

# A Novel Metric for Assessing User Influence based on User Behaviour

---

ANTONELA TOMMASEL – DANIELA GODOY



# Table of Contents

1. Introduction
2. Information-based User Characterisation
3. Quantitatively Assessing User Influence
4. Data Analysis
5. Summary

# Defining User Influence

- The potential or ability of an individual to:
  - Engage others in a certain act.
  - Induce others to behave in a particular manner.
- In micro-blogging sites, the definition relies on status attributes. (e.g. followers, re-tweets, mentions)
- The behaviour of users regarding the information diffusion might be related to their influence.

# Defining User Influence

- The potential or ability of an individual to:
  - Engage others in a certain act.
  - Induce others to behave in a particular manner.
- In micro-blogging sites, the definition relies on status attributes. (e.g. followers, re-tweets, mentions)
- The behaviour of users regarding the information diffusion might be related to their influence.

**However, a high level of popularity is NOT sufficient for also being influential in terms of triggering social responses.**

# Defining User Influence

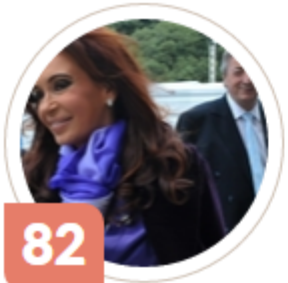
## Commercial Metrics

- They do not publicly state how scores are computed.
- They have received several critics.
  - How measurements are computed.
  - Effect of spam-bots on the algorithms.



# Defining User Influence

## Comparing Commercial Metrics



### Cristina Fernandez

Cristina Fernández de Kirchner, Presidenta de la Nación Argentina. En facebook:  
<http://www.facebook.com/CFKArgentina>

Apartments and Home Rentals ★ Expert    Argentina ★ Expert  
 Cardinal Stritch University ★ Expert    CoTweet ★ Expert    Edward Burns ★ Expert  
 Joseph Stalin ★ Expert    Lighthouses ★ Expert    Mar Del Plata ★ Expert  
 Michael Franti ★ Expert    Michael Moore ★ Expert    Mozilla Firefox ★ Expert  
 Paul McCartney ★ Expert    Polar Bears ★ Expert    Time Warner ★ Expert  
 TweetDeck ★ Expert    Wachovia ★ Expert

### My Kred Communities Score



CFKArgentina

Follow

999

of 1,000  
Influence

Outreach  
Level  
5 of 10

Global Kred

Yoga

Marketing

Travel

### My Kred Communities Score



CFKArgentina

Follow

998

of 1,000  
Influence

Outreach  
Level  
0 of 10

Global Kred

Yoga

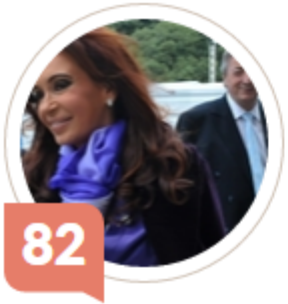
Marketing

Travel

Radio

# Defining User Influence

## Comparing Commercial Metrics



### Cristina Fernandez

Cristina Fernández de Kirchner, Presidenta de la Nación Argentina. En facebook:  
<http://www.facebook.com/CFKArgentina>

Apartments and Home Rentals ★ Expert    Argentina ★ Expert  
 Cardinal Stritch University ★ Expert    CoTweet ★ Expert    Edward Burns ★ Expert  
 Joseph Stalin ★ Expert    Lighthouses ★ Expert    Mar Del Plata ★ Expert  
 Michael Franti ★ Expert    Michael Moore ★ Expert    Mozilla Firefox ★ Expert  
 Paul McCartney ★ Expert    Polar Bears ★ Expert    Time Warner ★ Expert  
 TweetDeck ★ Expert    Wachovia ★ Expert

### My Kred Communities Score



CFKArgentina

Follow

999

of 1,000  
Influence

Outreach  
Level  
5 of 10

Global Kred

Yoga

Marketing

Travel

### My Kred Communities Score



CFKArgentina

Follow

998

of 1,000  
Influence

Outreach  
Level  
0 of 10

Global Kred

Yoga

Marketing

Travel

Radio

**There is no consensus on what it means to be an influential user!!**

# User Influence

**A novel metric of user influence based on analysing user behaviour regarding the patterns of information diffusion.**



# User Influence

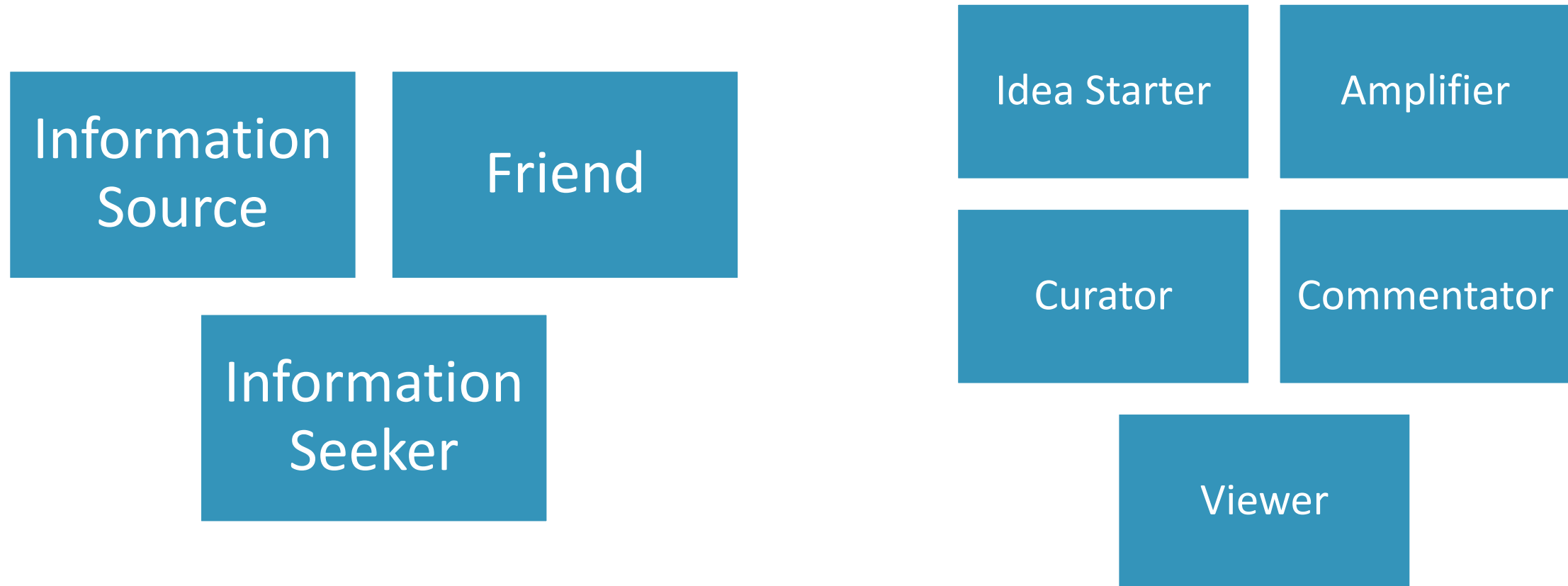
**A novel metric of user influence based on analysing user behaviour regarding the patterns of information diffusion.**

- It considers both content-based and topological factors.
- It aims at assessing the interactions with his/her neighbourhood.

# Table of Contents

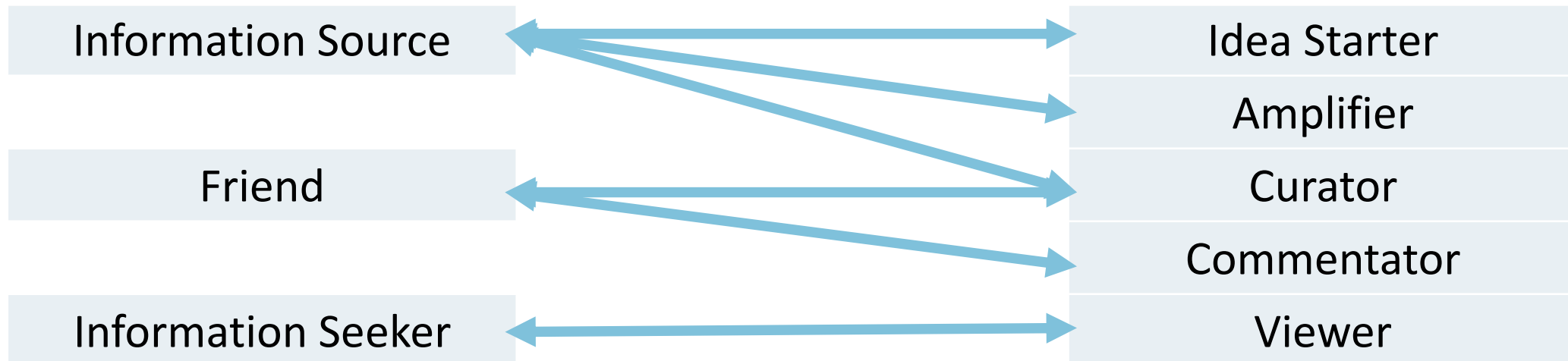
1. Introduction
- 2. Information-based User Characterisation**
3. Quantitatively Assessing User Influence
4. Data Analysis
5. Summary

# Behavioural-based User Characterisation



# Behavioural-based User Characterisation

## Association Between Characterisations



# Table of Contents

1. Introduction
2. Information-based User Characterisation
- 3. Quantitatively Assessing User Influence**
4. Data Analysis
5. Summary

# Quantitative Assessment of Personality

- Novel definitions for each dimension are presented.
- A quantitative assessment of user influence is presented.
- Scores are constrained to the interval  $[0,1]$ .
- Corrections are applied in order to avoid undetermined values.

# Quantitative Assessment of Personality

## Idea Starter

- Highly engaged with the media.
- The majority of their posts correspond to original content.
- They interact with a limited and selected group of users.

$$\frac{|Tweets_{ORIGINAL}\{Tweets_{ORIGINAL}RT \geq \mu - \sigma\}|}{|Tweets|} * \frac{\sum Tweets_{ORIGINAL}RT}{|ReTweets|}$$

- Assesses the proportion of original posts.
- Assesses the distribution of received retweets over the published tweets.
- Assesses the impact of posts in the neighbourhood of the user.
- *The higher the score, the more the user behaves as an Idea Starter, and thus as an Information Source.*

# Quantitative Assessment of Personality

## Amplifier

- Share ideas and opinions posted by other users.
- The majority of their posts correspond to retweets of Idea Starters.
- Posts are also highly retweeted by their followers.

$$\frac{|Tweets_{ORIGINAL}| + |Tweets_{REPLY}| + |Tweets_{MENTION}|}{|Tweets|} * \frac{\sum Tweets_{ORIGINAL}RT + \sum Tweets_{REPLY}RT + \sum Tweets_{MENTION}RT}{|ReTweets|}$$

- Assesses the interaction between a user and his/her social network.
- Assesses the impact of posts in the neighbourhood.
- *The higher the score, the more a user behaves as an Amplifier, and thus as an Information Source.*



# Quantitative Assessment of Personality

## Curator

- Interact with both Idea Starters and Amplifiers.
- Share a lot of content and interact with a large number of users.

$$\frac{|Interactions \in \{Followees \cup Followers\}|}{|Interactions|} * \frac{|Interactions \in \{Followees \cup Followers\}|}{|Followees| + |Followers|}$$

- Assesses the number of interactions with other users.
- Assesses to what extent a user interacts only with his/her neighbourhood.
- *The higher the number of interactions, the higher the score, and the less the user behaves as an Information Source.*

# Quantitative Assessment of Personality

## Follower/Followee Ratio (FF-Ratio)

$$\frac{|Followers|}{|Followers| + |Followees|}$$

- The content-related dimensions do not consider the size of the neighbourhood of a user.
- Two users might achieve the same score but have a totally different neighbourhood.
- Leverages the importance of the neighbourhood size.

# Quantitative Assessment of Personality

## Information Source Index

$$\mu_{IDAC} = \frac{Idea\ Starter + Amplifier + (1 - Curator)}{3}$$

$$IS = \frac{2 * \mu_{IDAC} * FF_{Ratio}}{\mu_{IDAC} + FF_{Ratio}}$$

- High values of IS denote users behaving as Information Sources, whereas low values of IS denote users behaving as Information Seekers.
- Information Sources meet all the requirements for being regarded as influential users.
- *The higher the IS of a user, the higher the influence of the user.*

# Table of Contents

1. Introduction
2. Information-based User Characterisation
3. Quantitatively Assessing User Influence
- 4. Data Analysis**
5. Summary

# Data Analysis Settings

## Data Collections

Argentina

(13 users)

Miscellaneous

(8 users)

Music

(8 users)

Politics

(7 users)

Sports

(7 users)

Technology

(7 users)

TV Shows –  
Movies

(10 users)

# Data Analysis Settings

## Human Assessment

- Considering that there is no consensus on what means to be an influential user the human perception of influence was analysed.
- 31 Undergraduate and graduate students from an Artificial Intelligence course at UNICEN University (Argentina) ranked Twitter users according to their perceived influence.
- To combine the rankings provided by the students, Twitter users were assigned the mode of the provided rankings.

# Data Analysis Settings

## Web Site

The screenshot shows a website titled "Análisis de la Influencia de Usuarios en Twitter". At the top right, there is a search bar with the text "Search this site". On the left side, there is a navigation menu with the following items: HOME, ARGENTINA (with sub-items FORMULARIO and TIMELINES), DEPORTES (with sub-items FORMULARIO and TIMELINES), MISCELÁNEO (with sub-items FORMULARIO and TIMELINES), MÚSICA (with sub-items FORMULARIO and TIMELINES), POLÍTICA (with sub-items FORMULARIO and TIMELINES), TECNOLOGÍA (with sub-items FORMULARIO and TIMELINES), TV - PELÍCULAS (with sub-items FORMULARIO and TIMELINES), and SITEMAP.

The main content area is titled "Home" and contains the following text:

Bienvenidos! En esta página encontrarán los diferentes formularios a completar respecto a la influencia de determinados usuarios en *Twitter*.

- **Cómo se suele definir la influencia de usuarios en redes sociales?**

La influencia de un usuario puede ser definida como el prestigio que un usuario tiene sobre su entorno social, y entonces, como el potencial o la habilidad de un individuo en generar ciertas acciones en los demás o incluso a comportarse de una manera en particular. La influencia también puede ser asociada a la popularidad de los usuarios, es decir, el grado con el cual un individuo es aceptado por sus amistades.

En el contexto de las redes sociales, la definición de influencia recae en atributos como la cantidad de followers (es decir, el tamaño del grupo de influencia de un usuario), re-tweets (es decir, la habilidad para generar contenido que resulte atractivo de ser distribuido) o la cantidad de menciones recibidas (es decir, la habilidad para involucrar otros usuarios en conversaciones). Asimismo, existen numerosas métricas comerciales que intentan medir la influencia de los usuarios (por ejemplo Klout, Kred o WeFollow), sin embargo, las mismas no han hecho público los algoritmos mediante los cuales miden la influencia. En consecuencia, sus cálculos no pueden ser reproducidos, lo que puede afectar su legitimidad y confiabilidad. Incluso, lo que es más importante, puede no haber un consenso entre ellas respecto de lo que significa ser influyente.

- **Cuál es el objetivo de la actividad?**

Considerando que la definición de cuán influyentes son ciertos usuarios, esta actividad tiene como objetivo analizar la percepción humana de dicha influencia. Luego, dicha percepción de influencia, definida en forma de rankings de influencia, será comparada tanto con los rankings de influencias generados por diversas métricas comerciales, como por una nueva métrica. Es a partir de la comparación de los diferentes rankings que se podrá obtener la correlación entre los mismos y así analizar cuán similar es la apreciación humana de influencia a la sugerida por las métricas.

- **Qué van a encontrar en el sitio?**

El sitio cuenta con diferentes secciones, cada una de ellas se corresponde con un tópico y un conjunto de usuarios de *Twitter* diferente. Por cada una de las secciones van a encontrar dos sub-páginas. En la primera de ellas van a encontrar los timelines resumidos de cada uno de los usuarios que tienen que analizar. En la segunda de ellas se encuentra el formulario que se debe completar.

- **En qué consiste la actividad?**

La actividad consiste en analizar a los usuarios de cada uno de los conjuntos y compararlos de acuerdo a la influencia de los mismos. Es importante destacar que no

<https://sites.google.com/site/influenciatwitterusers/>

# Data Analysis Settings

## Web Site

▼ **POLÍTICA**

FORMULARIO

**TIMELINES**

@barackObama  
 @algore  
 @tonyblairoffice  
 @PrensaMichelle  
 @PutinRF\_Eng  
 @dilmabr  
 @NicolasMaduro  
 @Pontifex\_es

FORMULARIO

TIMELINES

▼ **DEPORTES**

FORMULARIO

TIMELINES

▼ **MISCELÁNEO**

FORMULARIO

TIMELINES

▼ **MÚSICA**

FORMULARIO

TIMELINES

▼ **POLÍTICA**

FORMULARIO

**TIMELINES**

▼ **TECNOLOGÍA**

FORMULARIO

TIMELINES

▼ **TV - PELÍCULAS**

FORMULARIO


TIMELINES


**SITEMAP**

### Timelines

#### DILMA ROUSSEFF

**Tweets** [Follow](#)

 **Dilma Rousseff** @dilmabr 9h  
 Confira a íntegra da mensagem na minha página no Facebook: goo.gl/qqjllub #NãoAReduçãoDaMaioridadePenal  
 Expand

 **Dilma Rousseff** @dilmabr 9h  
 Lugar de meninos e meninas é na escola. Chega de impunidade para aqueles que aliciam crianças e adolescentes para o crime.  
 Expand

 **Dilma Rousseff** @dilmabr 9h  
 É preciso endurecer a lei, mas para punir com mais rigor os adultos que aliciam menores para o crime organizado.  
 Expand

Tweet to @dilmabr

#### NICOLÁS MADURO

**Tweets** [Follow](#)

 **VTVCANAL8** @VTVcanal8 20h  
 #VictoriaEnPanama pic.twitter.com/i5cfD1rF36  
 Retweeted by Nicolás Maduro



#### BARACK OBAMA

**Tweets** [Follow](#)

 **Barack Obama** @BarackObama 4h  
 Climate change deniers are running out of excuses. #ActOnClimate pic.twitter.com/quYIY2dyBQ



Expand

 **The White House** @WhiteHouse 10h

Tweet to @BarackObama

#### PAPA FRANCISCO

**Tweets** [Follow](#)

 **Papa Francisco** @Pontifex\_es 10 Apr  
 Si el Evangelio arraiga profundamente en nuestras vidas, seremos capaces de llevarlo a los demás.  
 Expand

 **Papa Francisco** @Pontifex\_es 9 Apr

<https://sites.google.com/site/influenciatwitterusers/>



# Data Analysis Settings

## Web Site

**\* Required**

**Ranking de Influencia de Users \***

Para cada usuario seleccione la posición en el ranking de influencias, siendo la posición 1 la correspondiente al usuario más influyente y 10 la del usuario menos influyente. Tener en cuenta que cada posición del ranking puede ser utilizada una única vez.

	1	2	3	4	5	6	7
Barack Obama (@baracObama)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Papa Francisco (@Pontifex_es)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vladimir Putin (@PutinRF_Eng)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dilma Rousseff (@dilmabr)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Michelle Bachelet (@prensaMichelle)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nicolás Maduro (@NicolasMaduro)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tony Blair (@tonyblairoffice)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Submit**

<https://sites.google.com/site/influenciatwitterusers/>

# Metrics Used for Comparison

Klout  
Score

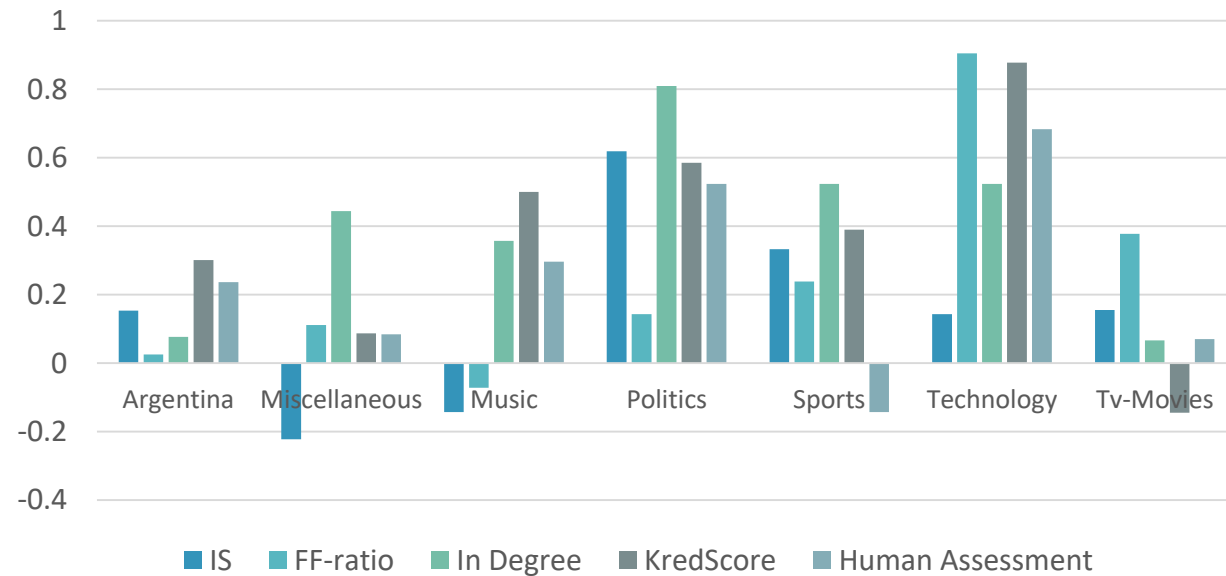
Kred  
Score

In  
Degree

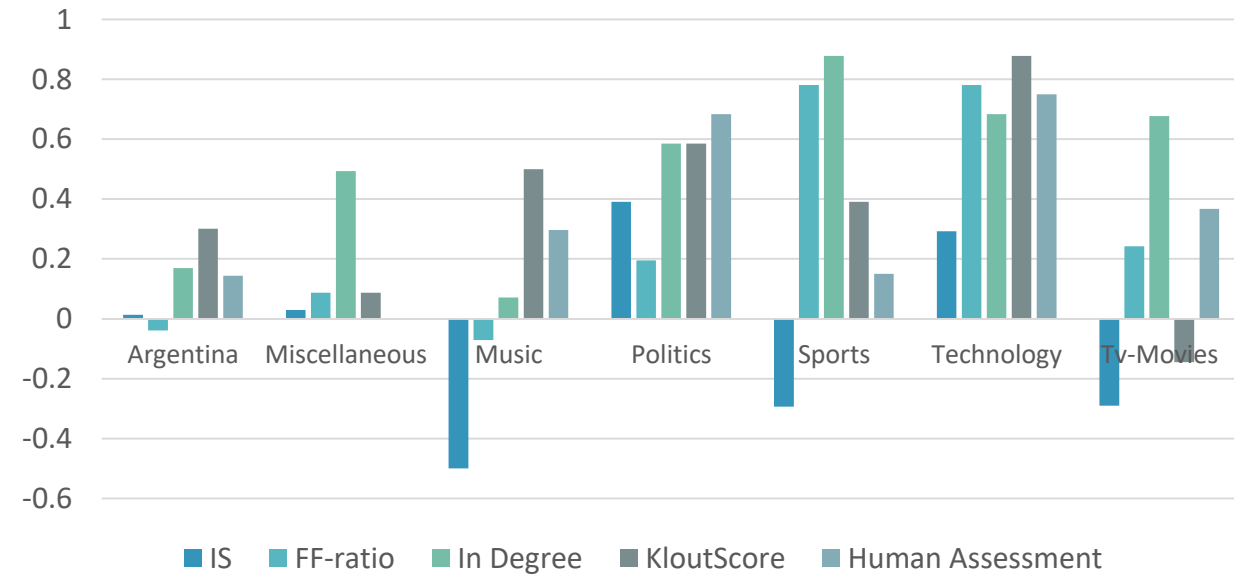
FF-  
Ratio

# Analysis of Influence Metrics

## Klout Vs. All

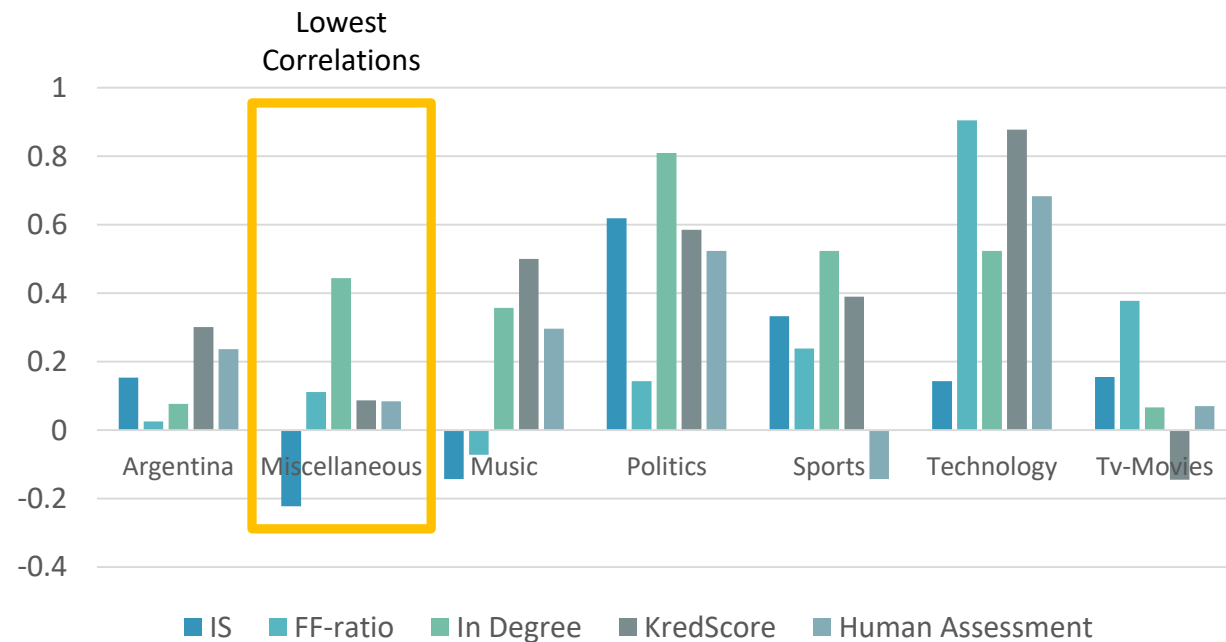


## Kred Vs. All

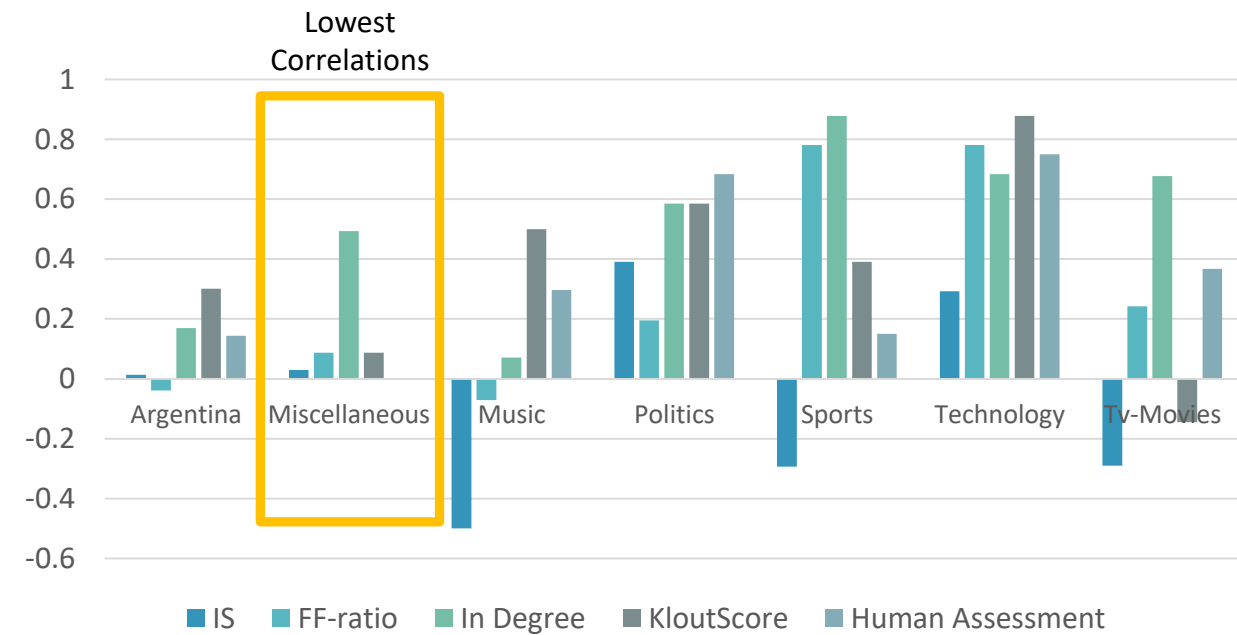


# Analysis of Influence Metrics

## Klout Vs. All

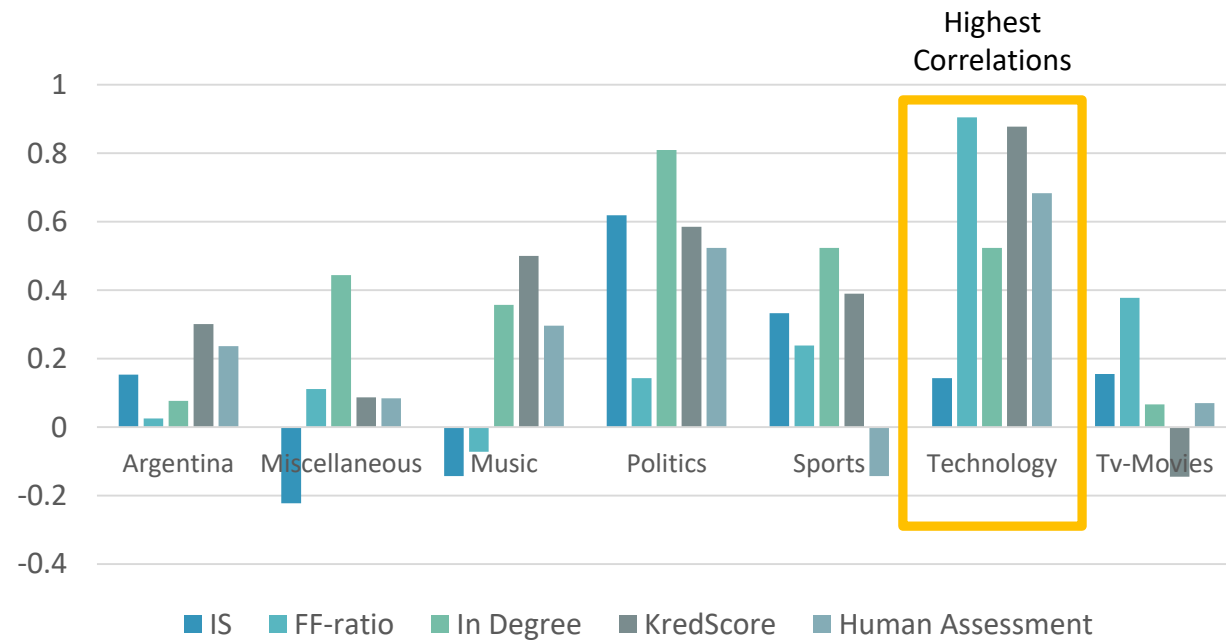


## Kred Vs. All

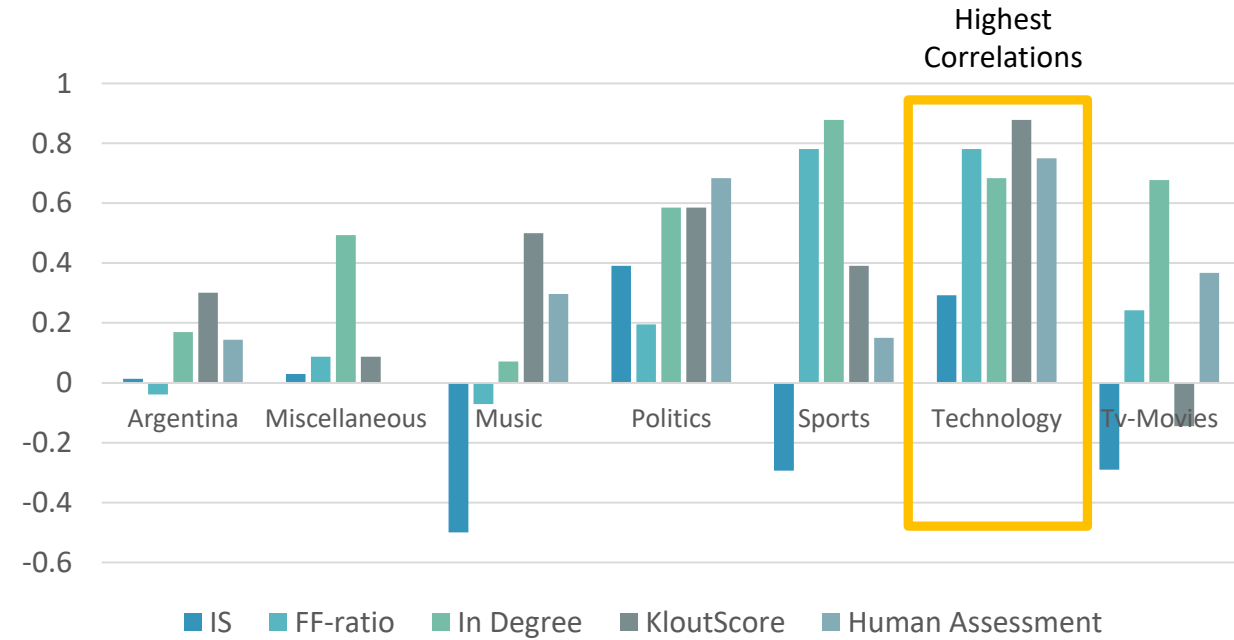


# Analysis of Influence Metrics

## Klout Vs. All

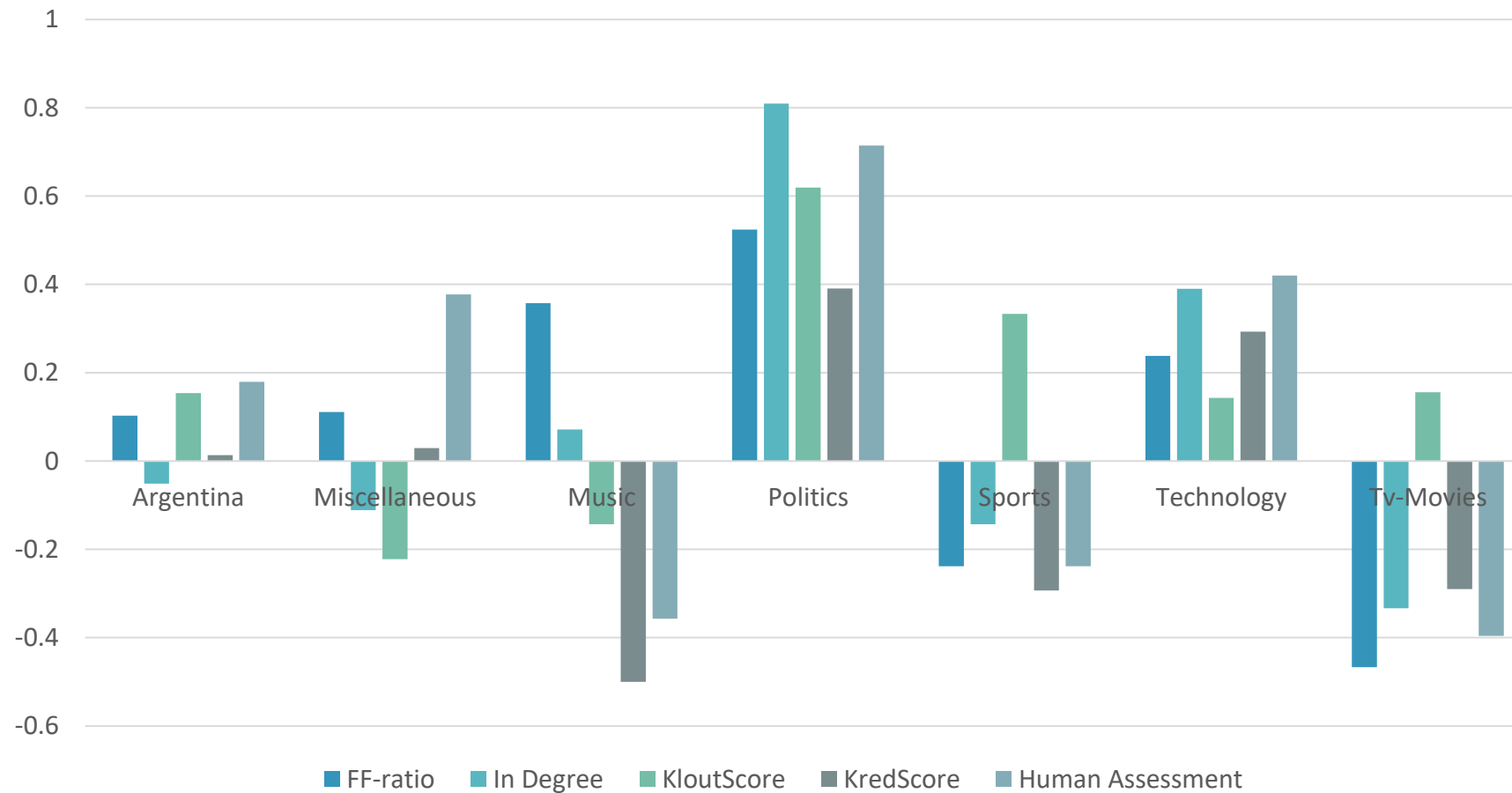


## Kred Vs. All



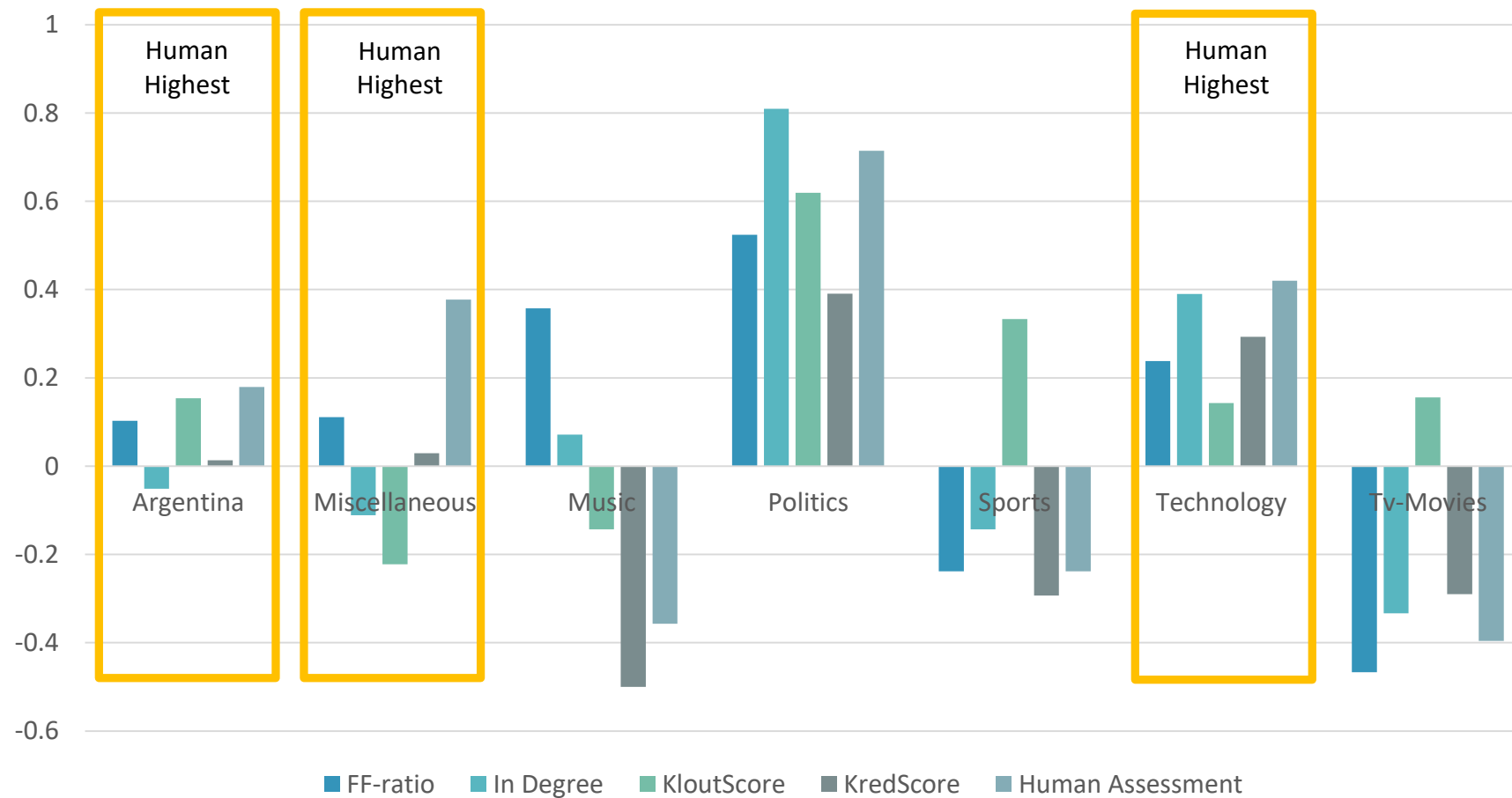
# Analysis of Influence Metrics

## IS Index Vs. All



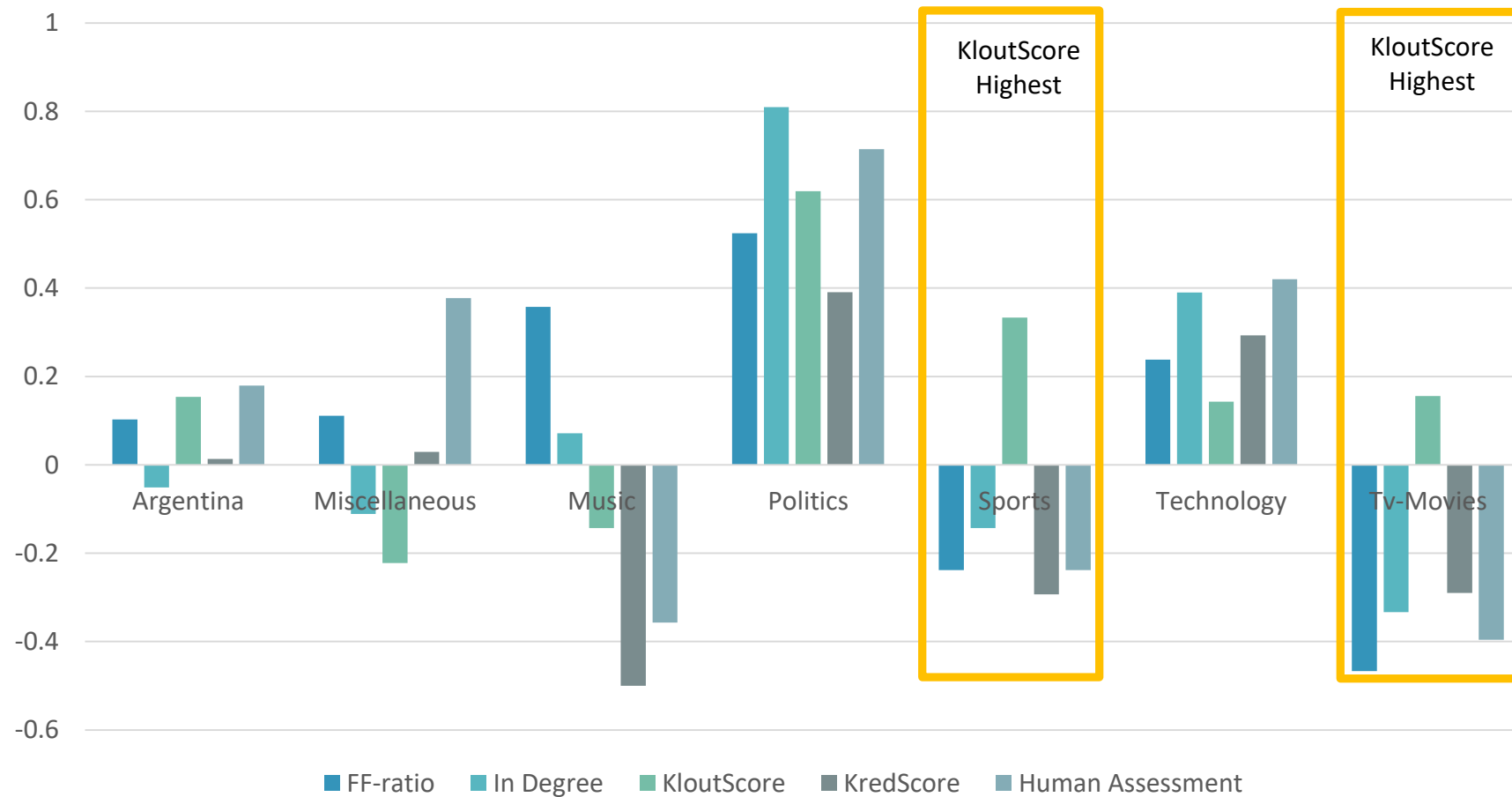
# Analysis of Influence Metrics

## IS Index Vs. All



# Analysis of Influence Metrics

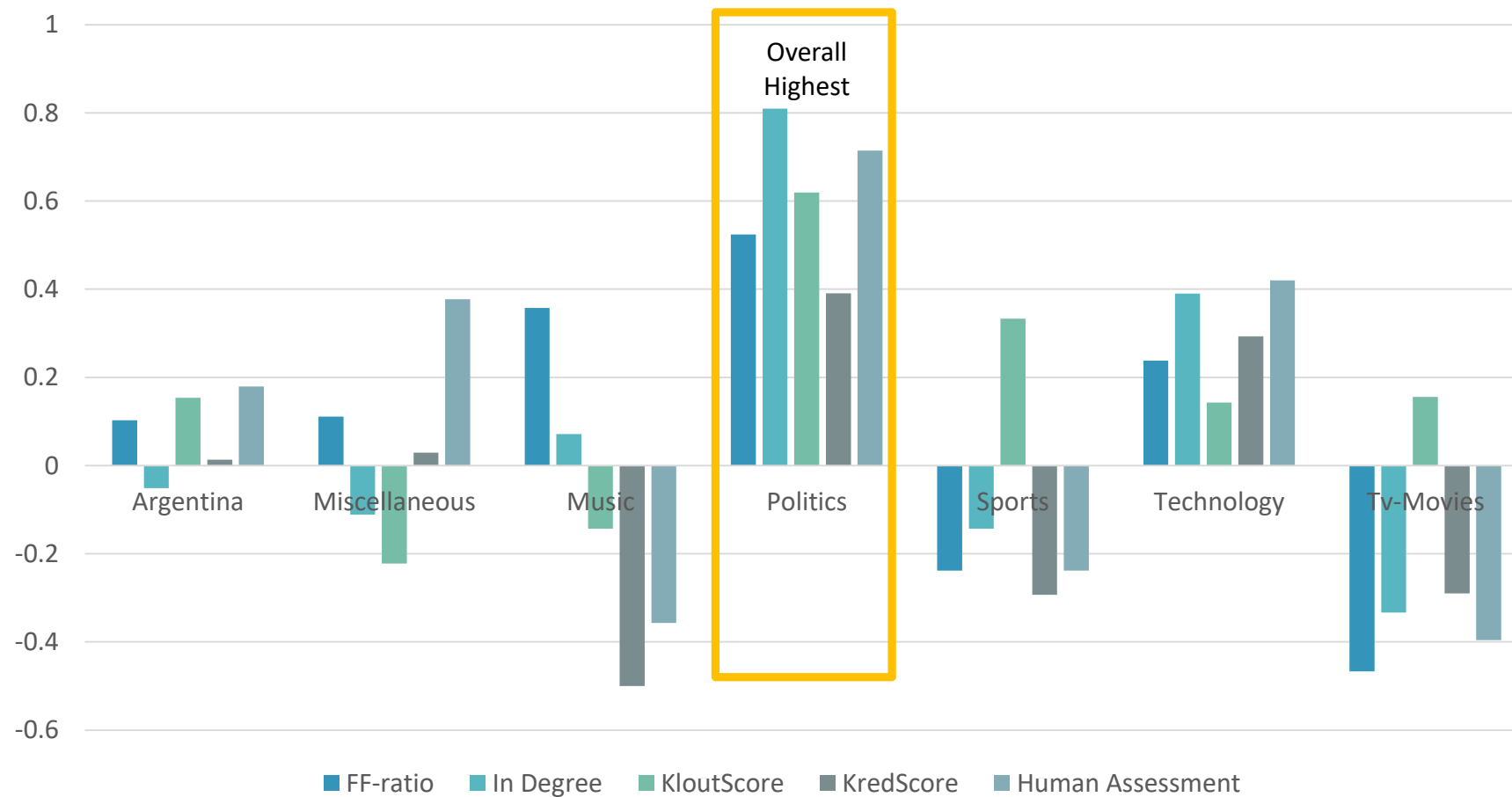
## IS Index Vs. All





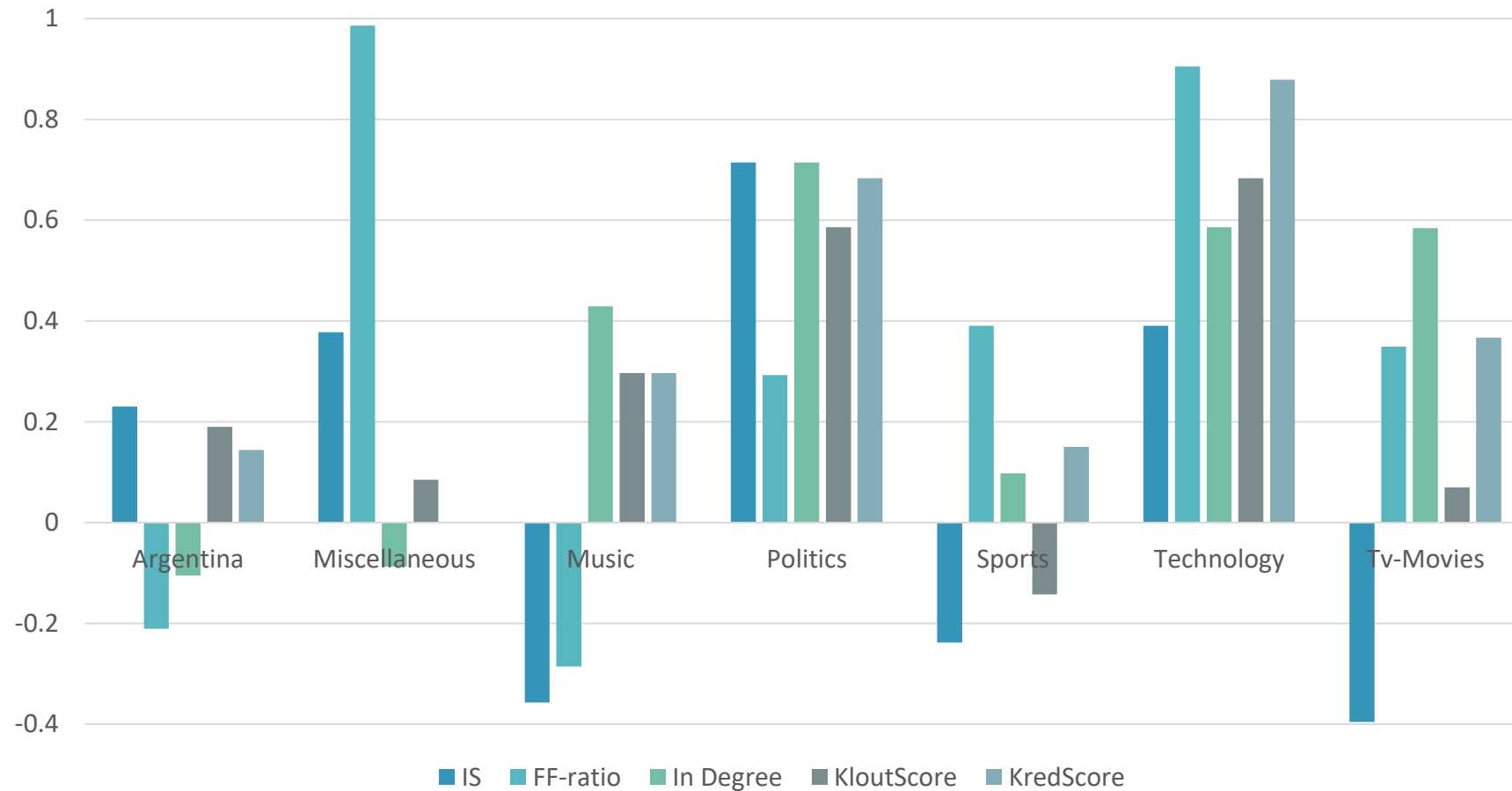
# Analysis of Influence Metrics

## IS Index Vs. All



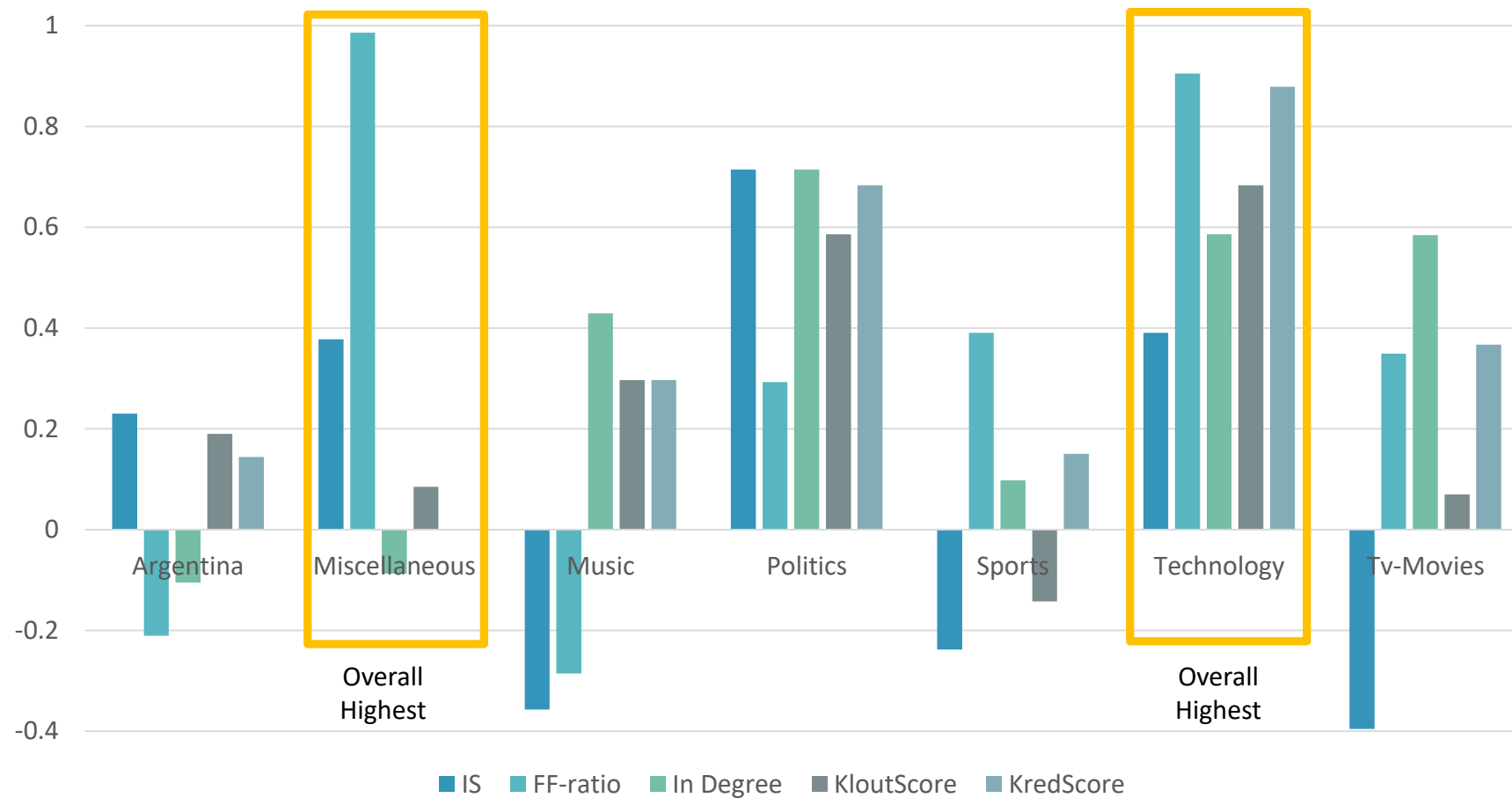
# Analysis of Influence Metrics

## Human Assessment Vs. All



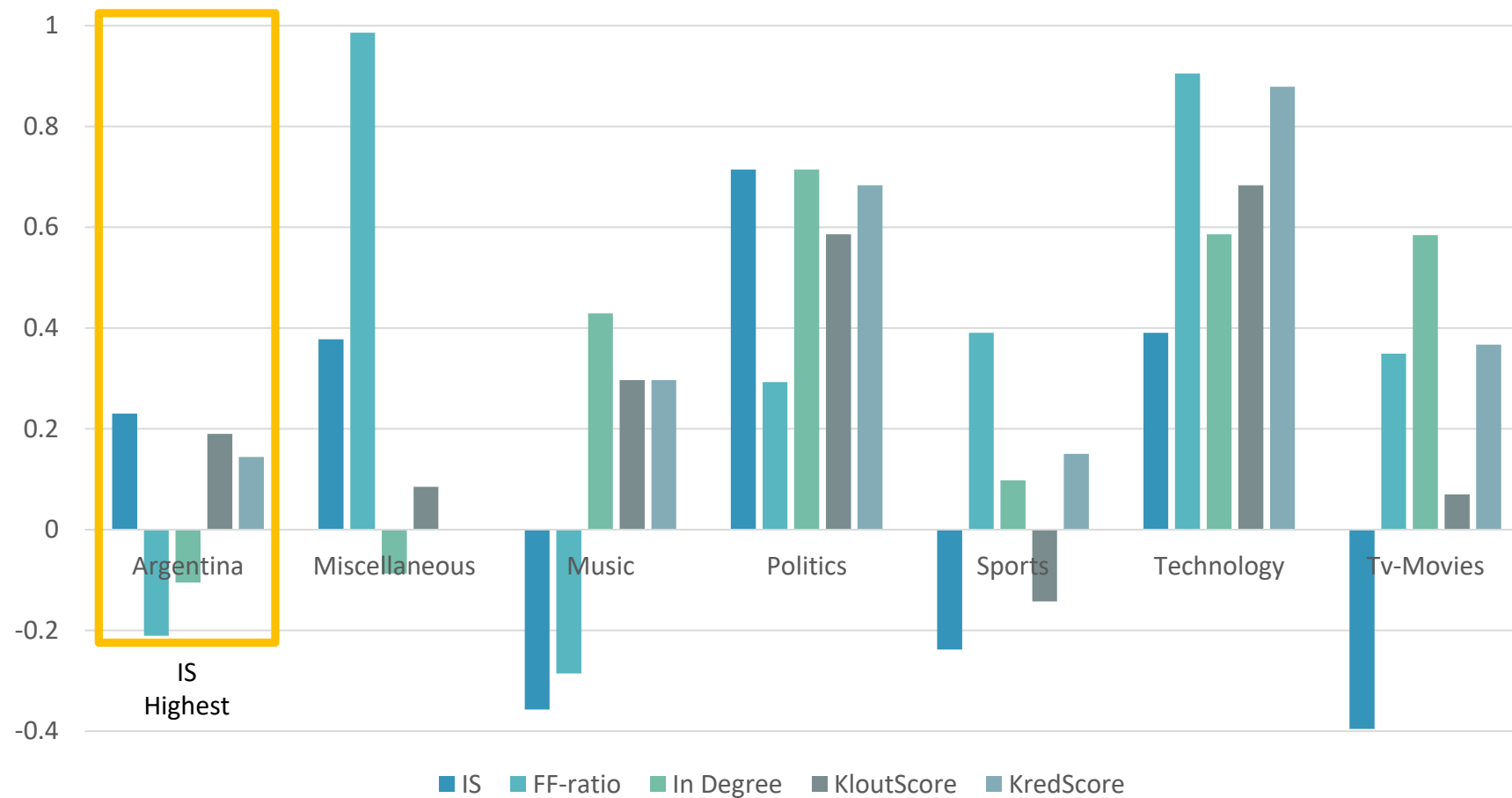
# Analysis of Influence Metrics

## Human Assessment Vs. All



# Analysis of Influence Metrics

## Human Assessment Vs. All



# Table of Contents

1. Introduction
2. Information-based User Characterisation
3. Quantitatively Assessing User Influence
4. Data Analysis
5. Summary

# Summary

- This work proposes a novel metric for measuring user influence.
- The metric is based on analysing user behaviour in the diffusion of information.
- The metric considers both content-based and topological factors.
- The metric was compared to traditional means for assessing user influence, commercial metrics and a human assessment of user influence.

# Conclusions

- There is no unique definition of user influence.
- User influence might differ according to the analysed topic.
- In most cases, the commercial metrics proved to be independent from each other.
- The highest correlations for the IS were achieved regarding the human assessment.
- Only analysing the topological factors is not sufficient for accurately assessing the influence of users.
- An accurate assessment of user influence might also consider the published content and its influence over the neighbourhood of users.

# Conclusions

- There is no unique definition of user influence.
- User influence might differ according to the analysed topic.
- In most cases, the commercial metrics proved to be independent from each other.
- The highest correlations for the IS were achieved regarding the human assessment.
- Only analysing the topological factors is not sufficient for accurately assessing the influence of users.
- An accurate assessment of user influence might also consider the published content and its influence over the neighbourhood of users.



# Conclusions

- There is no unique definition of user influence.
- User influence might differ according to the analysed topic.
- In most cases, the commercial metrics proved to be independent from each other.
- The highest correlations for the IS were achieved regarding the human assessment.
- Only analysing the topological factors is not sufficient for accurately assessing the influence of users.
- An accurate assessment of user influence might also consider the published content and its influence over the neighbourhood of users.

# Questions?



# A Novel Metric for Assessing User Influence based on User Behaviour

---

ANTONELA TOMMASEL – DANIELA GODOY

