

Governance regels in het Linksight platform

Versie 1.0

In dit artikel leggen we uit waarom er governance regels zijn in het Linksight platform en hoe deze governance regels in een datasamenwerking kunt toepassen.

De noodzaak van governance regels

Governance regels zijn een set aan afspraken die de organisaties onderling afspreken. De regels worden opgesteld omdat een organisatie de controle wil houden over haar eigen data en geen 'carte blanche' wil geven aan de andere organisaties in de datasamenwerking. De regels bepalen dan ook de speelruimte waarbinnen een organisatie mee doet in een datasamenwerking.

In eerdere datasamenwerkingen kan je deze afspraken bijvoorbeeld terugvinden in de bijlage van een samenwerkingsovereenkomst. In de datasamenwerkingen die werken met het Linksight platform worden deze regels in het platform vastgelegd en ook technisch afgedwongen. Dit wil zeggen, als een regel wordt overtreden betekent dit dat een gebruiker een bepaalde actie niet kan uitvoeren of dat een analyse wordt gestopt.

De governance regels die gezamenlijk worden opgesteld, worden door iedere partij bevestigd door middel van een stemprocedure. Pas nadat iedereen ingestemd heeft, zijn de (nieuwe) regels van kracht en kunnen analyses gedaan worden.

Voor volledige duidelijkheid en transparantie wordt in het auditlog bijgehouden wie welke regels heeft voorgedragen en het stemgedrag van de organisaties in de datasamenwerking.

Naast deze governance regels kan elke organisatie nog aan elke dataset policies koppelen die het gebruik ervan beperken (o.a. of met een attribuut gerekend mag worden, of dat het slechts als filter gebruikt mag worden).

Als je benieuwd bent naar de werking van governance rules in het Linksight platform, klik dan op [deze link](#) voor een korte demo over de governance rules.

Toepassing van de governance regels

Er zijn diverse categorieën governance rules:

1. Generiek
2. Analysis
3. Output controle
4. Ethisch

Per categorie en per regel leggen we in document het doel en de werking uit.

1 Generiek

1.1 Regelset vervaldatum

Het doel van deze regel is om af te dwingen dat elke periode de regels weer tegen het licht worden gehouden. Bijvoorbeeld om te checken of er aanpassingen nodig zijn. Daarnaast is deze regel nuttig in situaties waar bijvoorbeeld voor een eenmalig onderzoek analyses uitgevoerd worden en een einddatum meegegeven kan worden. Na die datum kan geen analyse meer gedaan worden, maar natuurlijk zijn de resultaten van eerdere analyses nog wel beschikbaar voor verwerking in bijvoorbeeld rapportages.

1.2 Maximaal aantal analyses per tijdseenheid per organisatie

Het doel van deze regel is het risico op onbedoelde onthulling nog verder te verkleinen. Het voorkomt het (bedoeld of onbedoeld) “afpellen” van de onderliggende databronnen doormiddel van grote aantallen analysevragen met kleine variaties. Daarnaast worden organisaties gedwongen om gericht vragen te stellen over het onderwerp.

2 Analyse

2.1 Analyse starten

Het doel van deze regel is om te beperken welke organisaties een analyse mogen initiëren. Dit geeft meer controle of hoe de onderliggende databronnen gebruikt worden.

2.2 Toegestane protocollen

Het doel van deze regel is om vast te stellen welke soort analyseprotocol binnen de datasamenwerking is toegestaan. Hiermee wordt de speelruimte bepaald en dit geldt voor alle organisaties in de datasamenwerking.

2.3 Handmatige goedkeuring

Het doel van deze regel is extra controle te geven aan een organisatie. Met deze regel kan één of meerdere organisaties binnen de samenwerking aangewezen worden om elke ingestuurde analyse eerst handmatig goed te keuren voordat deze wordt uitgevoerd.

2.4 Handmatige goedkeuring van alle partijen in een query

Het doel van deze regel is extra controle te geven aan de organisaties. Met deze regel actief moeten alle organisaties wiens data wordt gebruikt in een analyse de ingestuurde analyse eerst handmatig goedkeuren voordat deze wordt uitgevoerd.

2.5 Handmatige goedkeuring van alle partijen en een datasamenwerking

Het doel van deze regel is extra controle te geven aan alle organisaties. Met deze regel actief moeten alle organisaties in een datasamenwerking de ingestuurde analyse eerst handmatig goedkeuren voordat deze wordt uitgevoerd (dus ongeacht of data van een organisatie wordt gebruikt).

2.6 Selectiebeleid voor attribuut

Het doel van deze regel is om organisaties nog meer controle te geven over hoe bepaalde attributen uit de gekoppelde dataset in de datasamenwerking gebruikt kunnen worden. Met deze regel kan ingesteld worden of en hoe een attribuut in de selectie van een analyse gebruikt mag worden. De opties zijn op dit moment “weigeren” en “alleen aggregaat”. Deze regel beperkt eventueel bestaande policies van een dataset binnen de datasamenwerking.

2.7 Filterbeleid voor attribuut

Het doel van deze regel is om organisaties nog meer controle te geven over hoe bepaalde attributen uit de gekoppelde dataset in de datasamenwerking gebruikt kunnen worden. Met deze regel kan ingesteld worden of en hoe een attribuut als filter gebruikt mag worden. De 3

opties zijn “weigeren”, “toestaan als filter”, “niet-exacte filters toestaan”. Deze regel beperkt eventueel bestaande policies van een dataset binnen de datasamenwerking.

2.8 Aggregatiefunctie in query

Het doel van deze regel is om organisaties nog meer controle te geven over hoe bepaalde attributen uit de gekoppelde dataset in de datasamenwerking gebruikt kunnen worden. Met deze regel kan ingesteld worden welke aggregatiefuncties gebruikt mogen worden in analyses. De opties zijn “tellen”, “som”, “gemiddelde”, “gemiddelde en standaarddeviatie”.

2.9 Federated learning regressie; minimale grootte subpopulatie

Het doel van deze regel is om bij gebruik van regressie-analyse de kans op onbedoelde onthulling te verkleinen door een minimale grootte van de subpopulatie in te stellen.

2.10 Federated learning regressie; maximaal aantal onafhankelijke variabelen

Het doel van deze regel is om organisaties nog meer controle te geven en om bij gebruik van regressie-analyse de kans op onbedoelde onthulling te verkleinen door het aantal onafhankelijk variabelen te maximaliseren. Op die manier wordt de kans op “overfitting” verkleind en is daarmee het risico dat uit het resultaatmodel over de gebruikte data af te leiden is.

2.11 Federated learning regressie; niet toegestaan voor attribuut y

Het doel van deze regel is om organisaties nog meer controle te geven en om bij gebruik van regressie-analyse de kans op onbedoelde onthulling te verkleinen door te voorkomen dat bepaalde attributen als “afhankelijke variable” worden gebruikt in de regressie.

2.12 Federated learning regressie; niet toegestaan voor attribuut x

Het doel van deze regel is om organisaties nog meer controle te geven en om bij gebruik van regressie-analyse de kans op onbedoelde onthulling te verkleinen door te voorkomen dat bepaalde attributen als “onafhankelijke variable” worden gebruikt in de regressie.

2.13 Federated learning regressie; filterbeleid voor attribuut

Het doel van deze regel is om organisaties nog meer controle te geven en om bij gebruik van regressie-analyse de kans op onbedoelde onthulling te verkleinen door te voorkomen dat bepaalde attributen als filter attribuut worden gebruikt.

2.14 Lijst ontdebelling; filterbeleid voor attribuut

Het doel van deze regel is om organisaties nog meer controle te geven over hoe bepaalde attributen uit de gekoppelde dataset in de datasamenwerking gebruikt kunnen worden. Met deze regel kan ingesteld worden of en hoe een attribuut als filter gebruikt mag worden. De 3 opties zijn “weigeren”, “toestaan als filter”, “niet-exacte filters toestaan”. Deze regel beperkt eventueel bestaande policies van een dataset binnen de datasamenwerking.

3 Uitvoercontrole

3.1 Beschrijvende statistiek; grootte anonimiteitsset

Het doel van deze regel is om ongewenste onthulling tegen te gaan door de grootte van de anonimiteitsset in te stellen. De grootte kan verschillen per datasamenwerking en is mede afhankelijk van de totale onderzoekspopulatie.

De governance regels worden steeds verder uitgebreid, net als andere aspecten rondom “data colaberation governance”, bijvoorbeeld het stemmechanisme. De nieuwste versie van dit document is te vinden op: <https://www.linksight.nl/content/governance-rules>

