facility-operations reports provide a valuable reporting and learning baseline but do not, by themselves, supply (i) a standardized, auditable center-level survey instrument, (ii) field-adjusted targets and normalization suitable for cross-center comparison, (iii) transparent multi-criteria weighting of categories and indicators, or (iv) policy-alignment constructs required under national policy constraints. Accordingly, we treat PRACE themes (reliability/utilization, training, engagement, outputs/outcomes), PUE and standardized power-measurement discipline, and operations metrics (uptime, queue wait, allocation cycle, help-desk responsiveness, training volume; active users/supported projects) as selection priors for the Foundation/Process/Output layers, while our model supplies the missing operational elements: a structured, evidence-backed questionnaire; field-specific targets with target-based normalization; AHP-based hierarchical weighting; and policy-alignment checks (e.g., shared-utilization interconnects and standardized reporting), so that energy stewardship, service reliability, and user access are scored, auditable, and comparable across centers in Korea.

2.3 Weighting method

In this study, we investigate the following weighting methods to evaluate and synthesize key indicators across and within categories.

Originating with Saaty (1980) [13] and consolidated by methodological expositions and surveys [14, 15], AHP structures complex goals into a hierarchy, elicits ratio-scale priorities via pairwise comparisons, and verifies judgment quality through the consistency ratio (CR). In practice, AHP is used when heterogeneous criteria require transparent, stakeholder-involved weighting and when decision makers must defend how importance was assigned.

AHP organizes objectives into a hierarchical structure, starting with composite competencies (objectives) → categories → indicators, and derives pairwise comparisons based on the Saaty 1-9 scale. At a given level, experts provide a matrix of intercomparisons for a set of n criteria,

$$A = \left(a_{\{j\ \}}\right) \text{ with } a_{ij} > 0,\ a_{\{i\ \}} = 1, \text{ and } a_{\{j\ \}} = 1/a_{\{j\ \}} \tag{1}$$

where $a_{\{j\ \}}$ is the judged importance of i over j. The priority vector w is obtained from the principal right eigenvector of A (or the equivalent geometric-mean approximation) and normalized so that $\sum_i w_i = 1$. Judgment quality is screened via the Consistency Ratio;

$$CR = \frac{CI}{RI}, \text{ where } CI = \left(\lambda_{\{m\ ax\}} - n\right)/(n-1) \tag{2}$$

and RI is the Random Index for size n; a common acceptability guideline is CR≤0.10, with revision requested otherwise [14, 15]. For group AHP, individual pairwise matrices are typically aggregated by the geometric mean entry-wise before extracting the priority vector, preserving reciprocity and scale invariance [14, 15].

In this study, AHP is used to derive weights between categories and weights between indicators within categories.

3. Conceptual Model

This section presents the conceptual model for capability assessment for Korea's national domain-specific HPC centers. The model is organized around six evaluation categories (Foundation, Input, Process, Output, Outcome, and Policy Alignment) and links them to a policy-aware methodological pipeline.

Figure 1 summarizes the overall policy-linked capability assessment framework, connecting the theoretical foundation, the six evaluation categories, the methodological steps, and the assessment outputs.

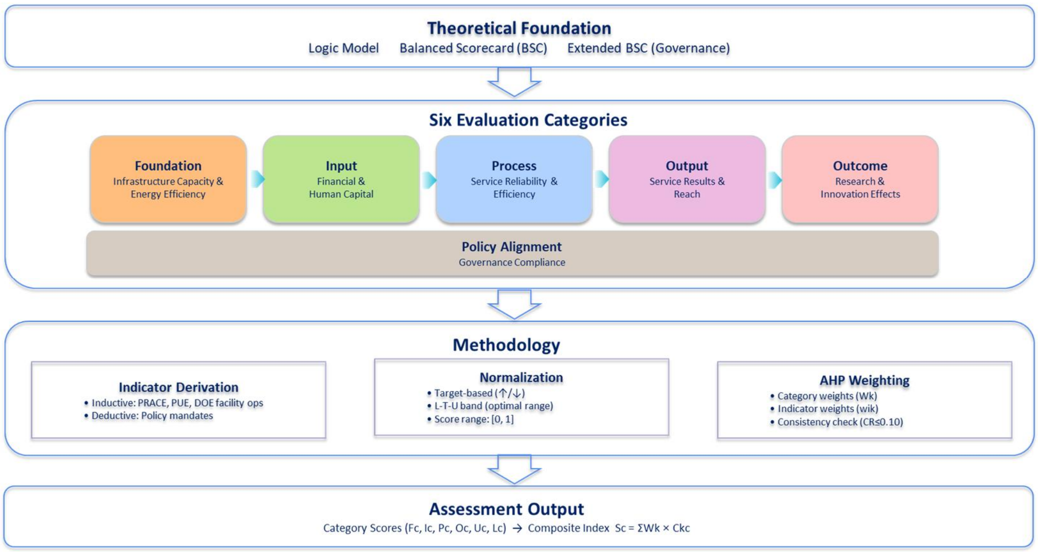

Figure 1. Conceptual Model Architecture

Based on the framework in Figure 1, this study designs an evaluation model according to the following five-step procedure:

1. Setting up categories for evaluation
2. Deriving key indicators through exploratory methods
3. Establishing a normalization framework for each field
4. Deriving weights between categories and indicators
5. Comparison of results between centers

3.1 Category architecture

Guided by the logic model and an extended BSC (governance lens added), we define six evaluation categories that preserve causal staging and strategic coverage: Foundation (infrastructure capacity & energy efficiency), Input (financial and human-capital readiness), Process (service reliability & efficiency), Output (near-term service results), Outcome (research/innovation effects), and Policy Alignment (governance compliance) [4-6], [8]. This mapping avoids siloed metrics, clarifies each indicator's position in the inputs→activities→outputs→outcomes chain, and makes statutory requirements first-class rather than implicit.

Foundation aggregates infrastructure capacity and energy efficiency (e.g., peak performance, storage, backbone, PUE); Input captures financial and human-capital readiness (OPEX per core-hour, electricity-cost share, staffing sufficiency, training budget); Process measures service reliability/efficiency (uptime, queue wait, allocation cycle, shared-utilization links); Output records near-term service results (active users, supported projects, allocated compute hours, industry collaboration); Outcome tracks research/innovation effects (publications, software releases, patents, technology transfers); and Policy Alignment scores governance compliance (shared utilization, standardized reporting, master-plan milestones, audit artifacts). This cross-walk clarifies placement (where each indicator sits in the causal chain and which strategic lens it informs), reduces overlap, specifies evidence requirements, and aligns the category set with national policy.

Table 1. Categories for evaluation

Category	Definition
Foundation	Infrastructure capacity and facility-level efficiency that enable all higher-order activity (capability "stock," e.g., compute/storage/network efficiency at the facility level).
Input	Period resources and organizational readiness—financial and human capital—that condition operational quality and service delivery.
Process	Quality and efficiency of operational service delivery as experienced by users (availability, timeliness, throughput, fairness).
Output	Immediate, countable products of service delivered within the period (scope and reach of services provided).
Outcome	Downstream research and innovation effects plausibly attributable to service use (knowledge and technology results beyond service volumes)
Policy Alignment	Explicit conformance to national mandates captured as scorable endpoints (e.g., shared utilization, standardized reporting, master-plan milestones, audit artifacts).

3.2 Indicator set

We inductively compile recurrent, auditable indicators from community practice—PRACE usage/engagement KPIs, PUE & standardized power-measurement methods, and HPC facility-operations metrics (uptime, queue, allocation cycle, training/support)—because they are routinely collected and comparable across sites [9-12]. We then deductively check coverage against the logic-model × BSC frame and policy mandates, adding missing but theoretically warranted items (e.g., cost/resource inputs, staffing sufficiency, policy endpoints such as shared utilization and standardized reporting) so that the final list is both practice-grounded and frame-complete.

Key indicators extracted from related studies are mapped to six evaluation categories as follows:

Table 2. Key indicators

Category	Indicator families (from literature)	PRACE KPI [9]	Green500 / PUE [10, 11]	DOE facility ops [12]	Key Indicators (in this study)
Foundation	Compute capacity / infrastructure; Energy efficiency	Capacity statistics (system class, availability)	PUE; standardized power measurement	Facility capabilities (context); energy stewardship emphasis	Peak PF (↑); Storage PB (↑); Backbone Gb/s (↑); PUE (↓)
Input	Cost & resource inputs; Staffing sufficiency; Training enablement (budget)	Training/support staff context (as available)	-	Budgetary/ops planning; human resources for operations	Electricity-cost share (↓); OPEX per core-h (↓); Staffing sufficiency (↑); Training budget share (↑)
Process	Reliability; Process efficiency (queue / allocation); Shared-utilization interconnects (operational)	-	-	Uptime targets & reporting; Queue wait; Allocation cycle time; Ops connectivity	Uptime % (↑); Queue wait (↓); Allocation cycle (↓); Interconnect availability (↑)
Output	Usage / access; Training participants (delivered);	Supported projects; Allocated hours; Active users;	-	User program oversight; Training delivery;	Active users (↑); Supported projects (↑); CPU/GPU-normalized allocated compute

	Industry collaboration	Training events & participants; Industry/SME links		Partnership indicators	hours (↑); Training participants (↑)*; Collaboration events/MoUs (↑)
Outcome	Research outputs; Innovation / technology transfer	Publications; Software; Data products; Tech transfer	-	Outcome tracking (context)	Publications (↑); Software releases (↑); Patents (↑); Technology transfers (↑)

.

3.3 Normalization Framework (Field-Specific Targets and Directions)

Once each indicator value is received from each center, standardization is performed for each indicator. Target values for each indicator are set according to each center's field, ensuring that the characteristics of each field are reflected.

We transform each raw indicator value $x_{i,c}$(indicator i for center c) into a unit-free score $s_{I,c} \in [0,1]$ using field-specific targets and directionality. Here we allow for the evaluation and comparison of heterogeneous centers and equipment, and we make the comparison easy on a scale of 0-1.

For each indicator i, we define a target $T_i > 0$ and a direction flag $\sigma_i \in \{+1, -1\}$ where $\sigma_i = +1$ means higher is better and $\sigma_i = -1$ means lower is better. The default normalization is ratio:

$$s_{i,c} = \min\left(1, \left(\frac{x_{i,c}}{T_i}\right)^{\sigma_i}\right) \in [0,1] \tag{3}$$

If $\sigma_i = +1(\uparrow): s_{i,c} = \min(1, \frac{x_{i,c}}{T_i})$

If $\sigma_i = -1(\downarrow): s_{i,c} = \min(1, \frac{T_i}{x_{i,c}})$

This preserves order(monotonicity), caps over-performance at 1(prevents single metrics from dominating).

For rare indicators where neither larger nor smaller is intrinsically better (an optimum band is desired), we reserve the L–T–U (lower–target–upper) normalization for indicators whose desirable performance lies within a bounded operating range rather than increasing or decreasing monotonically. Values substantially below the target indicate under-provision or under-utilization, whereas values substantially above the target introduce congestion, risk, or the crowding-out of other priorities. Let T_I denote the nominal target for indicator i, and let $L_i < T_i < U_I$ be the lower and upper tolerance bounds that delimit the acceptable band. We then compute a closeness-to-target score $s_{I,c}$ for center c that attains 1 at T_i and declines linearly toward 0 at the tolerance limits;

$$s_{i,c} = (0,1 - |x_{\dot{c}} - T_i|/\max\{U_i - T_i, T_i - L_i\}), L_i < T_i < U_i \tag{4}$$

The banded form is applied sparingly—most indicators in this instrument are strictly monotone and are normalized with the default target-based ratio—yet it is essential where both extremes are operationally undesirable and policy seeks stability around a target.

For pass/fail mandates, we use an indicator function or proportion:

$$s_{i,c} = 1\{x_{i,c} \geq T_i\} \tag{5}$$

3.4 Weighting of categories and indicators

This section describes how to assign weights between categories and between indicators within each category. As discussed above, weighting is done using the AHP method.

We derive both the category weights W_k and the within-category indicator weights $w_{i|k}$ using the Analytic Hierarchy Process (AHP) with expert pairwise comparisons; individual matrices are aggregated by the entry-wise geometric mean, and weights are accepted only when the consistency ratio satisfies $CR \leq 0.10$ [14, 15]. For center c, the category score and composite index are

$$C_{kc} = \sum_{i \in k} w_{i|k} s_{\dot{c}} \tag{6}$$

$$S_c = \sum_k W_k C_{kc} = \sum_k \sum_{i \in k} (W_k w_{i|k}) s_{\dot{c}} \tag{7}$$

The above contents are organized by indicator as shown in the table below.

Table 3. Normalization summary

Category	**Indicator (↑/↓)**	**Target T**	**Normalization formula**
Foundation	Peak PF (↑)	T_{PF}	$s = \min(1, \frac{x_{PFc}}{T_{PF}})$
	Storage PB (↑)	T_{ST}	$s = \min(1, \frac{x_{STc}}{T_{ST}})$
	Backbone Gb/s (↑)	T_{BW}	$s = \min(1, \frac{x_{BW\,c}}{T_{BW}})$
	PUE (↓)	T_{PUE}	$s = \min(1, \frac{T_{PUE}}{x_{PUEc}})$
	Category score	-	$F_c = \sum_j w_{j\|F}{}^{s_{i_jc}}$
Input	Elec-cost share (↓)	T_{EC}	$s = \min(1, \frac{T_{EC}}{x_{ECc}})$
	Staffing sufficiency (↑)	T_{STF}	$s = \min(1, \frac{x_{STFc}}{T_{STF}})$
	OPEX per core-h (↓)	T_{OPEX}	$s = \min(1, \frac{T_{OPEX}}{x_{OPEXc}})$
	Training budget share (↑)	T_{TB}	$s = \min(1, \frac{x_{TBc}}{T_{TB}})$
	Category score		$I_c = \sum_j w_{j\|I}{}^{s_{i_jc}}$
Process	Uptime % (↑)	T_{UP}	$s = \min(1, \frac{x_{UPc}}{T_{UP}})$
	Queue wait (↓)	T_{QW}	$s = \min(1, \frac{T_{QW}}{x_{QW\,c}})$
	Allocation cycle (↓)	T_{AC}	$s = \min(1, \frac{T_{AC}}{x_{ACc}})$
	Shared-utilization links (↑)	T_{SU}	$s = \min(1, \frac{x_{SUc}}{T_{SU}})$
	Category score		$P_c = \sum_j w_{j\|P}{}^{s_{i_jc}}$
Output	Active users (↑)	T_{AU}	$s = \min(1, \frac{x_{AUc}}{T_{AU}})$
	Supported projects (↑)	T_{SP}	$s = \min(1, \frac{x_{SPc}}{T_{SP}})$
	Allocated compute hours (↑)	T_{CH}	$s = \min(1, \frac{x_{CH\,c}}{T_{CH}})$
	Industry collaborations (↑)	T_{IC}	$s = \min(1, \frac{x_{ICc}}{T_{IC}})$
	Category score		$O_c = \sum_j w_{j\|O}{}^{s_{i_jc}}$
Outcome	Publications (↑)	T_{PUB}	$s = \min(1, \frac{x_{PUBc}}{T_{PUB}})$
	Patents (↑)	T_{PAT}	$s = \min(1, \frac{x_{PATc}}{T_{PAT}})$
	Software releases (↑)	T_{SW}	$s = \min(1, \frac{x_{SW\,c}}{T_{SW}})$

	Tech transfers (↑)	T_{TT}	$s = \min(1, \frac{x_{TTc}}{T_{TT}})$
	Category score		$U_c = \sum_j w_{j\|U} s_{i_j c}$
Policy	Shared utilization (↑)	T_{SUt}	$s = \min(1, \frac{x_{SUtc}}{T_{SUt}})$
	Standardized reporting (↑)	T_{SR}	$s = \min(1, \frac{x_{SRc}}{T_{SR}})$
	Master-plan milestones (↑)	T_{MP}	$s = \min(1, \frac{x_{MPc}}{T_{MP}})$
	Audit artifacts (↑)	T_{AA}	$s = \min(1, \frac{x_{AAc}}{T_{AA}})$
	Category score		$L_c = \sum_j w_{j\|L} s_{i_j c}$

Note: x_{i} *=observed value;* T_i*=target/benchmark;* s_{i} *=normalized score;* w_{k} *=indicator weight within category* k*(AHP);* W_k*=category weight(*$\sum_k W_k$*=1);* F_c*,*I_c*,*P_c*,*O_c*,*U_c*,*L_c*=category scores;* S_c*=composite capability score*

3.5 Comparison of results between centers

The six category scores (F_c,I_c,P_c,O_c,U_c,L_c) and the composite index (S_c) enable systematic comparisons across centers. We propose two primary visualization methods.

First, a radar chart (Figure 2) plots each center's category scores on six axes, allowing stakeholders to quickly identify relative strengths and weaknesses. Centers with balanced profiles appear as regular hexagons, whereas uneven profiles reveal specific categories that require targeted improvement.

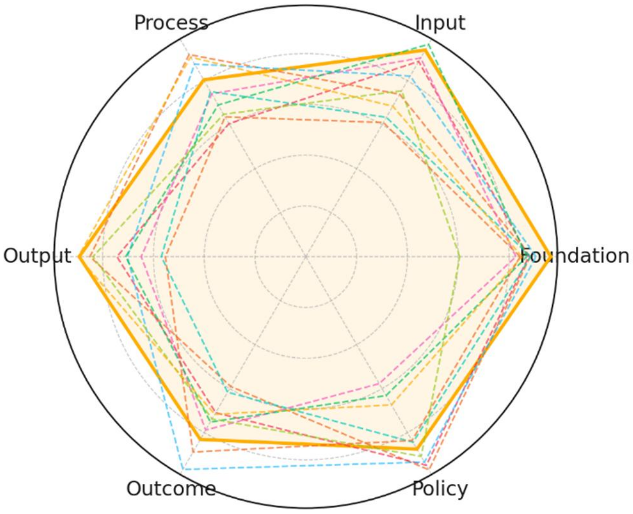

Figure 2. Example comparison of centers across the six evaluation categories

Second, a rank-order table lists centers by the composite index S_c, providing a transparent, single-number benchmark. This ranking supports governance decisions such as resource allocation, priority setting, and performance-based incentives.

Both methods avoid disclosure of raw indicator values, protecting operational confidentiality while maintaining comparability. The normalized [0, 1] scale ensures that scores are interpretable regardless of field-specific target differences, and the AHP-derived weights guarantee that the composite reflects expert-validated priorities. Together, these outputs support transparent benchmarking, targeted improvement planning, and evidence-based policy discussion across Korea's national HPC ecosystem.

4. Conclusions

This study designs a reproducible capability-assessment model for Korea's domain-specific HPC centers. Building on the logic model and an extended Balanced Scorecard, the framework organizes six categories (Foundation, Input, Process, Output, Outcome, and Policy Alignment) and specifies indicator-level targets, direction-aware normalization, and AHP-based weights, thereby providing a transparent pipeline from indicator → category → composite index. The contributions are as follows.

International comparability: by aligning the indicator families with widely used international practices, the model enables cross-national reading of domestic results.

Policy contextualization: a dedicated Policy Alignment category renders statutory requirements (e.g., shared utilization, standardized reporting, master-plan milestones) visible and scorable alongside technical and service metrics.

Field-specific targets: directionality (↑/↓) and, where appropriate, L–T–U bands allow targets to reflect disciplinary conditions, yielding fair comparisons on a 0–1 scale.

AHP-based integration and salience: AHP derives weights at both levels (across categories and within categories), with consistency screening, so that the composite is traceable and the relative importance of categories and indicators are explicit.

Taken together, the design converts KPIs that were previously reported as descriptive lists into a re-computable, audit-ready capability index—one that supports benchmarking, performance tracking, and governance use.

The study has limitations. We have not yet conducted a large-scale empirical application with operational data. Future work will therefore: (i) collect center-level observations and test indicator validity and reliability; (ii) calibrate targets and any L–T–U bands by field; (iii) broaden the AHP panel and conduct sensitivity analyses; and (iv) use the composite index to diagnose cross-center capability and identify priority areas for improvement. These steps will mature the instrument into a decision-support tool for policy execution and strategic management.

Acknowledgments: This research was supported by Korea Institute of Science and Technology Information(KISTI).(No.(KISTI) K25L2M1C1)

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Korea National Law Information Center, Act on the Utilization and Promotion of National Supercomputing, 2011.
[2] Relevant Ministries Joint Committee, National Supercomputing Innovation Strategy for Quantum Leap in the Fourth Industrial Revolution (2021–2030), Ministry of Science and ICT, Republic of Korea, 2021.
[3] M. J. Koh, J. G. Hahm, and Y. H. Jung, "A Conceptual Model for Assessing the Capabilities of National Supercomputing Centers in Korea," in Proc. ICCC 2024—International Conference on Convergence Content (ICCC 2024), Danang, Vietnam, Dec. 17-19, pp. 281–282, 2024.
[4] W.K. Kellogg Foundation, Logic Model Development Guide, 2004.
[5] R. S. Kaplan and D. P. Norton, "The Balanced Scorecard—Measures that Drive Performance," Harvard Business Review, vol. 70, no. 1, pp. 71–79, 1992.
[6] S. Kumar, W. M. Lim, R. Sureka, C. J. Chiappetta Jabbour, and U. Bamel, "Balanced scorecard: trends, developments, and future directions," Review of Managerial Science, vol. 18, no. 8, pp. 2397–2439, 2024, doi: https://doi.org/10.1007/s11846-023-00700-6.
[7] J. C. Panitz, M. Wiener, and M. Amberg, "A Balanced Scorecard for Compliance—Requirements of a Comprehensive Compliance-Reporting," In AMCIS 2010 Proceedings, p. 160, 2010.
[8] R. S. Kaplan and M. E. Nagel, Improving Corporate Governance with the Balanced Scorecard, Harvard Business School Working Paper 04-044, 2003/2004.
[9] PRACE (Partnership for Advanced Computing in Europe), Statistics & KPIs / Impact materials (various releases), 2021–.
[10] Green500 & Energy Efficient HPC Working Group, Energy-Efficient High-Performance Computing Power Measurement Methodology, 2014.
[11] Energy Efficient HPC Working Group, Energy Efficiency Considerations for HPC Procurement Document, Dec. 17, 2019.

[12] U.S. Department of Energy, Office of Science (ASCR). 2023 Oak Ridge Leadership Computing Facility (OLCF) Operational Assessment Report, 2024.

[13] T. L. Saaty, The Analytic Hierarchy Process, McGraw-Hill, 1980.

[14] E. H. Forman and S. I. Gass, “The Analytic Hierarchy Process—An Exposition,” Operations Research, vol. 49, no. 4, pp. 469–486, 2001.

[15] O. S. Vaidya and S. Kumar, “Analytic hierarchy process: An overview of applications,” European Journal of Operational Research, vol. 169, no. 1, pp. 1–29, 2006.