DEGEVAL … INFO

Zeitschrift für Evaluation | 25. Jahrgang | 2026
Heft 1 | S. 163–167 | https://doi.org/10.31244/zfe.2026.01.10 |
© 2026 Waxmann

Institutionalisierung und Professionalisierung von Evaluation

Bericht zur 28. Jahrestagung der DeGEval – Gesellschaft für Evaluation e. V. vom 17. bis 19. September 2025, Universität des Saarlandes, Saarbrücken

Hamadou Ngoei[1]*, Bouba Abdouramane*[2]

Einleitendes

Die 28. Jahrestagung der DeGEval – Gesellschaft für Evaluation e. V. fand vom 17. bis 19. September 2025 an der Universität des Saarlandes in Saarbrücken statt und vereinte Expert:innen aus Wissenschaft, Praxis und Politik aus zahlreichen Ländern. Unter dem Leitthema: „Institutionalisierung und Professionalisierung von Evaluation" – bot die Tagung nicht nur ein Forum für den fachlichen Austausch, sondern auch ein breites Spektrum an Weiterbildungsangeboten – von praxisorientiertem Workshop bis hin zu methodischen Vertiefungen. Die internationale Zusammensetzung der Teilnehmenden und die Vielfalt der Beiträge spiegelten die globale Relevanz des Themas wider. Evaluation ist nicht nur ein rein nationales Anliegen, sondern ein zentraler Bestandteil evidenzbasierter Steuerung in unterschiedlichsten gesellschaftlichen Kontexten weltweit. Besonders hervorzuheben ist das Treffen der MABLE- und MEval-Alumni und Alumnae, das den interkulturellen Dialog und die Vernetzung von Fachkräften aus verschiedenen Regionen förderte. Neben einer Keynote zu globalen Trends und makrosoziologischen Analysen der Professionalisierung von Evaluation wurden unter anderem Sessions zu den öffentlichen und privaten Sektoren, zu Institutionen und Evaluationsmethoden, zur Entwicklung nationaler Evaluationskapazitäten, zur Integration von KI in Evaluationsprozesse sowie zur Rolle der Evaluation in den Bereichen Bildung, Kultur und Entwicklungszusammenarbeit angeboten. Die Tagung verdeutlichte, dass die Evaluation längst eine globale Disziplin ist, die nicht nur technische Expertise, sondern auch kulturelle Sensibilität und strategisches Denken erfordert. Die Veranstaltung setzte Impulse für die Weiterentwicklung von Qualitätsstandards, ethischen Leitlinien und Qualifizierungsstrategien und eröffnete

1 Hochschule der Bildenden Künste Saar, Saarbrücken
2 Action Medeor, Tönisvorst

Perspektiven, wie Evaluation in Zeiten globaler Herausforderungen als Instrument für Lernen, Transparenz und Entwicklung genutzt werden kann. Dieser Text ist ein persönlicher Bericht der Autoren auf Basis der von ihnen besuchten Sessions.

Makrosoziologische Trends und der Status der Profession im Evaluationsbereich

In ihrer makrosoziologischen Analyse untersuchten Reinhard Stockmann und Wolfgang Meyer den globalen Institutionalisierungsgrad der Evaluation in 50 Ländern. Die Autoren stellten fest, dass die Etablierung von Evaluationen primär eine ‚policy'-getriebene Entwicklung darstellt, die maßgeblich durch das politische System sowie externe Faktoren wie Geberrichtlinien oder das New Public Management vorangetrieben wird. Während sich das professionelle System, das sich durch internationale Vernetzung und Ausbildungsstrukturen auszeichnet, relativ eigenständig entwickelt, bleibt die Verankerung im sozialen System marginal. Zivilgesellschaftliche Akteure treten kaum als treibende Kraft auf und der öffentliche Diskurs über Evaluation ist schwach ausgeprägt. Die Studie belegte zudem eine Korrelation zwischen der Qualität der Regierungsführung (Good Governance) und demokratischen Strukturen einerseits sowie der tatsächlichen Nutzung von Evaluationsergebnissen andererseits, warnt jedoch vor einer reinen Bürokratisierung ohne strategische Nutzung. Komplementär dazu hinterfragt Stefanie Krapp den aktuellen Status der Professionalisierung kritisch und konstatiert eine Diskrepanz zwischen Anspruch und Wirklichkeit. Sie diagnostiziert eine Stagnation der Fachgesellschaften (VOPEs) sowie eine Fragmentierung globaler Bemühungen, begleitet von einem bedenklichen Rückbau akademischer Ausbildungsangebote an Universitäten. Dies führe zu einer ‚Competency Training Gap'. Ferner argumentiert Krapp, dass die derzeitige technische Professionalisierung angesichts globaler ‚Polykrisen' unzureichend ist. Sie plädiert stattdessen für eine transformative Neuausrichtung: Evaluator:innen müssten ihre Rolle von reinen Methodiker:innen zu wissenschaftlich fundierten, ethisch orientierten und systemisch denkenden Akteuren erweitern. Nur so könnten sie Entscheidungsprozesse in komplexen Kontexten wirksam unterstützen und die gesellschaftliche Relevanz der Disziplin durch eine klare professionelle Identität stärken.

Methodische Pluralität: zwischen Rigorosität und technologischer Innovation

Ein zentraler Diskursstrang der Tagung betraf die methodische Qualitätssicherung. Judith Ihl und Carolin Wicke plädierten angesichts der Informationsflut für die stärkere Nutzung systematischer Synthesemethoden (Systematic Reviews), um die Validität von Aussagen zu erhöhen. Dass auch quasi-experimentelle Designs wie das Regression-Discontinuity-Design (RDD) belastbare kausale Aussagen ermöglichen, demonstrierte Christoph E. Müller am Beispiel des EXIST-Gründungsstipendiums, wobei er die ethischen Vorteile gegenüber randomisierten Kontrollstudien hervorhob.

Die Grenzen methodischer Rigorosität in fragilen Kontexten zeigten Nico Herforth und sein Team am Beispiel einer Agrarstudie in Uganda auf, wo soziale Barrieren und Kontaminationseffekte pragmatische Anpassungen erforderten. Eine gänzlich neue Dynamik bringt die Künstliche Intelligenz (KI) in das Feld. Armen Grigorian machte deutlich, dass KI die Effizienz von Datenerhebung und -analyse revolutionieren kann, jedoch das Risiko intransparenter ‚Black-Box'-Algorithmen birgt. Der Konsens war, dass Evaluator:innen künftig stärker als strategische Interpret:innen fungieren müssen, die die ethische Integrität KI-generierter Daten sicherstellen. Christoph Glauser ergänzte das methodische Spektrum innovativ durch die Analyse digitaler Suchvolumina (‚Active Search') als Indikator für gesellschaftliche Relevanz.

Institutionalisierung im öffentlichen Dienst und in der Politik

Der Transfer von Evaluation in die administrative Praxis erweist sich als vielschichtiger Prozess, der oft an kulturellen Hürden scheitert. Vivien Benert und Carolina Gironés-Fraile identifizierten in der deutschen Ministerialverwaltung mittels des „E-Valuate Wirkungsindex" starre Hierarchien und eine fehlende Fehlerkultur als Hemmnisse für einen lernenden Staat. Ähnliche Herausforderungen skizzierte Ingo Dungs für die Polizei NRW, wo die Balance zwischen wissenschaftlichem Anspruch und operativer Pragmatik gefunden werden muss. Dass fehlende Evaluation zu fragmentierten Politikzyklen führen kann, illustrierte Angela Wroblewski am Beispiel der europäischen Gleichstellungspolitik. Ohne systematisches Monitoring verbleiben politische Bekenntnisse oft auf symbolischer Ebene. Um dieser Lücke methodisch zu begegnen, stellten Matthias Klapproth, Susanne Johanna Väth und Hansjörg Gaus das FIV-Modell des Bundesamts für Naturschutz vor, das Forschung, Integration und Verwertung verknüpft, um wissenschaftliche Erkenntnisse effektiv in Gesetzgebungsverfahren zu überführen. Auch Christelle Nowack betonte im Kontext der Weiterentwicklung der DeGEval-Empfehlungen für Auftraggebende die Notwendigkeit niederschwelliger, aber standardbasierter Leitfäden für Auftraggebende zu haben.

Bildung, Kultur und Gesundheit

In den Sektoren Bildung und Kultur zeigt sich ein Spannungsfeld zwischen Autonomie und Rechenschaftspflicht. Thomas Renz und Thomas Heskia diskutierten die Institutionalisierung im Kultursektor unter Rückgriff auf die kritische Theorie Adornos, wobei sie die Gefahr hervorheben, dass quantitative Indikatoren die ästhetische Autonomie untergraben könnten. Im Schulbereich wiesen Dirk Groß und Ralf Schattschneider darauf hin, dass Standards allein nicht genügen; vielmehr müsse Evaluation als soziale Intervention verstanden werden, die ‚Sensemaking' bei den Akteur:innen ermöglicht. Sarah Schmidt mahnte im Hochschulbereich eine stärkere Verankerung evaluativer Kompetenzen in den Curricula an, um dem Fachkräftemangel entgegenzuwirken. Aus der Perspektive der Schweizerischen Evaluationsgesellschaft (SEVAL) erörterten Tanja Schnoz und ihre Kolleg:innen die Professionalisierung innerhalb des

Hochschulbereichs selbst und unterstrichen die Notwendigkeit standardisierter Kompetenzprofile sowie den Austausch über Methoden wie den ‚Teaching Analysis Poll'. Manuela Pötschke thematisierte die begleitende Evaluation von Schulkonzepten und verdeutlichte, dass für eine nachhaltige Qualitätsentwicklung eine dauerhafte, wissenschaftlich fundierte Begleitung sowie eine Entlastung des Lehrpersonals unabdingbar sind. Réné Krempkow, Kerstin Janson und Julia Rathke widmeten sich der Professionalisierung von Evaluation im Wissenschaftsmanagement und analysierten die damit verbundenen Kompetenzprofile. Auf Basis empirischer Erhebungen differenzierten sie zwischen allgemeinen Wissenschaftsmanager:innen und spezialisierten Evaluator:innen. Letztere nehmen höhere Anforderungen in generischen und systemischen Kompetenzbereichen wie Problemlösung und Zusammenhangserkennung wahr. Gleichzeitig identifizierten die Autor:innen Unsicherheiten im Bereich digitaler Kompetenzen.

Im Gesundheitswesen forderten Ludwig Grillich und sein Team sowie Kerstin Schwarz und Jörg Rech robustere Designs, um Designfehler von Umsetzungsdefiziten zu unterscheiden. Hier wird Evaluation nicht nur als Qualitätskontrolle, sondern zunehmend als Instrument strategischer Steuerung verstanden, wobei die Datenverfügbarkeit eine konstante Herausforderung bleibt.

Globale Perspektiven und Entwicklungszusammenarbeit

Die internationale Dimension der Tagung beleuchtete, dass Institutionalisierung kontextabhängig verläuft. Während Rebecca Wagner und ihre Kolleg:innen den EU-Beitrittsprozess in der Ukraine und Moldawien als Reformmotor identifizierten, verwiesen Hamadou Ngoei und Abdouramane Bouba auf die Notwendigkeit einer ‚Evaluation Made in Africa', die westliche Wissenschaftsmodelle durch lokale Philosophien wie Ubuntu ergänzt. Sarah Klier (DEval) stellte mit dem INCE-Index ein Instrument vor, das nationale Evaluierungskapazitäten global vergleichbar macht und Peer-Learning fördert. Susanne Johanna Väth und Janis Wicke sowie Katharina Kreutz und Sebastian Schuster (Welthungerhilfe) betonten, dass in der Entwicklungszusammenarbeit (EZ) flexible Antragsphasen und eine Trennung von Durchführung und Nutzung essenziell sind, um eine echte Lernkultur zu etablieren. Brigitte Ecker und Mario Steyer beschrieben am Beispiel des FH Campus Wien, wie Evaluation als Instrument des organisatorischen Lernens und Wachstums dienen können, wenngleich die Herausforderung besteht, fragmentierte Instrumente zu einem systemischen Ganzen zu verknüpfen. Jörg Faust machte begreiflich, dass Evaluierung in der EZ stärker verankert ist, als in vielen anderen Politikfeldern. Gründe dafür sind die Debatten um Wirksamkeit und Legitimität, die Komplexität externer Interventionen sowie der hohe Bedarf an Rechenschaft und Lernen.

Nachhaltigkeit als multidimensionales Paradigma

Abschließend widmete sich die Tagung der Operationalisierung von Nachhaltigkeit. Die Ad-hoc-Gruppe Nachhaltigkeit sowie Beiträge von Reinhard Zweidler und Kers-

tin Schwarz kritisierten die oft eindimensionale ökologische Betrachtung. Gefordert wurde ein holistischer Ansatz, der soziale Aspekte und gesundheitliche Folgekosten (‚Health in All Policies') integriert. Christiane Kerlen und Christina Bantle zeigten auf, dass dies methodisch durch partizipative Ansätze im Sozialraum und differenzierte Förderkriterien unterstützt werden kann.

Konklusion und Learned Lessons

Die 28. Jahrestagung der DeGEval verdeutlichte, dass sich die Evaluation in einer entscheidenden Transformationsphase befindet. Aus den vielfältigen Beiträgen lassen sich folgende zentrale Erkenntnisse ableiten:

1. *Institutionalisierung ohne Professionalisierung ist riskant:* Es existiert eine Diskrepanz zwischen der politischen Implementierung von Evaluationssystemen und der Verfügbarkeit qualifizierten Personals. Ohne eine Stärkung akademischer Ausbildungsstrukturen droht die Evaluation zur reinen Verwaltungsroutine zu verkommen.
2. *Kulturwandel vor Methodik:* Erfolgreiche Evaluation in Verwaltung und Organisationen hängt weniger von perfekten Methoden ab als von einer gelebten Fehlerkultur und der Bereitschaft zum organisationalen Lernen (‚Sensemaking').
3. *Technologie als Gamechanger:* Der Einsatz von KI und Big Data (z. B. Active Search) erweitert das methodische Repertoire massiv, erfordert jedoch eine neue ethische Kompetenz und ein Rollenverständnis von Evaluator:innen als strategische Einordner:innen von Evidenz.
4. *Kontextsensibilität:* Ob in der EZ, im Kultursektor oder in der Schule – Standards sind notwendig, müssen aber an die spezifischen sozialen und kulturellen Logiken der Felder adaptiert werden, um Akzeptanz und Wirksamkeit zu entfalten.

Die diesjährige Tagung sendete somit das klare Signal, dass die Zukunft der Evaluation in der Verbindung von methodischer Rigorosität mit systemischem Denken und ethischer Verantwortung liegt.

Hamadou Ngoei | Hochschule der Bildenden Künste Saar | Stabstelle QM und Akkreditierung | Keplerstraße 3–5, D-66117 Saarbrücken | E-Mail: h.ngoei@hbksaar.org

Bouba Abdouramane | Action Medeor | Programmadministration – Frankophones Afrika | St. Töniser Str. 21, D-47918 Tönisvorst |
E-Mail: bouba.abdouramane@medeor.de