

Inclusiveness in Russian NLP conferences



Double-blind reviewing

- ▶ **Double-blind** means the authors do not know the names of the reviewers, and the reviewers do not know the names of the authors.
- ▶ This is pretty standard for **all** top-tier international CL/NLP conferences.

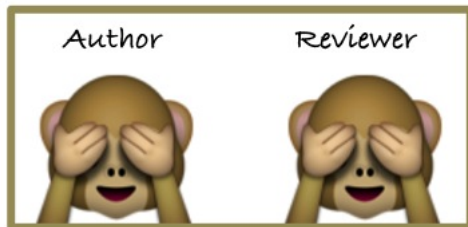
Inclusiveness in Russian NLP conferences



Double-blind reviewing

- ▶ **Double-blind** means the authors do not know the names of the reviewers, and the reviewers do not know the names of the authors.
- ▶ This is pretty standard for **all** top-tier international CL/NLP conferences.

Inclusiveness in Russian NLP conferences



Double-blind reviewing

- ▶ **Double-blind** means the authors do not know the names of the reviewers, and the reviewers do not know the names of the authors.
- ▶ This is pretty standard for **all** top-tier international CL/NLP conferences.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
 - ▶ but it allows to significantly alleviate this bias,
 - ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Why do we use double-blind reviewing?

- ▶ Humans are **subjective** (consciously or unconsciously);
- ▶ If authors are not anonymized, the reviewers' decisions are inevitably **biased** [Tomkins et al., 2017]:
 - ▶ papers by **well-known authors** are accepted 1.5 times **more often**,
 - ▶ papers by **well-known companies** are accepted 2 times **more often**,
 - ▶ papers with a **female first author** are accepted 20% **less often**.
- ▶ Anonymization is not ideal...
- ▶ but it allows to significantly alleviate this bias,
- ▶ and as a result make the scientific program more diverse, and the **conference itself more inclusive**.

Inclusiveness in Russian NLP conferences

Russian NLP conferences try to keep up with that

- ▶ *AIST* switched completely to double-blind reviewing starting from 2017,
- ▶ *Dialogue* did the same starting from 2019.

We (*Irina Nikishina* and *Andrey Kutuzov*) set to find out whether this changed anything in terms of **inclusiveness**.

- ▶ Did the **number of female authors** increased?
- ▶ Did the **number of 'non-centrally located' authors** increased?

Inclusiveness in Russian NLP conferences

Russian NLP conferences try to keep up with that

- ▶ *AIST* switched completely to double-blind reviewing starting from 2017,
- ▶ *Dialogue* did the same starting from 2019.

We (*Irina Nikishina* and *Andrey Kutuzov*) set to find out whether this changed anything in terms of inclusiveness.

- ▶ Did the number of female authors increased?
- ▶ Did the number of 'non-centrally located' authors increased?

Inclusiveness in Russian NLP conferences

Russian NLP conferences try to keep up with that

- ▶ *AIST* switched completely to double-blind reviewing starting from 2017,
- ▶ *Dialogue* did the same starting from 2019.

We (*Irina Nikishina* and *Andrey Kutuzov*) set to find out whether this changed anything in terms of inclusiveness.

- ▶ Did the number of female authors increased?
- ▶ Did the number of 'non-centrally located' authors increased?

Inclusiveness in Russian NLP conferences

Russian NLP conferences try to keep up with that

- ▶ *AIST* switched completely to double-blind reviewing starting from 2017,
- ▶ *Dialogue* did the same starting from 2019.

We (*Irina Nikishina* and *Andrey Kutuzov*) set to find out whether this changed anything in terms of **inclusiveness**.

- ▶ Did the **number of female authors** increased?
- ▶ Did the **number of 'non-centrally located' authors** increased?

Inclusiveness in Russian NLP conferences

Russian NLP conferences try to keep up with that

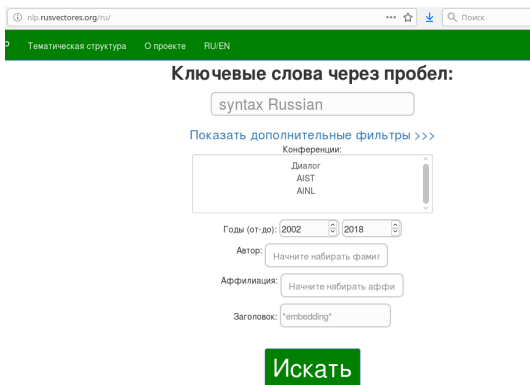
- ▶ *AIST* switched completely to double-blind reviewing starting from 2017,
- ▶ *Dialogue* did the same starting from 2019.

We (*Irina Nikishina* and *Andrey Kutuzov*) set to find out whether this changed anything in terms of **inclusiveness**.

- ▶ Did the **number of female authors** increased?
- ▶ Did the **number of 'non-centrally located' authors** increased?

Inclusiveness in Russian NLP conferences

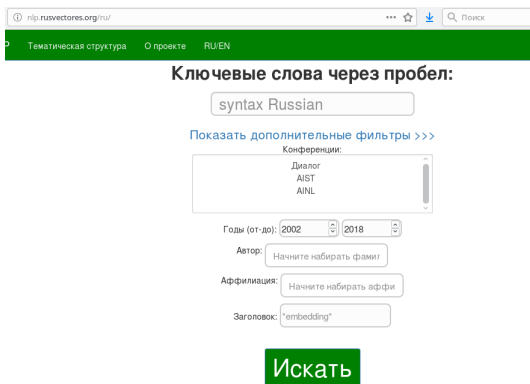
- ▶ Data from the **RusNLP** project
- ▶ <http://nlp.rusvectors.org/>
- ▶ ...plus fresh numbers from 2019
- ▶ Gender and geography annotated manually, based on authors' names and affiliations.
 - ▶ 'Center' is Moscow, Saint-Petersburg and foreign countries,
 - ▶ 'Province' is everything else.



The screenshot shows a web browser window with the URL nlp.rusvectors.org/. The page has a green header with navigation links: "Тематическая структура", "О проекте", and "RU/EN". The main heading is "Ключевые слова через пробел:" (Keywords separated by spaces:). Below it is a search input field containing "syntax Russian". A blue link "Показать дополнительные фильтры >>>" (Show additional filters >>>) is present. Underneath, there are several filter sections: "Конференции:" (Conferences) with a list box containing "Диалог", "AIST", and "AINL"; "Годы (от-до):" (Years (from-to):) with dropdown menus for "2002" and "2018"; "Автор:" (Author) with a text input field containing "Начните набирать фамил"; "Аффилиация:" (Affiliation) with a text input field containing "Начните набирать аффи"; and "Заголовок:" (Title) with a text input field containing "embedding". At the bottom center is a large green button with the text "Искать" (Search). The browser's address bar and search bar are visible at the top.

Inclusiveness in Russian NLP conferences

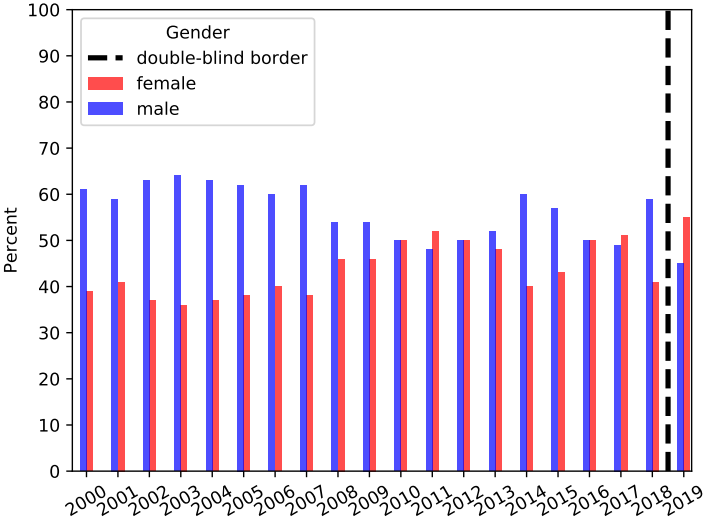
- ▶ Data from the **RusNLP** project
- ▶ <http://nlp.rusvectors.org/>
- ▶ ...plus fresh numbers from 2019
- ▶ Gender and geography annotated manually, based on authors' names and affiliations.
 - ▶ **'Center'** is Moscow, Saint-Petersburg and foreign countries,
 - ▶ **'Province'** is everything else.



The screenshot shows a web browser window with the URL nlp.rusvectors.org/. The page has a green header with navigation links: "Тематическая структура", "О проекте", and "RU/EN". The main heading is "Ключевые слова через пробел:" (Keywords separated by spaces:). Below it is a search input field containing "syntax Russian". A blue link "Показать дополнительные фильтры >>>" (Show additional filters >>>) is present. Underneath, there are several filter sections: "Конференции:" (Conferences) with a list box containing "Диалог", "AIST", and "AINL"; "Годы (от-до):" (Years (from-to):) with dropdown menus for "2002" and "2018"; "Автор:" (Author) with an input field "Начните набирать фамию" (Start typing the surname); "Аффилиация:" (Affiliation) with an input field "Начните набирать аффи" (Start typing the affiliation); and "Заголовок:" (Title) with an input field containing "embedding". At the bottom center is a large green button with the text "Искать" (Search). The browser's address bar and search bar are visible at the top.

Gender distribution

Gender distribution among Dialogue authors from 2000 to 2019



Gender distribution

Dialogue males

- ▶ Average yearly percentage:
- ▶ 57 (before double blind)
- ▶ 45 (after double blind)
- ▶ Difference: -12

Dialogue females

- ▶ Average yearly percentage:
- ▶ 43 (before double blind)
- ▶ 55 (after double blind)
- ▶ Difference: +12

Naively comparing with the standard deviation before 2019: 5
The difference seems significant: more female authors.

Gender distribution

Dialogue males

- ▶ Average yearly percentage:
- ▶ 57 (before double blind)
- ▶ 45 (after double blind)
- ▶ Difference: -12

Dialogue females

- ▶ Average yearly percentage:
- ▶ 43 (before double blind)
- ▶ 55 (after double blind)
- ▶ Difference: +12

Naively comparing with the standard deviation before 2019: 5
The difference seems significant: more female authors.

Gender distribution

Dialogue males

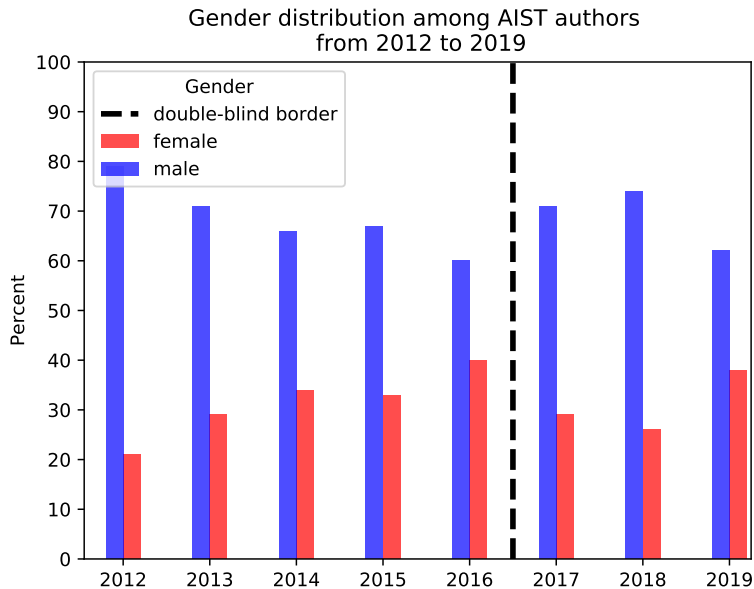
- ▶ Average yearly percentage:
- ▶ 57 (before double blind)
- ▶ 45 (after double blind)
- ▶ Difference: -12

Dialogue females

- ▶ Average yearly percentage:
- ▶ 43 (before double blind)
- ▶ 55 (after double blind)
- ▶ Difference: +12

Naively comparing with the standard deviation before 2019: 5
The difference seems significant: more female authors.

Gender distribution



Gender distribution

AIST males

- ▶ Average yearly percentage:
- ▶ 69 (before double blind)
- ▶ 69 (after double blind)
- ▶ Difference: 0

AIST females

- ▶ Average yearly percentage:
- ▶ 31 (before double blind)
- ▶ 31 (after double blind)
- ▶ Difference: 0

Welch T-test: `Ttest_indResult(statistic=-0.08, pvalue=0.94)`

No difference.

Gender distribution

AIST males

- ▶ Average yearly percentage:
- ▶ 69 (before double blind)
- ▶ 69 (after double blind)
- ▶ Difference: 0

AIST females

- ▶ Average yearly percentage:
- ▶ 31 (before double blind)
- ▶ 31 (after double blind)
- ▶ Difference: 0

Welch T-test: `Ttest_indResult(statistic=-0.08, pvalue=0.94)`

No difference.

Gender distribution

AIST males

- ▶ Average yearly percentage:
- ▶ 69 (before double blind)
- ▶ 69 (after double blind)
- ▶ Difference: 0

AIST females

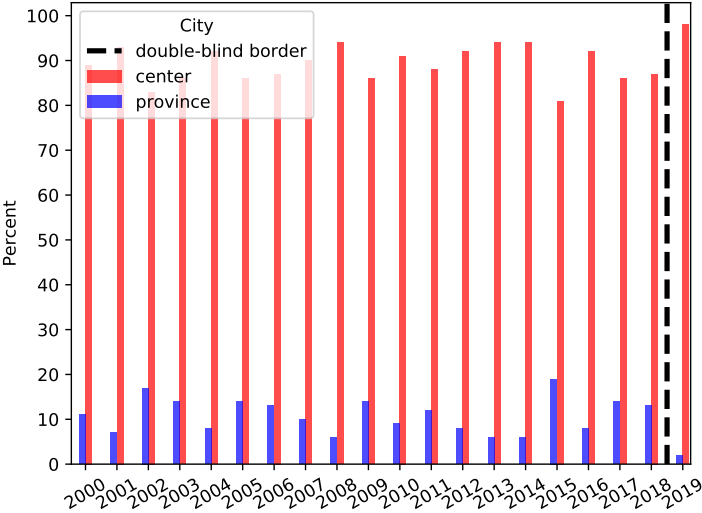
- ▶ Average yearly percentage:
- ▶ 31 (before double blind)
- ▶ 31 (after double blind)
- ▶ Difference: 0

Welch T-test: `Ttest_indResult(statistic=-0.08, pvalue=0.94)`

No difference.

Location distribution

City distribution among Dialogue authors from 2000 to 2019



Location distribution

Dialogue Center

- ▶ Average yearly percentage:
- ▶ 89 (before double blind)
- ▶ 98 (after double blind)
- ▶ Difference: +9

Dialogue Province

- ▶ Average yearly percentage:
- ▶ 11 (before double blind)
- ▶ 2 (after double blind)
- ▶ Difference: -9

Naively comparing with the standard deviation before 2019: 4
The difference seems significant (although in an unexpected direction: more 'central' authors).

Location distribution

Dialogue Center

- ▶ Average yearly percentage:
- ▶ 89 (before double blind)
- ▶ 98 (after double blind)
- ▶ Difference: +9

Dialogue Province

- ▶ Average yearly percentage:
- ▶ 11 (before double blind)
- ▶ 2 (after double blind)
- ▶ Difference: -9

Naively comparing with the standard deviation before 2019: 4
The difference seems significant (although in an unexpected direction: more 'central' authors).

Location distribution

Dialogue Center

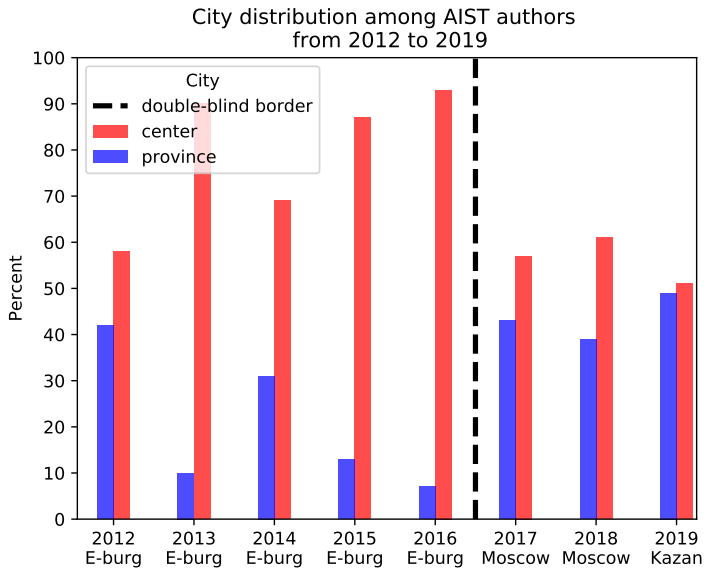
- ▶ Average yearly percentage:
- ▶ 89 (before double blind)
- ▶ 98 (after double blind)
- ▶ Difference: +9

Dialogue Province

- ▶ Average yearly percentage:
- ▶ 11 (before double blind)
- ▶ 2 (after double blind)
- ▶ Difference: -9

Naively comparing with the standard deviation before 2019: 4
The difference seems significant (although in an unexpected direction: more 'central' authors).

Location distribution



Location distribution

AIST Center

- ▶ Average yearly percentage:
- ▶ 79 (before double blind)
- ▶ 56 (after double blind)
- ▶ Difference: -23

AIST Province

- ▶ Average yearly percentage:
- ▶ 21 (before double blind)
- ▶ 44 (after double blind)
- ▶ Difference: +23

Welch T-test: `Ttest_indResult(statistic=2.48, pvalue=0.048)`

Significantly more 'non-central' authors!

Location distribution

AIST Center

- ▶ Average yearly percentage:
- ▶ 79 (before double blind)
- ▶ 56 (after double blind)
- ▶ Difference: -23

AIST Province

- ▶ Average yearly percentage:
- ▶ 21 (before double blind)
- ▶ 44 (after double blind)
- ▶ Difference: +23

Welch T-test: `Ttest_indResult(statistic=2.48, pvalue=0.048)`
Significantly more 'non-central' authors!

Location distribution

AIST Center

- ▶ Average yearly percentage:
- ▶ 79 (before double blind)
- ▶ 56 (after double blind)
- ▶ Difference: -23

AIST Province

- ▶ Average yearly percentage:
- ▶ 21 (before double blind)
- ▶ 44 (after double blind)
- ▶ Difference: +23

Welch T-test: `Ttest_indResult(statistic=2.48, pvalue=0.048)`

Significantly more 'non-central' authors!

Inclusiveness in Russian NLP conferences

Summary of changes after introducing double-blind reviewing

- ▶ Significant **increase in the number of 'non-central' authors** for *AIST*
- ▶ Significant **increase in the number of female authors** for *Dialogue*
- ▶ Significant decrease in the number of 'non-central' authors for *Dialogue* (???)
- ▶ No changes in gender distribution for *AIST*.

NB: for *Dialogue*, we rely on only **one** year after the introduction of double-blind reviewing (2019).



We welcome everyone, by the way :-)

Inclusiveness in Russian NLP conferences

Summary of changes after introducing double-blind reviewing

- ▶ Significant **increase in the number of 'non-central' authors** for *AIST*
- ▶ Significant **increase in the number of female authors** for *Dialogue*
- ▶ Significant decrease in the number of 'non-central' authors for *Dialogue* (???)
- ▶ No changes in gender distribution for *AIST*.

NB: for *Dialogue*, we rely on only **one** year after the introduction of double-blind reviewing (2019).



We welcome everyone, by the way :-)

Inclusiveness in Russian NLP conferences

Summary of changes after introducing double-blind reviewing

- ▶ Significant **increase in the number of 'non-central' authors** for *AIST*
- ▶ Significant **increase in the number of female authors** for *Dialogue*
- ▶ Significant decrease in the number of 'non-central' authors for *Dialogue* (???)
- ▶ No changes in gender distribution for *AIST*.

NB: for *Dialogue*, we rely on only **one** year after the introduction of double-blind reviewing (2019).



We welcome everyone, by the way :-)

Inclusiveness in Russian NLP conferences

Summary of changes after introducing double-blind reviewing

- ▶ Significant **increase in the number of 'non-central' authors** for *AIST*
- ▶ Significant **increase in the number of female authors** for *Dialogue*
- ▶ Significant decrease in the number of 'non-central' authors for *Dialogue* (???)
- ▶ No changes in gender distribution for *AIST*.

NB: for *Dialogue*, we rely on only **one** year after the introduction of double-blind reviewing (2019).



We welcome everyone, by the way :-)