Diversity, Equity and Inclusion Activities in Database Conferences: A 2022 Report

`https://dbdni.github.io`

Sihem Amer-Yahia, Divyakant Agrawal, Yael Amsterdamer, Sourav S. Bhowmick, Angela Bonifati, Renata Borovica-Gajic, Jesús Camacho-Rodríguez, Barbara Catania, Panos K. Chrysanthis, Carlo Curino, Jérôme Darmont, Gillian Dobbie, Amr El Abbadi, Avrilia Floratou, Juliana Freire, Alekh Jindal, Vana Kalogeraki, Sujaya Maiyya, Alexandra Meliou, Madhulika Mohanty, Behrooz Omidvar-Tehrani, Fatma Özcan, Liat Peterfreund, Wenny Rahayu, Shazia Sadiq, Sana Sellami, Utku Sirin, Wang-Chiew Tan, Bhavani Thuraisingham, Yuanyuan Tian, Pınar Tözün, Genoveva Vargas-Solar, Neeraja Yadwadkar, Victor Zakhary, Meihui Zhang

1. THE DEI@DB INITIATIVE

The Diversity, Equity and Inclusion (DEI) initiative started as the Diversity/Inclusion initiative in 2020 [4]. The current report summarizes our activities in 2022.

Our responsibility as a community is to ensure that attendees of DB conferences feel included, irrespective of their scientific perspective and personal background. One of the first steps was to establish the role of the DEI chairs at DB Conferences, with the DEI team dedicated to providing leadership to help our community achieve this goal. In this leadership role, the DEI team is advising DEI chairs at DB conferences, serving as a memory of DEI events at conferences, building an agreed-upon vision, and committing to working together to devise a set of measures for achieving DEI. That is pursued via actions led by our core members (Figure 1) and liaisons of individual executive bodies (Figure 2): **REACH OUT** collects data and experiences from our community. **INCLUDE** monitors and recommends inclusion efforts. **ORGANIZE** focuses on in-conference organization efforts, such as adopting a code of conduct. **INFORM** communicates through various channels. **SUPPORT** coordinates DEI support from executive bodies and sponsors. **SCOUT** collates DEI efforts from other communities. **COORDINATE** manages all actions. Two new actions: **MEDIA** preserves and disseminates the digital media produced by DEI@DB events. **ETHICS** establishes and promotes ethics guidelines for publications in our community.

What did we achieve last year? Last year's focus was on raising awareness by introducing special DEI sessions at conferences and collecting statistics and wishes through surveys. Additional actions were also taken at some of our conferences, such as making deliberate efforts to address diversity and gender balance of conference officers, student volunteers, and invited speakers, and drafting and displaying a Code of Conduct. We also published tips on inclusive writing and video captioning and developed and deployed tools that facilitate our DEI efforts.

What did we achieve this year? The response from the DB community in various events in 2021 conferences has
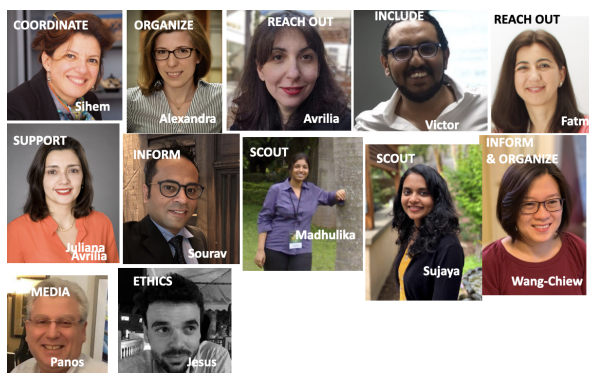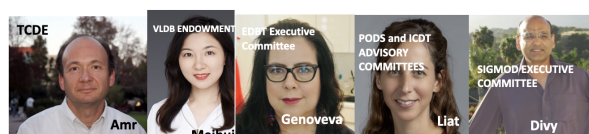

Figure 1: DEI@DB Core Members

Figure 2: DEI@DB Liaisons Members

been extremely positive and experiences shared across DB conferences have already enabled the scaling up of DEI activities in 2022 conferences. Two new actions were introduced this year: **MEDIA** and **ETHICS**, listed above. We worked on a proposal to extend DEI to workshops and journals. We re-branded D&I to DEI and merged DEI and DBCARES. We drafted our bylaws to renew the initiative and introduced a Code of Ethics. We also drafted the job descriptions of all DEI members and of each year's DEI conference chairs.

Our proposal of extending DEI to journals consists of several concrete actions: (1) promote the use of inclusive and bias-free language by providing authors and reviewers inclusivity guidelines; (2) appoint a journal's DEI lead; (3) establish a DEI mentoring scheme for early-career researchers; (4) conduct voluntary surveys to understand the diversity of authors and reviewers.

Figure 3: DEI Chairs of 2022-2023 DB Conferences

To draft our Code of Ethics (CoE) [8], we scouted for Codes of Ethics from larger organizations such as ACM [3]. Additionally, we identified a set of DEI best practices for conference organization from various conferences within and beyond Computer Science [5] and created guidelines for DB conference organizers. Lastly, some of our venues continued to use CLOSET [6] towards fairer reviewer assignment.

DEI statistics. In 2022, we conducted surveys at EDBT, ICDT, SIGMOD, VLDB, ADBIS, MDM, and DEBS. We had 104, 30, 48, 261, 53, and 16 participants, respectively. Among them 46.7% to 75% were males, 18.8% to 50% were females, and 3.77% were LGBTQ+; the majority of the participants are either from academia (20.8-43.3%) or work in both academia and industry (25-43.8%), followed by students (19.2-37.5%), and industry-only participants (2.9-14.6%). Authors came from about 30 countries and submitted papers had an average of 3 authors each, most are based on international collaboration. According to these surveys, people are most interested in research (finding topics, setting goals, determining success) followed by work-life balance and career paths as the top-3 topics they want to discuss at DEI sessions. About half of the survey participants were interested in a mentorship program, either as a mentee, a mentor, or both.

2. DEI@DB CONFERENCES 2022

A common action at DB conferences this year was drafting DEI guidelines and examples of inclusive and discriminatory behavior and language, encouraging authors and participants to consider them when writing their papers and preparing presentations, and when asking/answering questions in sessions. DEI reviewer guidelines were also drafted to spot exclusionary writing in papers and to write constructive and respectful reviews.

DEI@CIDR. The DEI co-chairs urged conference organizers to make sure the social event is inclusive in terms of diet options, beverage choices, and entertainment. They also organized a Women in DB lunch event in-person and online. There was no pre-set topics and anything related women in DB was discussed. They also collected statistics, such as female attendees in person and online, female presenters (including paper presentation, keynotes, panelists, and sponsor talks), female session chairs, and questions from female attendees during Q&A. The statistics will be used to track progress of DEI efforts in CIDR conferences.

DEI@EDBT/ICDT. The DEI chairs of EDBT/ICDT proposed a program under the premise that the first step to promoting DEI is to understand its meaning.

(1) **Talk:** The plenary talk addressed the algorithmic fairness challenges from the perspective of data querying, exploration and analytics. The talk was titled *Algorithmic Fairness for Graphs* delivered by Evaggelia Pitoura, Professor, University of Ioannina, Greece. She discussed different fairness criteria and presented fairness-aware PageRank algorithms that achieve fairness with minimum loss from the original PageRank.

(2) **Panel:** Diversity is a critical step towards achieving an environment with varied perspectives, and inclusion allows those perspectives to be shared, heard, and valued. Whereas diversity is a tangible, quantifiable metric, inclusion requires sophisticated methods for taking the community's pulse and engaging with community members to safely express their unique experiences within the collective culture. Towards this, the panel on *DEI Perspectives around the World* aimed to open borders and explore the meaning and perception of DEI in scientific contexts across different regions. The seven panelists responded to the key question: *What does DEI mean in Europe, Latin America, the Middle East, the north and south of Africa, Asia, Oceania, etc.?*

DEI@ICDE. Bhavani Thuraisigham from the University of Texas at Dallas gave a talk on *Why a Career in Data Science for a Woman + The Role of Mentoring for Career Success.* where she discussed the importance of mentoring for women in Data Science and gave examples of her personal story on how lack of mentoring was initially impacted her career and how she chose mentors who have then helped her thrive in her career. There was a mentoring workshop.

DEI@MDM.

(1) **Keynote:** The first event was a keynote talk on *Analysing Crowds Mobility from Decolonisation and Inclusive Perspectives* given by Genoveva Vargas-Solar, a principal scientist of the French Council of Scientific Research (CNRS) and a member of the DataBase group of the Laboratory on Informatics on Image and Information Systems (LIRIS). In her talk, Dr. Vargas-Solar addressed the following important aspects: the usage of crowd behaviour analytics techniques to model crowds as complex systems and of a decolonization perspective to perform data analysis more inclusively.

(2) **Workshop:** Falaah Arif Khan (Center for Data Science, NYU) organized a Workshop on *Demystifying AI: Comics as a Medium of Scientific Dissemination* revealed the limitations of the scientific method in addressing bias in algorithms. Participants were invited to contribute a creative response to the multi-faceted approach of the 'Data, Responsibly' comics.

(3) **Surveys:** The goal here was to provide a better understanding of our community to develop more ef-

fective strategies. Surveys included two questions: (1) How would you assess the current climate for women and under-represented minorities? (2) Does research in mobile data management & IoT consider DEI aspects?

(4) **DEI award:** To recognize students' efforts, MDM, granted *a Best DEI PhD Forum Presentation* awards, which was established in MDM 2021.

DEI@SIGMOD.

(1) **Keynote:** The DEI chairs opted for the topic of *Inclusive teaching in STEMM* as a keynote. This topic is crucial since without it, we cannot attract a diverse pool of young people to our field. With this motivation, Colleen Lewis from UIUC, who curates a popular website on tips and tricks for effective CS teaching called CSTeachingTips.org, gave a keynote on *Strategies for Creating Inclusive Learning Environments.*

(2) **Panel:** Following the success of panels on *Failures and Impostor Syndrome* at SIGMOD 2021, SIGMOD 2022 held a DEI panel on *Success & Impact Beyond Traditional Metrics.* This discussion served as a continuation across the two years while also pondering on how one can improve DEI by attributing a broader meaning to the terms success and impact.

(3) **Workshop:** The second iteration of the workshop on *Resume Review for Students from Underrepresented Groups* was organized by Victor Zakhary (Oracle). The goal is to reduce the inequalities across students applying for jobs by minimizing potential mistakes one makes while preparing a resume for a target industry job.

DEI@VLDB. Three DEI activities took place at VLDB 2022. The traditional first nation's welcome to country took place at the opening ceremony followed by a keynote from Distinguished Professor Maggie Walter on *Melding Indigenous Data Sovereignty with New Data Technologies*, which highlighted the risks and consequences of data technologies that may result from lack of inclusion of marginalized people, and provided pathways for mitigating these risks towards collective benefits. A working lunch session was organized on day 3 of the conference featuring the CEO of Australian Academy of Technology and Engineering, Kylie Walker. Kylie delivered a powerful message on why *Diversity in computer science is everybody's business.* A mentoring program was also offered to both in-person and virtual attendees. For in-person attendees a Grand Challenges session was organized giving a chance for students and early career researchers to gain first-hand guidance and inspiration on grand challenges and future research opportunities from experienced researchers who attended VLDB. Thirty five people attended this session including fifteen women. A Mentoring Journey program was offered to all, but especially targeted people who could not attend VLDB. The framework of mentorship is based on an 8-meetings' journey. Those present at VLDB were encouraged to have the first meeting in Sydney at the time of the conference. The remaining were connected via emails. Fourteen pairs of mentors and mentees are currently participating in the program.

DEI@ADBIS.

(1) **Actions.** We promoted the Code of Conduct on the first day of the conference. We also promoted diversity in conference and workshop committees, keynote and tutorial speakers, session chairs, notably including younger chairs. The Doctoral Consortium (DC) included a mentoring session and a DEI Panel. A PhD student was involved in local organization and some PhD students and postdocs in the Program Committees of the DC and of the main conference, respectively. We encouraged women and underrepresented communities to submit papers and asked authors of accepted papers to provide a short bio, getting to help the audience know them. Thanks to support from the Department of Control and Computer Engineering and the smart Data center at Politecnico di Torino, we offered grants assigned based on underrepresentation, gender and role/position. One went to a Master's student who contributed as author to an accepted paper.

(2) **DEI panel.** The panel was entitled *Promoting DEI: policies, strategies and future directions in higher education, research communities, and business.* It offered a multi-perspective view of DEI projects in organizations that aim to create environments where people can achieve positive professional and personal goals. We discussed best practices of projects implemented in multicultural organizations, including academia, universities, and research centers (Politecnico di Torino, Higher Education in France and the French CNRS), as well as leading international companies such as Accenture and Nestlé. The panel offered two pieces of advice to participants: 1) Beyond statistics and quantitative assessments of DEI indices, social awareness within an organization is important; 2) Institutions need to continually adapt their DEI agendas and define new goals: improve productivity through well-being, develop new markets, attract outstanding talent to promote meritocracy, empower underrepresented communities.

DEI@DEBS. DEBS joined DEI this year and focused raising the DEBS community's awareness.

(1) **Talk:** Falaah Arif Khan, a PhD student at NYU Center for Data Science delivered a talk titled *It's funny because it's true – confronting scientific catechisms.* Her message was that any work's scope of impact needs to be defined more broadly. Artistic intervention is a promising solution to catechisms in science.

(2) **Debrief:** Sihem Amer-Yahia (CNRS, Univ. Grenoble Alpes) presented a DEI debrief. She invited attendees to get involved, offering ideas and experiences.

(3) **DEI Compliance:** A shepherding DEI committee was formed to review camera-ready papers in research and industry tracks. The committee marked papers according to their DEI compliance in the proceedings.

DEI@WISE. WISE 2022 DEI agenda aimed to continue building on previous year's DEI activities. In particular, as part of our continuous efforts in raising awareness and sharing best practices, we developed a new award category called the Best DEI WISE paper award. To select the winning paper, we proposed a form that asked every author to outline how they have incorporated DEI considerations in their paper, including the use of inclusive languages, inclusive case studies and examples,

gender, unbiased data set, and use of images or graphs for the visually impaired. We received a response rate of 89%. The DEI team reviewed the papers based on the author survey form, evaluated the presentations according to inclusive practices, and awarded two best DEI paper awards. We also continued to maintain a balance of gender representations across the different roles at the conference, particularly keynote speakers and session chairs. Following the feedback from WISE 2021 to invite a keynote speaker who is also a leader and role model in promoting DEI within the computer science community, we invited Sihem Amer-Yahia who presented a talk on *Fairness on Online Labor Markets* and who also presented the DEI intiative.

DEI@SoCC. Ana Klimovic from ETH Zurich gave the first keynote at the conference on *Scalable Input Data Processing for Resource-Efficient Machine Learning*, where she discussed the characteristics of real ML input pipelines from production workloads and the challenges and opportunities that lie ahead in supporting these workloads on disaggregated system architectures. Kunle Olukun from Stanford gave the third keynote, *Systems for ML and ML for Systems: A Virtuous Cycle.* He discussed how ML-optimized systems can be used to improve training accuracy and how ML models can be used to improve upon system design and management, specifically using Reconfigurable Dataflow Accelerators and their use in the SambaNova systems as an example. Neeraja Yadwadkar gave one of the invited talks in the newly established Fresh Thinking track, further discussing opportunities at the intersection of ML and Systems. In addition, almost 60% of the conference organizers were female, and the organizers emphasized diversity in representation in the Session Chairs and the Program Committee. Diversity among the participants was further elevated via the selected poster presenters and the student travel grants.

3. TOWARDS FAIRER REVIEWER ASSIGNMENTS

In 2022, PC chairs of SIGMOD, VLDB, and ICDE requested to leverage CLOSET [6] for managing COI during their review process. Different venues used different variants of CLOSET. SIGMOD 2022 invoked *Mini CLOSET*, a "lighter" version, whereas VLDB and ICDE leveraged the complete version. Mini CLOSET takes as input authors and PC members in a venue (name, affiliation, email). It generates reports on co-authorship relationships between all reviewer-author pairs. SIGMOD 22 PC chairs used the output of Mini CLOSET as a guide to assign reviewers to submissions. In VLDB 2022, CLOSET was often invoked *after* reviewer-paper assignment to check for any COI violations whereas in ICDE 2022 it was invoked *prior* to the assignment.

The frequency of invocation of CLOSET varied across different venues. In SIGMOD 2022, Mini CLOSET was invoked only for the first submission cycle. In VLDB 2022, CLOSET was invoked 11 out of 12 cycles. ICDE 2022 leveraged it for all cycles.

CLOSET detected a significant number of unreported COIs in VLDB 2022 and ICDE 2022. In these venues, the PC chairs manually reassigned reviewers for submissions with COI issues. In practice, this can be onerous. For instance, ICDE 2022 invoked CLOSET *four* times in the second cycle as manual re-assignments gave rise to new COI violations. In particular, this process lasted for 10 days which adversely impacted the schedule of ICDE 2022 review process.

The output of CLOSET can now be ingested by CMT [1] and utilized in its automatic reviewer assignment. This interface was deployed for VLDB and EDBT 2023.

4. GOING FORWARD

Job descriptions. We are working with the ACM to make sure the job descriptions of DEI chairs at conferences and DEI initiative members (core and liaisons) are aligned with the ACM policies.

COIs. Although major venues utilize CLOSET regularly, several others still depend on conference management tools to manage COIs which is inadequate. Furthermore, different venues have different COI policies. We are working towards a consistent COI policy across all DB venues.

Code of Ethics. Going forward, our goal is to keep updating the CoE and conference guidelines by scouting other conferences and advocacy in the DB community. We are looking into organizing committees as in other conferences [1, 2]—particularly those that are targeted towards DEI efforts such as Accessibility Committee, Family Committee, and Sustainability Committee.

MEDIA action. To preserve and disseminate valuable digital media produced in the last two years as well as upcoming material, we have established a `DEI-DB-MEDIA` YouTube channel where we organized the recordings based on the event type (talk, brief, panel discussion and workshop), the conference, and the year.

ETHICS action. We are working on establishing and promoting ethics guidelines for publications in our community, similar to other ongoing efforts [7]. This involves creating a living document that specifies major aspects of ethics that authors and reviewers should consider, e.g., general conduct when dealing with data or potential negative societal impact of a research application, and designing and implementing the ethics review process (possibly multiple different processes depending on the conference format). To enhance inclusion during conference sessions, we also plan to compile a set of guidelines for session chairs, presenters, and participants for handling panels and Q&A. This action will unify author, reviewer, and presenter guidelines that have been written at individual conferences.

---

[1] We thank Wolfgang Lehner for leading the collaboration with CMT.

5. REFERENCES

[1] NeurIPS 2021. 2021 organizing committee.
    `https://nips.cc/Conferences/2021/Committees`.

[2] CHI 2022. Equity, justice, and access
    commitments. `https://chi2022.acm.org/2021/06/23/equity-justice-and-access-commitments`.

[3] ACM. Code of ethics.
    `https://www.acm.org/code-of-ethics`.

[4] Sihem Amer-Yahia, Yael Amsterdamer, Sourav S.
    Bhowmick, Angela Bonifati, Philippe Bonnet,
    Renata Borovica-Gajic, Barbara Catania, Tania
    Cerquitelli, Silvia Chiusano, Panos K. Chrysanthis,
    Carlo Curino, Jérôme Darmont, Amr El Abbadi,
    Avrilia Floratou, Juliana Freire, Alekh Jindal,
    Vana Kalogeraki, Georgia Koutrika, Arun Kumar,
    Sujaya Maiyya, Alexandra Meliou, Madhulika
    Mohanty, Felix Naumann, Nele Sina Noack, Fatma
    Özcan, Liat Peterfreund, Wenny Rahayu,
    Wang-Chiew Tan, Yuanyuan Tian, Pinar Tözün,
    Genoveva Vargas-Solar, Neeraja J. Yadwadkar, and
    Meihui Zhang. Diversity and inclusion activities in
    database conferences: A 2021 report. *SIGMOD
    Rec.*, 51(2):69–73, 2022.

[5] American Sociological Association (ASA). ASA
    Diversity, Equity & Inclusion. `https://www.asanet.org/diversity-equity-inclusion`.

[6] Sourav S Bhowmick. Closet: Conflict of interest
    detection & management system. 2019.
    `https://personal.ntu.edu.sg/assourav/research/DARE/closet.html`.

[7] NeurIPS Foundation. Ethics guidelines.
    `https://nips.cc/public/EthicsGuidelines`.

[8] DB DEI Initiative. DB DEI Materials.
    `https://dbdni.github.io/#materials`.