

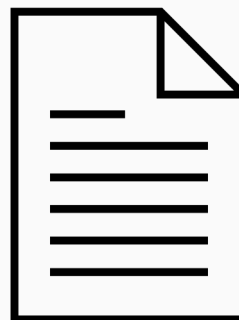
Modelado de Tópicos

Mtro. René Rosado González
Director de Programa LTP

Problemática



Problemática



Problemática

P Proceso

AMLO se deslinda de iniciativa de Morena que limita facultades de la Corte

No se actúa por consigna, no se dan órdenes. No es como antes o como actuaba el presidente en otros tiempos (...) Era el presidente el que...



Problemática

P Proceso

AMLO se deslinda de iniciativa de Morena que limita facultades de la Corte

No se actúa por consigna, no se dan órdenes. No es como antes o como actuaba el presidente en otros tiempos (...) Era el presidente el que...

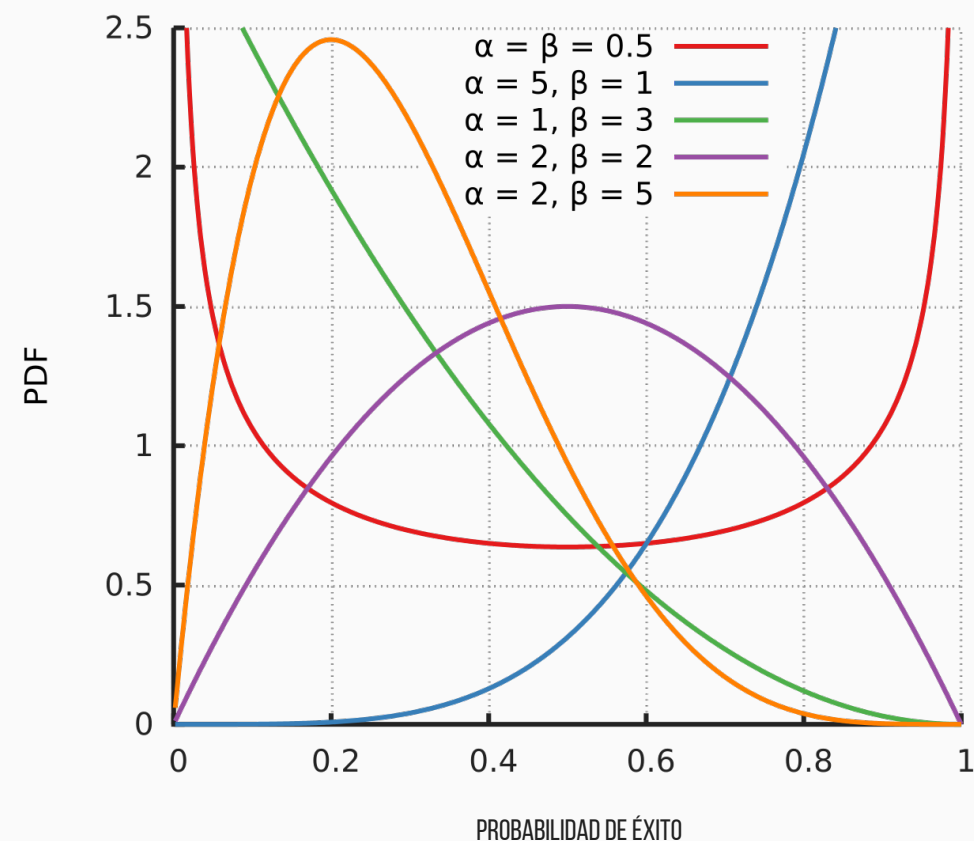


La Distribución Beta

Beta(α, β)

La distribución Beta es una distribución de probabilidad sobre probabilidades dada por:

$$\text{PDF} = \frac{x^{\overset{\text{NÚMERO DE ÉXITOS}}{\downarrow} \alpha - 1} (1 - x)^{\overset{\text{NÚMERO DE FALLOS}}{\downarrow} \beta - 1}}{\underset{\substack{\uparrow \\ \text{NORMALIZA ENTRE} \\ \text{0 Y 1}}}{B(\alpha, \beta)}}$$



Problemática

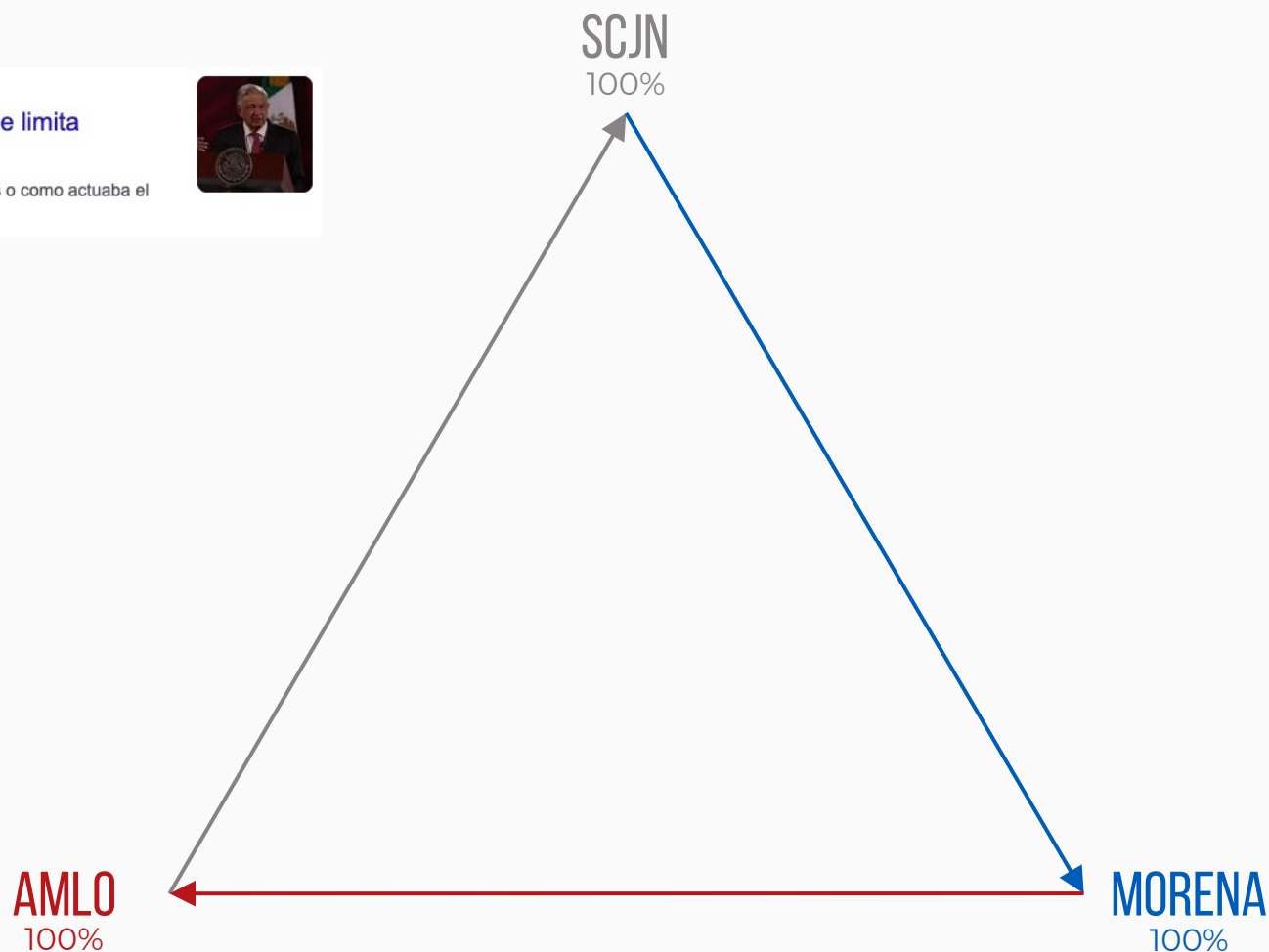
Proceso

AMLO se deslinda de iniciativa de Morena que limita facultades de la Corte

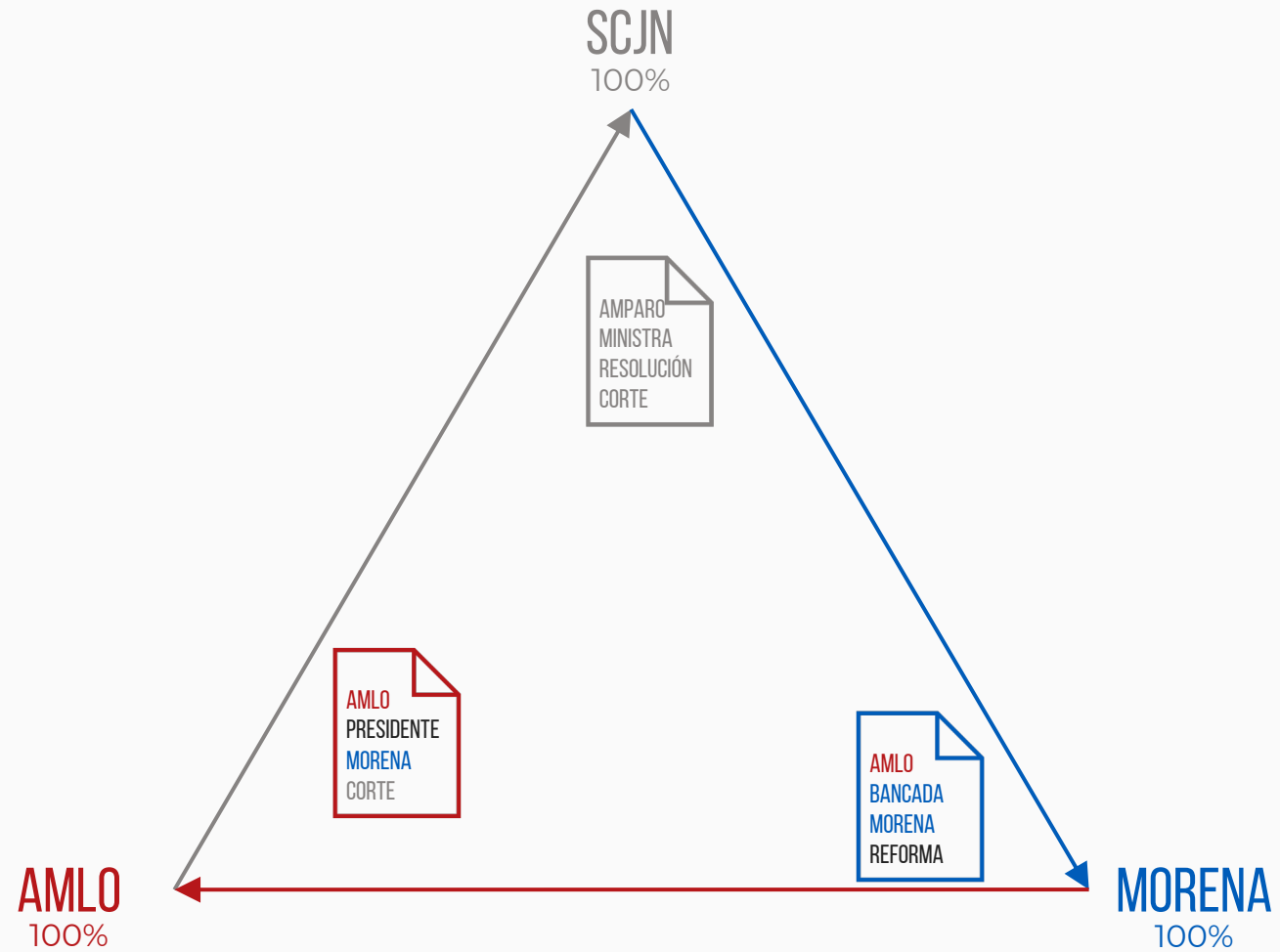
No se actúa por consigna, no se dan órdenes. No es como antes o como actuaba el presidente en otros tiempos (...) Era el presidente el que...



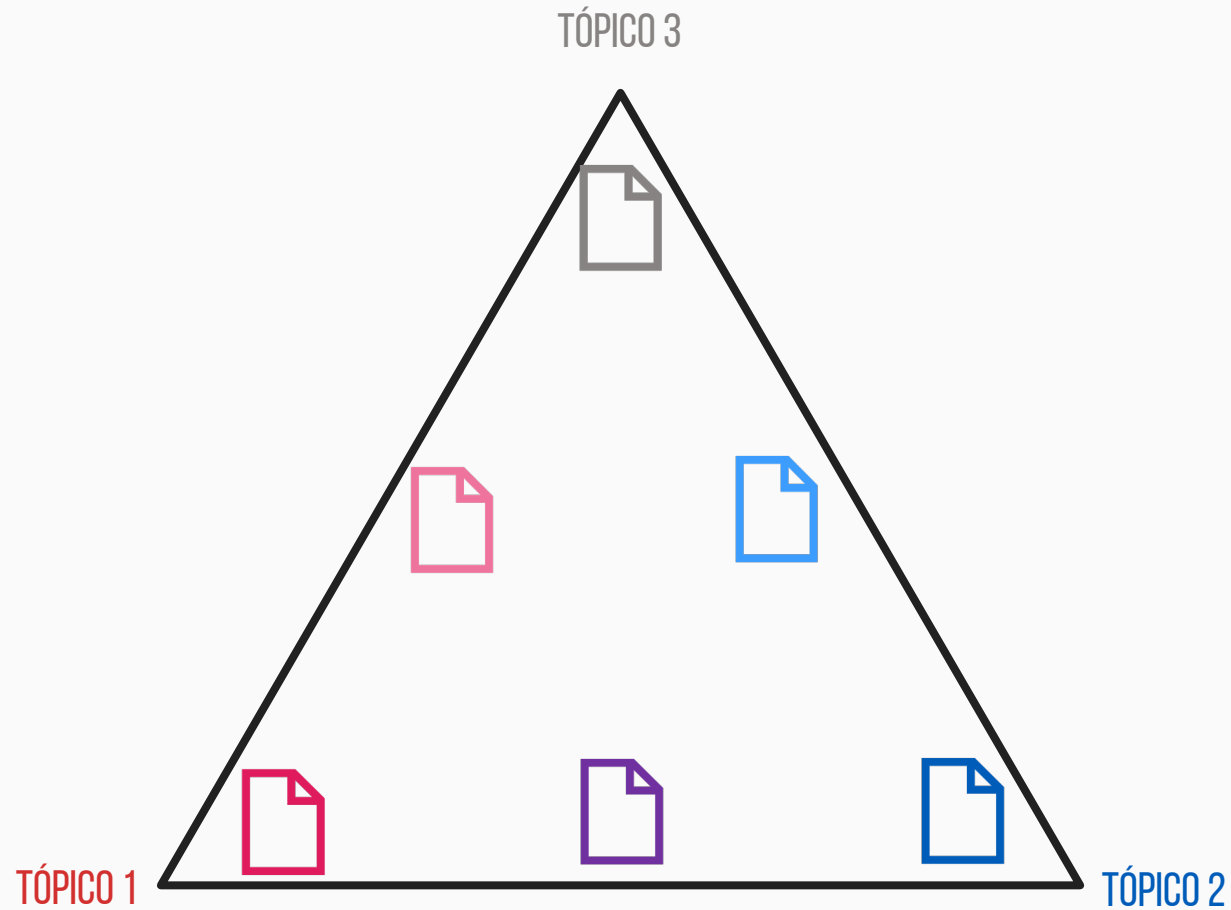
AMLO
PRESIDENTE
MORENA
CORTE



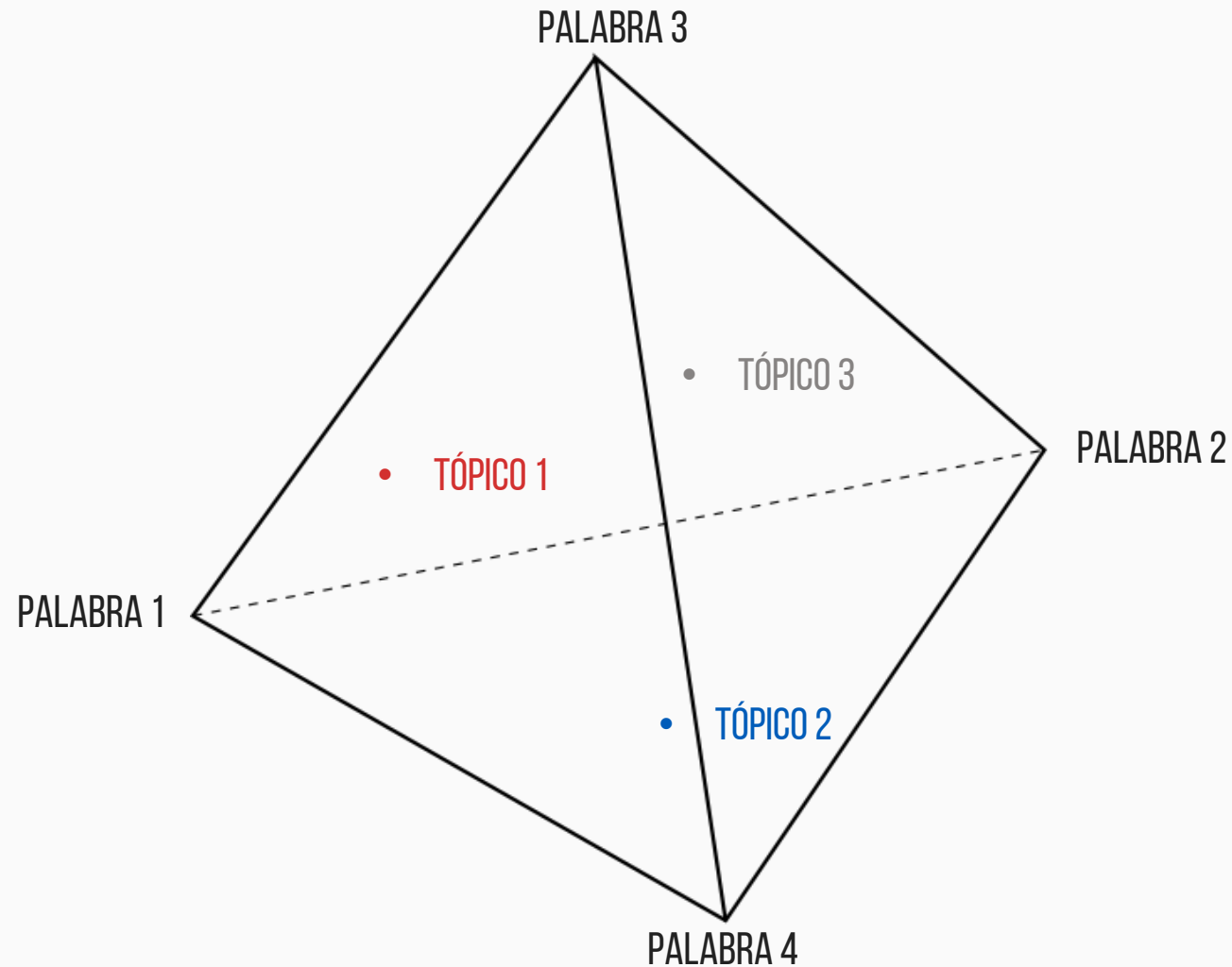
Problemática



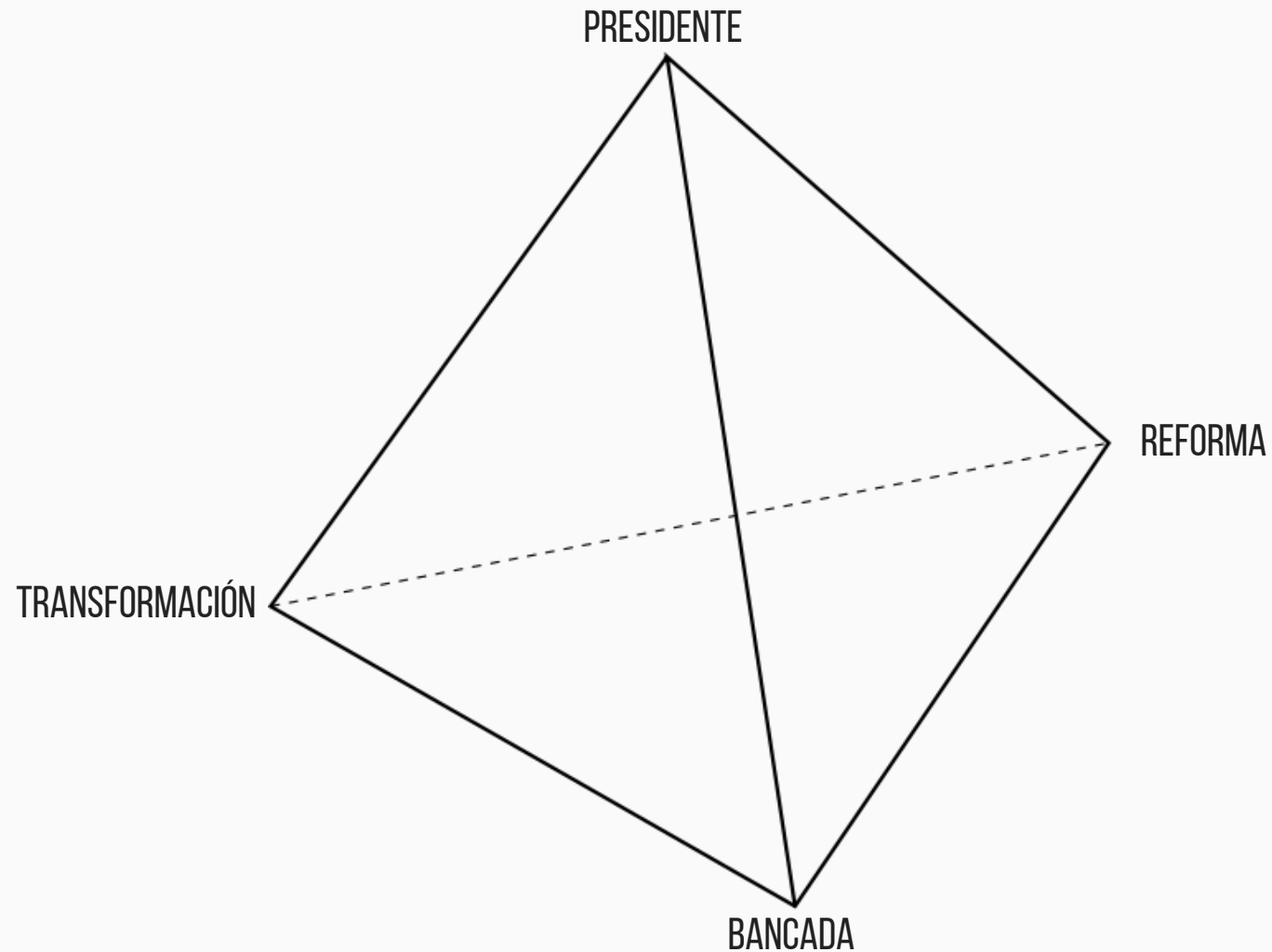
Distribución Documentos-Tópicos



Distribución Tópicos-Palabras

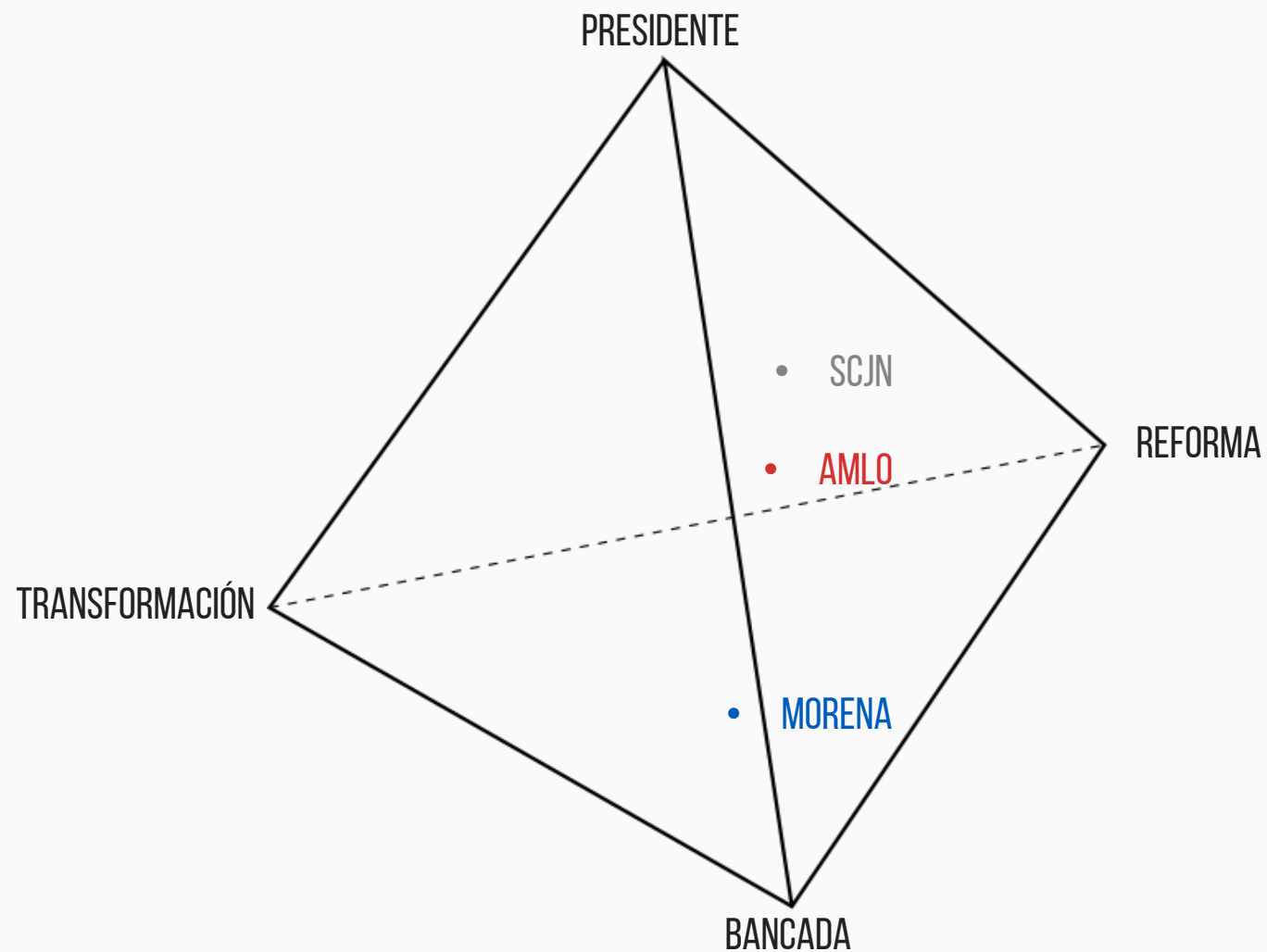


Distribución Tópicos-Palabras



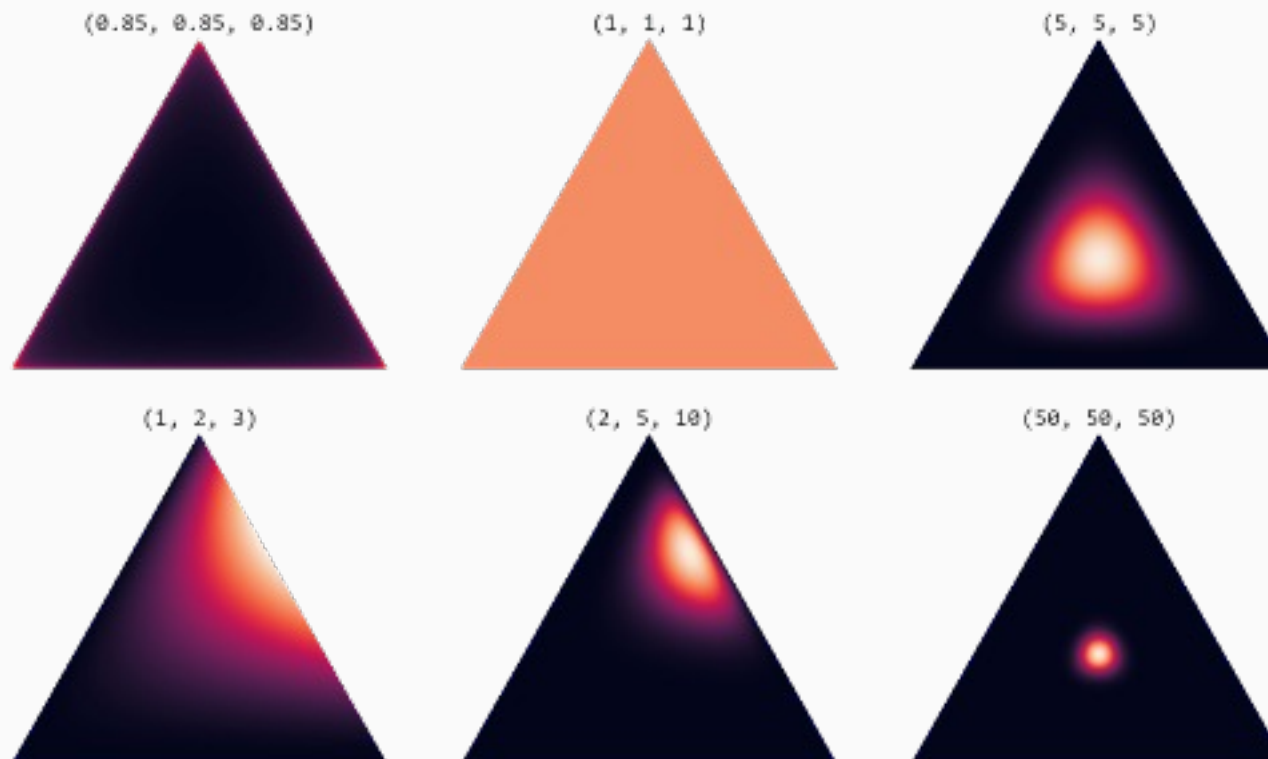
- SCJN
- AMLO
- MORENA

Distribución Tópicos-Palabras



Distribución de Dirichlet

“Beta Multinomial”

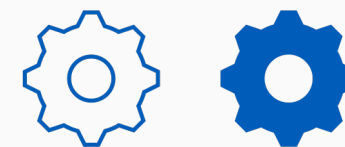
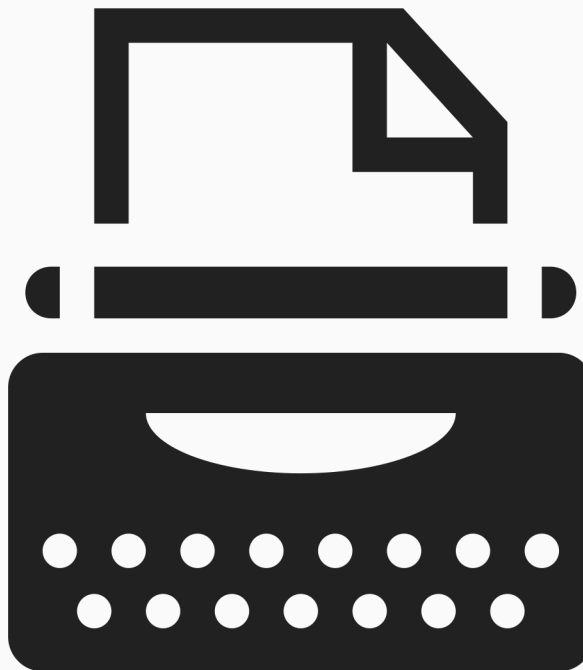


Modelado de Tópicos

Topic Modeling



Configuración



Engranaje

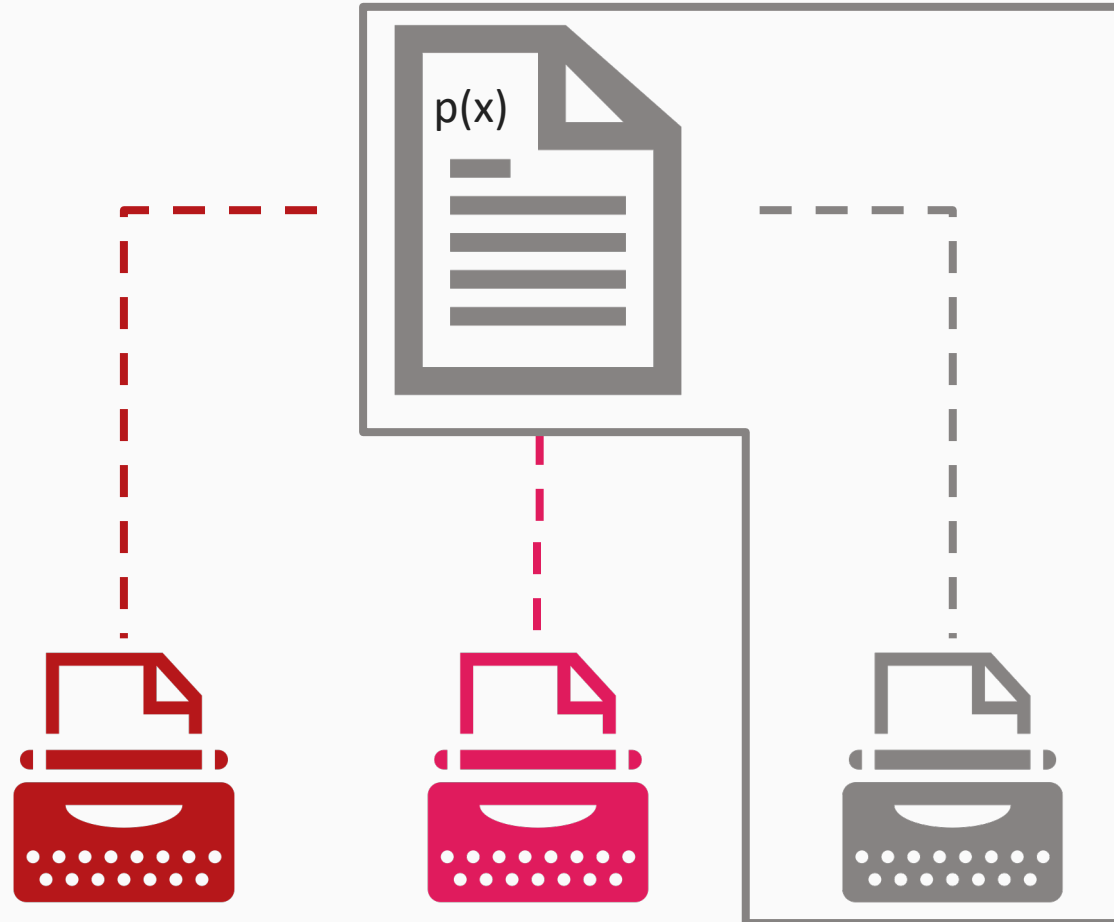
Asignación Latente de Dirichlet

Latent Dirichlet Allocation (LDA)



Asignación Latente de Dirichlet

Latent Dirichlet Allocation (LDA)



Modelo Generativo del LDA

LDA Generative Approach

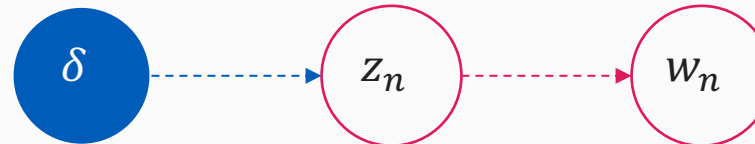
- Cada tópico $k \in \{1, \dots, K\}$ es una distribución multinomial de palabras θ_k , que a su vez proviene de una distribución de Dirichlet con parámetro λ .



- Cada documento $d \in \{1, \dots, D\}$ distribución multinomial de tópicos que δ_d , que a su vez proviene de una distribución de Dirichlet con parámetro α .



- Cada coordenada de una palabra $n \in \{1, \dots, N\}$ pertenece a un tópico latente (“escondido”) z_n que proviene de la distribución multinomial parametrizada por δ .

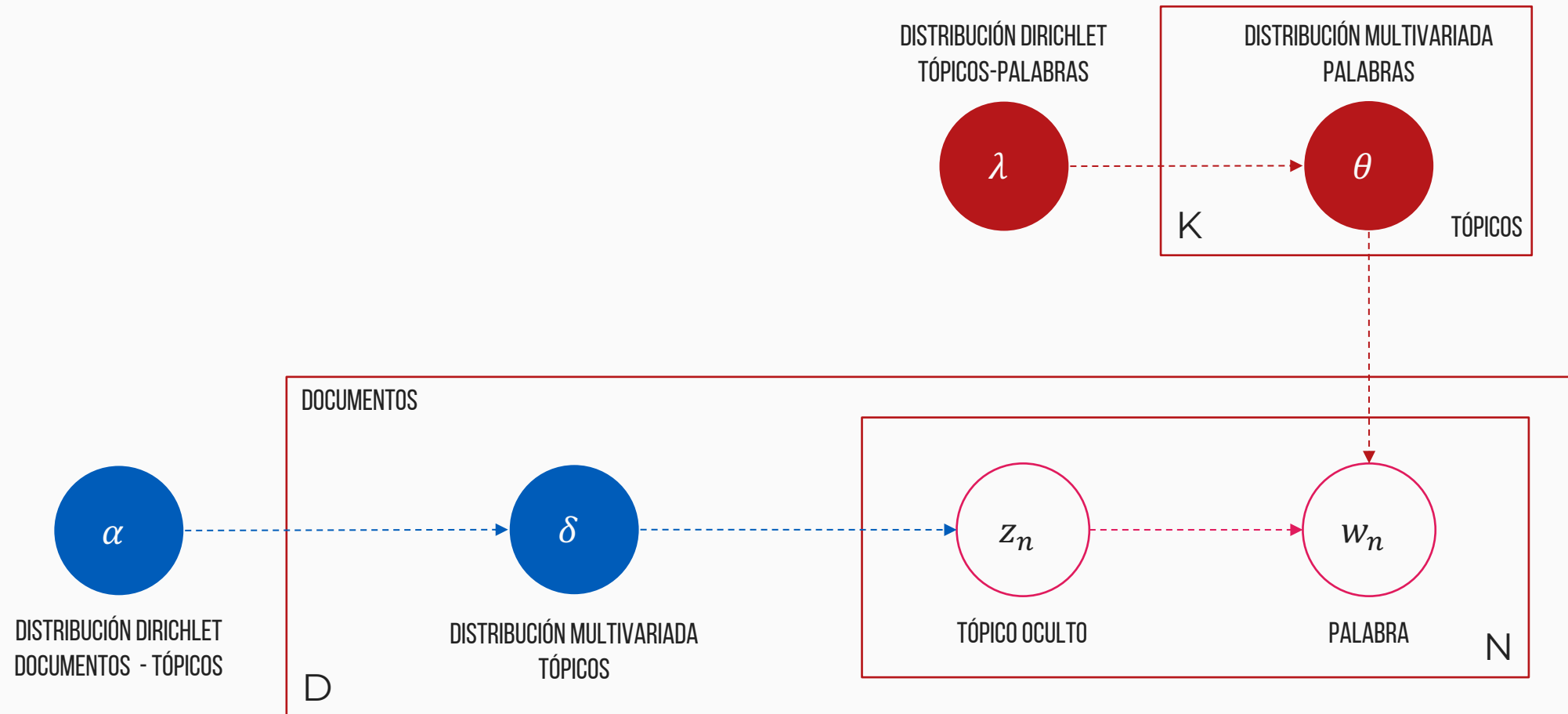


- Cada palabra observada w_n proviene de la distribución θ_{z_n}



Asignación Latente de Dirichlet

Latent Dirichlet Allocation (LDA)

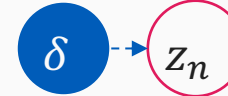
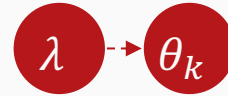
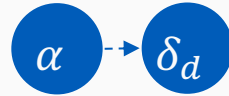


La Función de Probabilidad

$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{z_n} | \theta_{z_n})$$

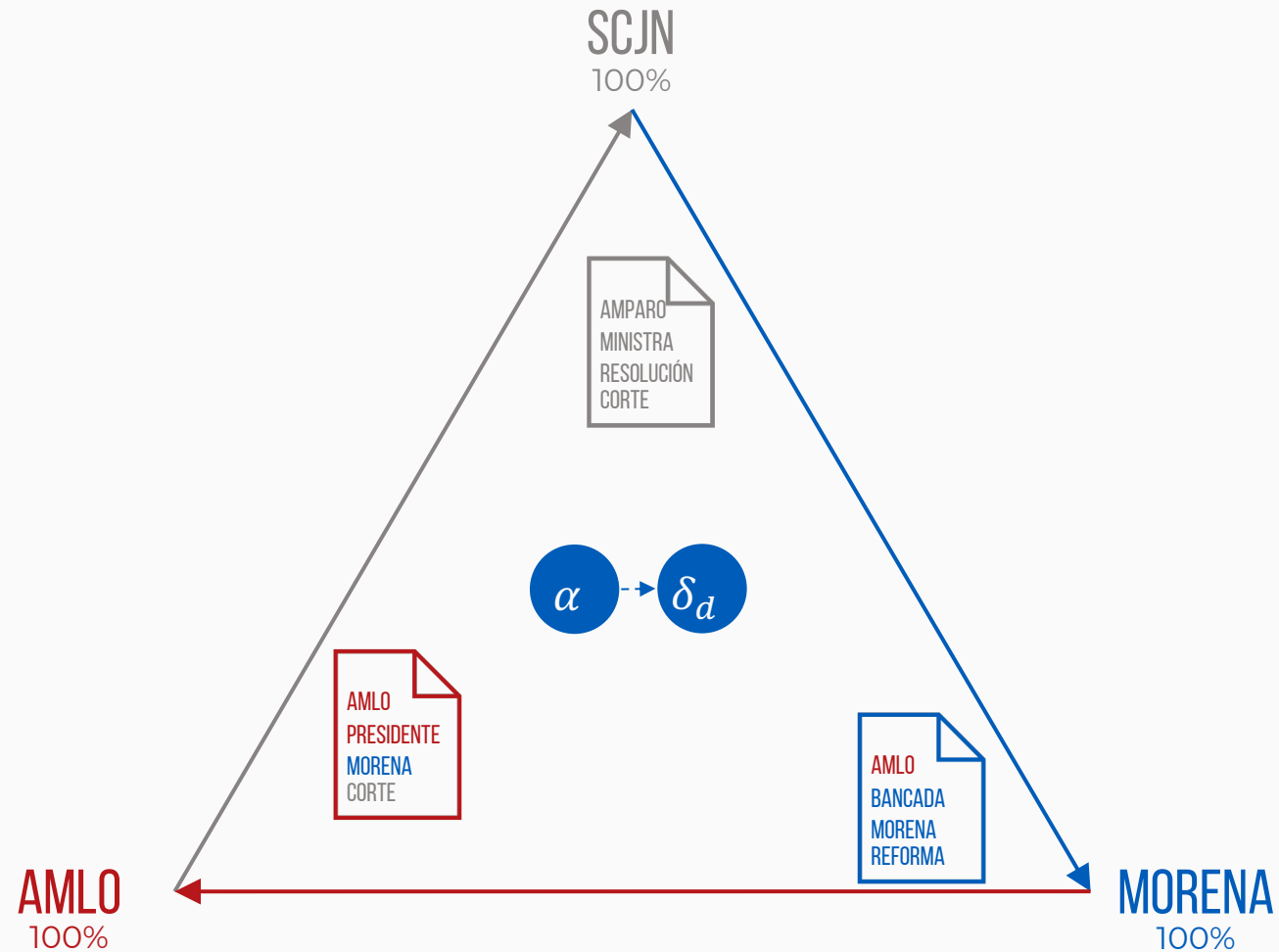
La Función de Probabilidad

$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



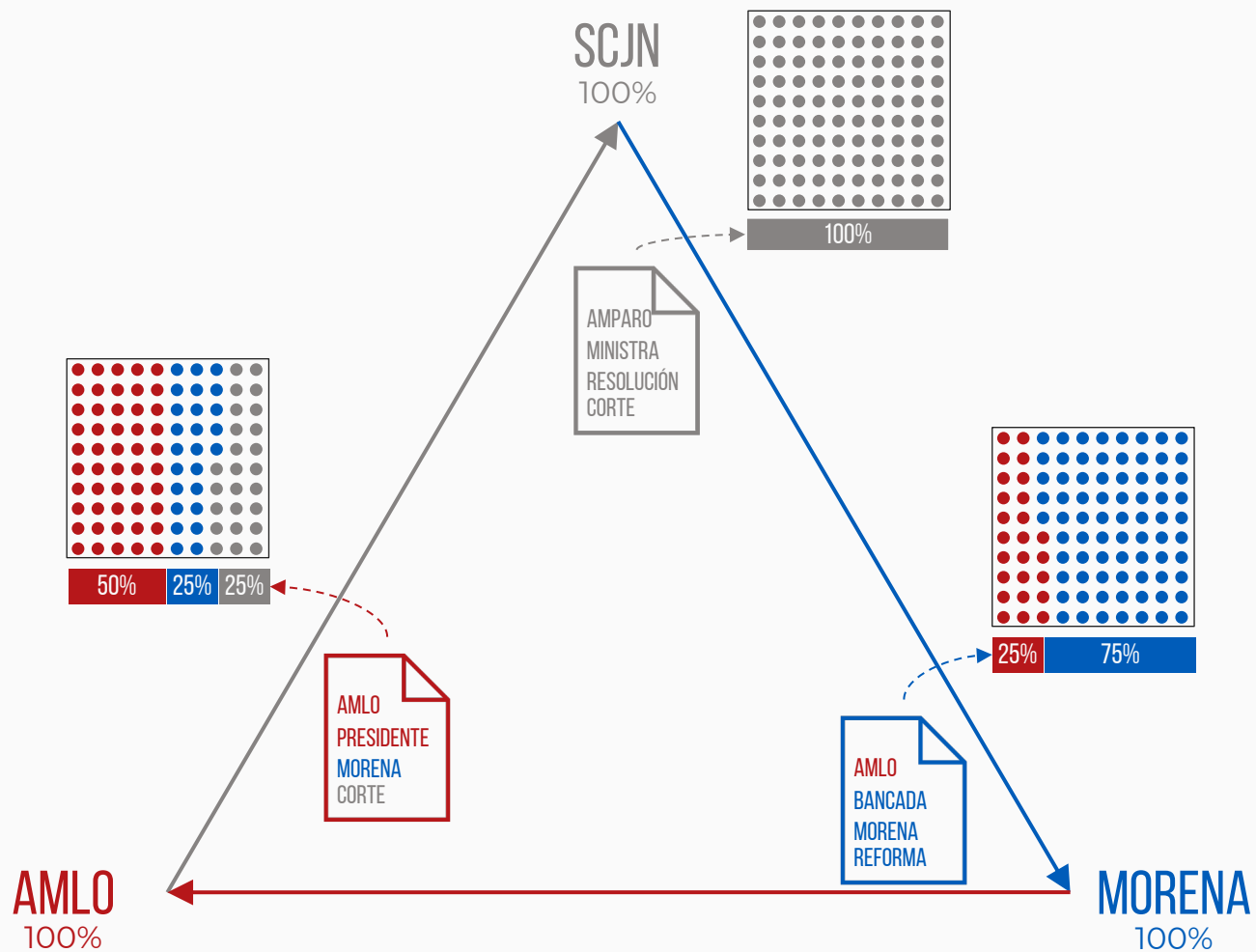
Documentos-Tópicos

Distribución de Dirichlet



Documentos-Tópicos

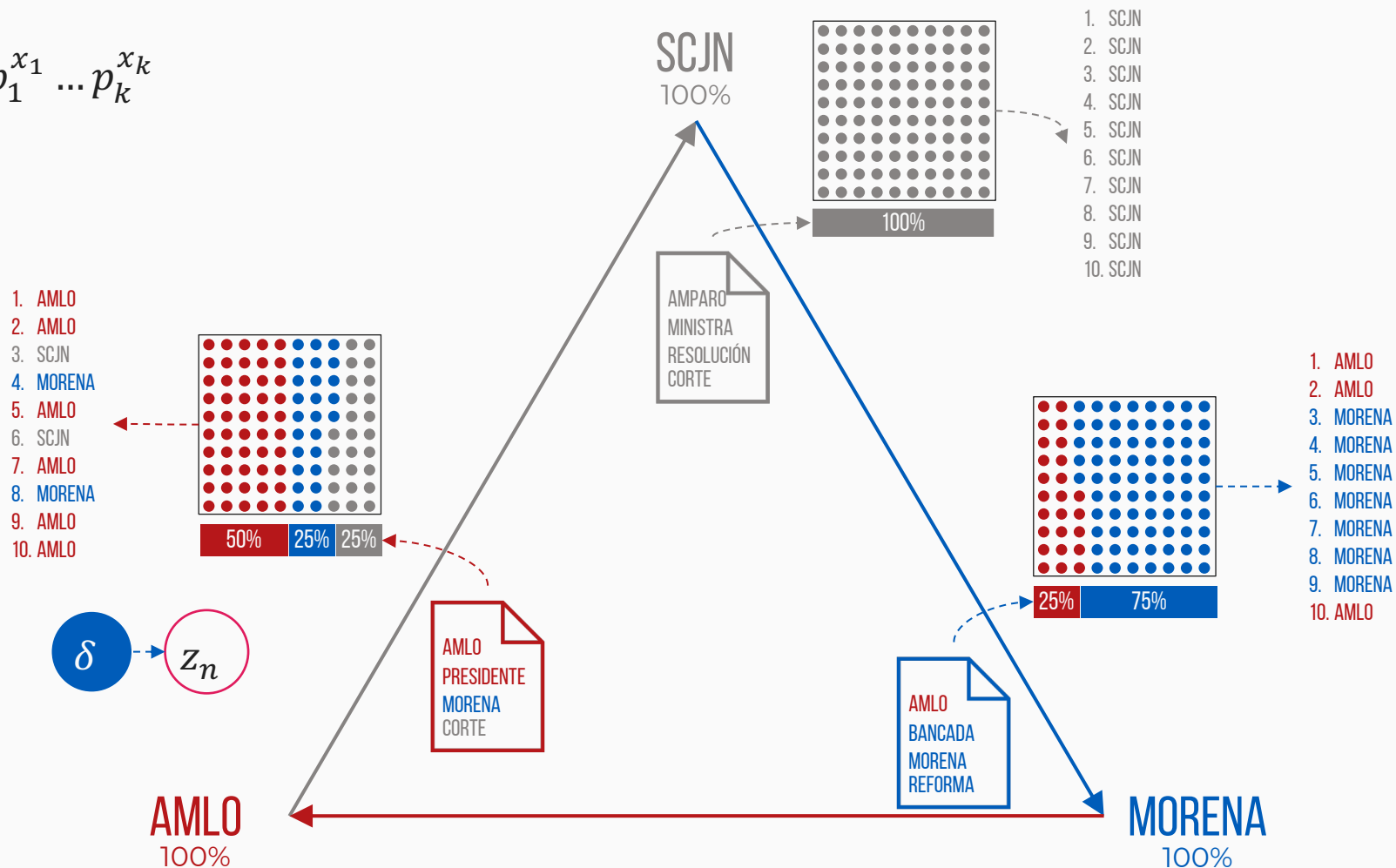
Distribución Multinomial



Documentos-Tópicos

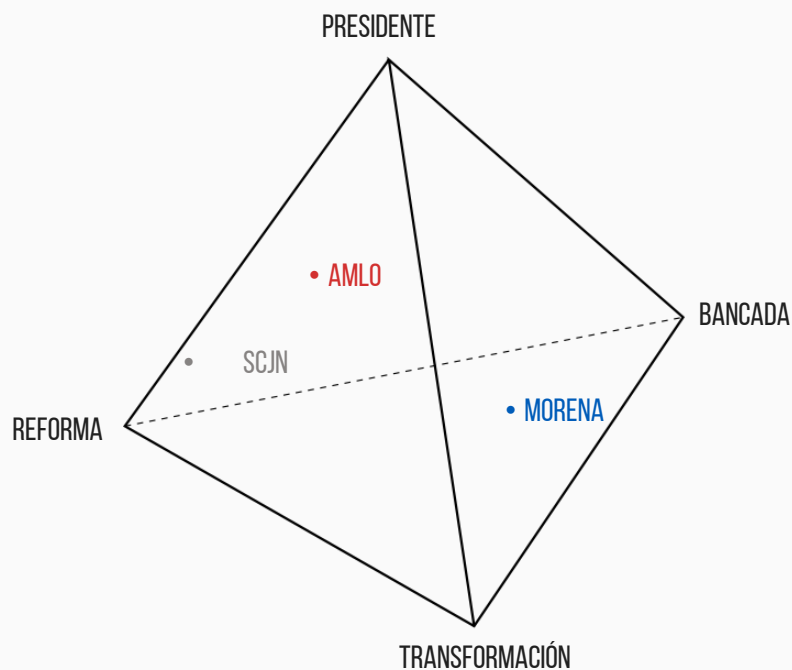
Distribución Multinomial

$$\text{PDF} = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}$$



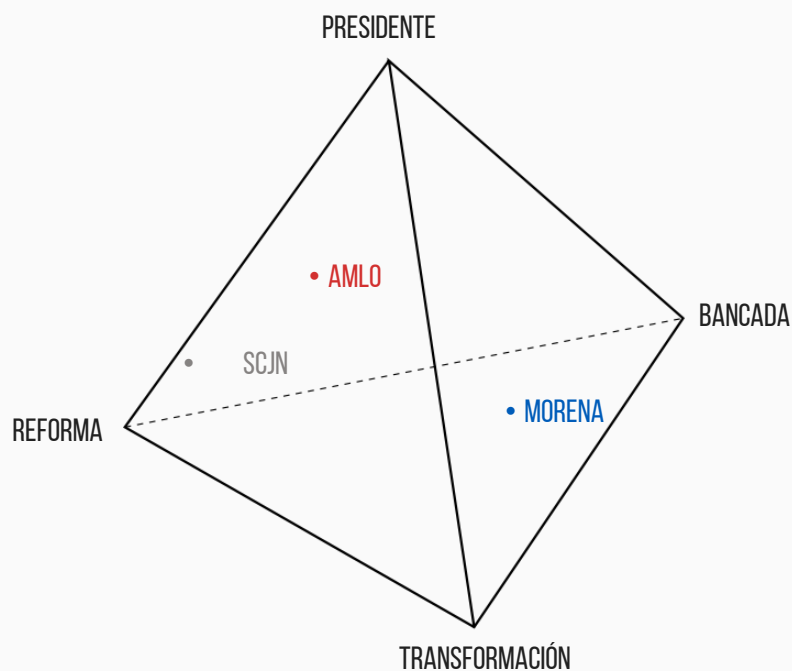
Distribución Tópicos-Palabras

Distribución de Dirichlet



Distribución Tópicos-Palabras

Distribución Multinomial



$$\theta_{z_n} \rightarrow w_n$$

AMLO

PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2

PRESIDENTE	PRESIDENTE
TRANSFORMACIÓN	TRANSFORMACIÓN
PRESIDENTE	PRESIDENTE
REFORMA	PRESIDENTE
REFORMA	TRANSFORMACIÓN

MORENA

PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2

TRANSFORMACIÓN	BANCADA
PRESIDENTE	TRANSFORMACIÓN
BANCADA	REFORMA
PRESIDENTE	BANCADA
REFORMA	TRANSFORMACIÓN

SCJN

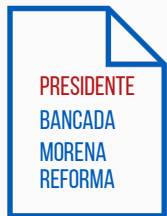
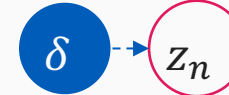
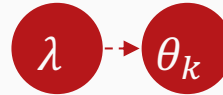
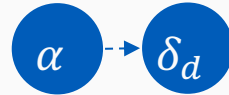
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7

REFORMA	PRESIDENTE
REFORMA	REFORMA
PRESIDENTE	REFORMA
REFORMA	PRESIDENTE
REFORMA	REFORMA

El Modelo Generativo

LDA Topic Modeling

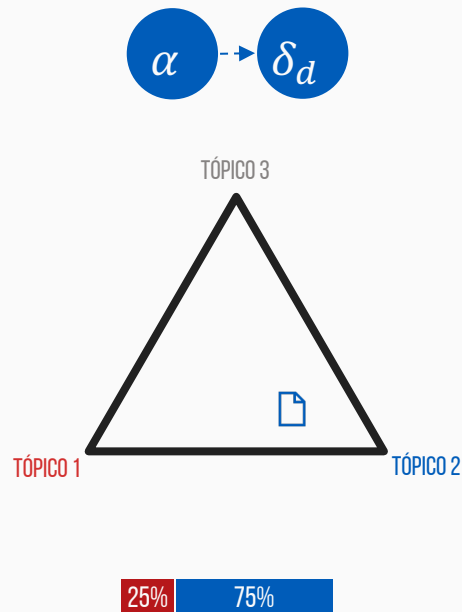
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



El Modelo Generativo

LDA Topic Modeling

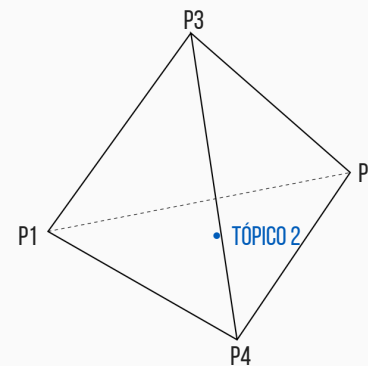
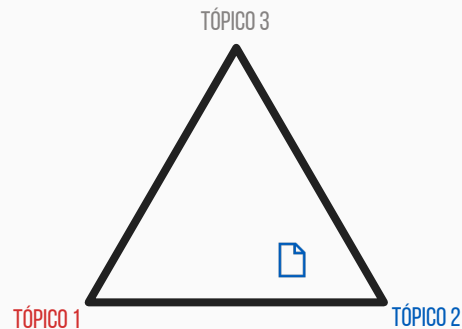
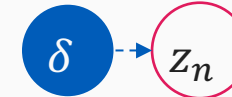
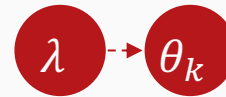
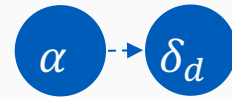
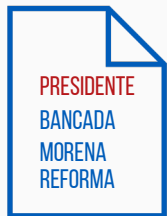
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



El Modelo Generativo

LDA Topic Modeling

$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$

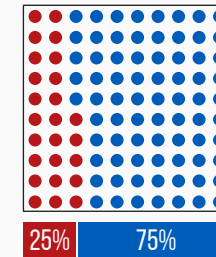
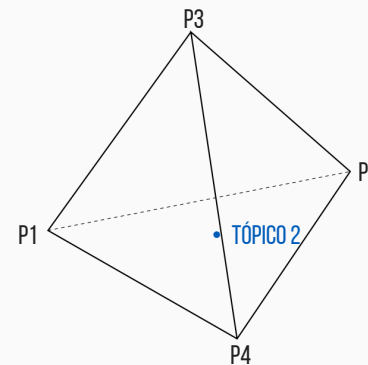
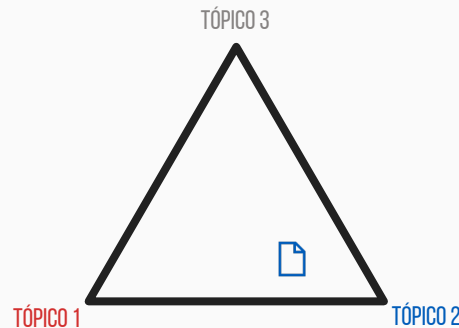
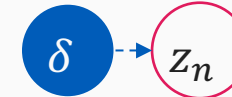
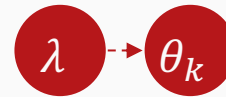
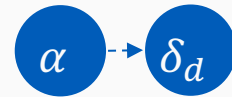


PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7

El Modelo Generativo

LDA Topic Modeling

$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$

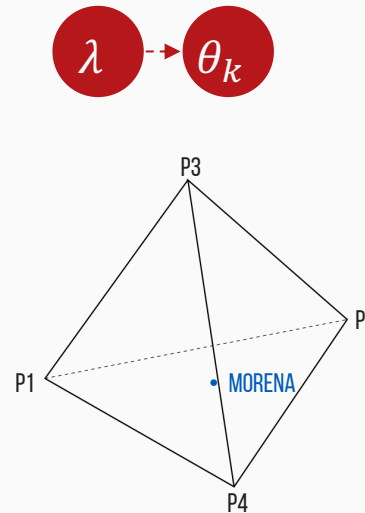
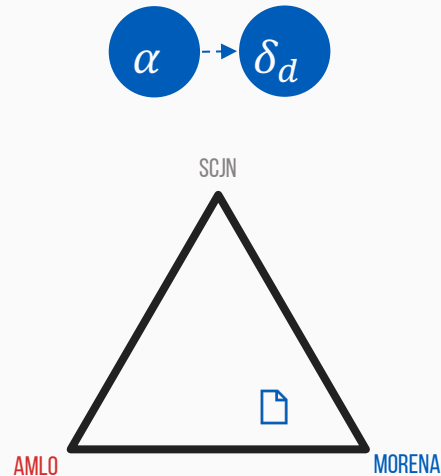
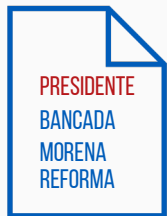


PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7

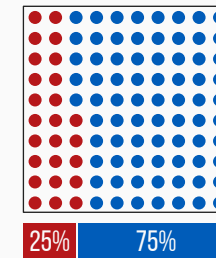
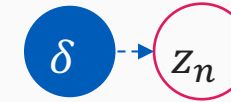
El Modelo Generativo

LDA Topic Modeling

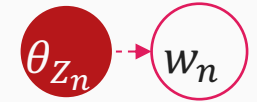
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7



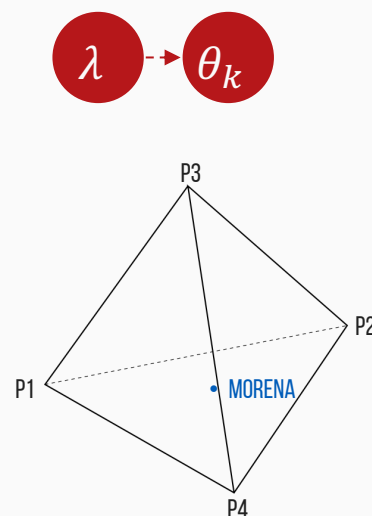
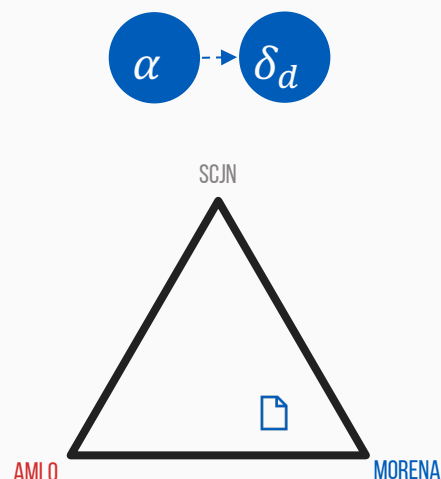
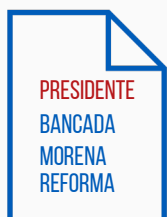
MORENA



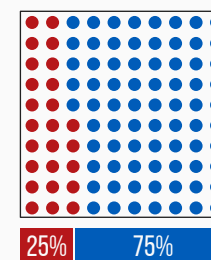
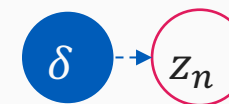
El Modelo Generativo

LDA Topic Modeling

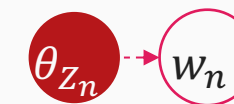
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7



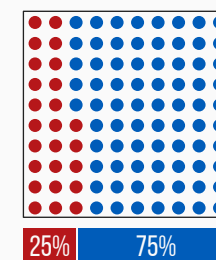
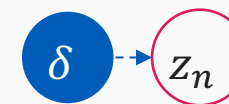
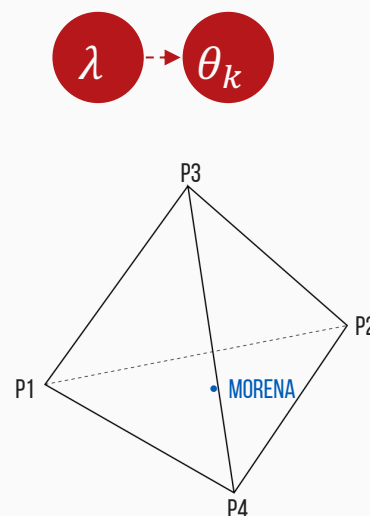
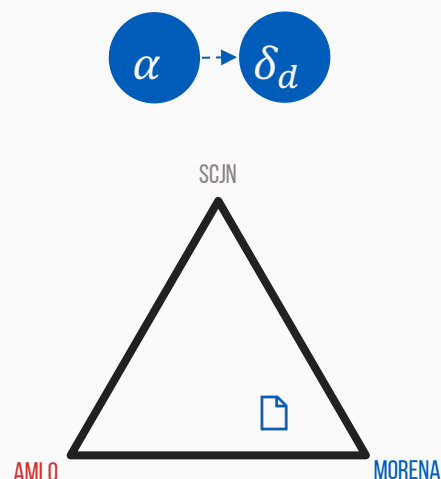
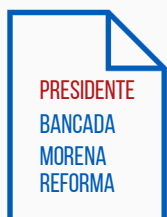
MORENA
MORENA



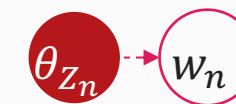
El Modelo Generativo

LDA Topic Modeling

$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{z_n} | \theta_{z_n})$$



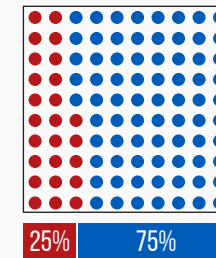
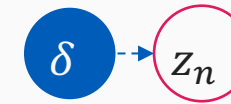
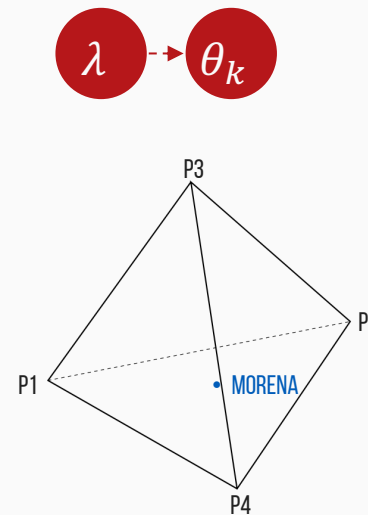
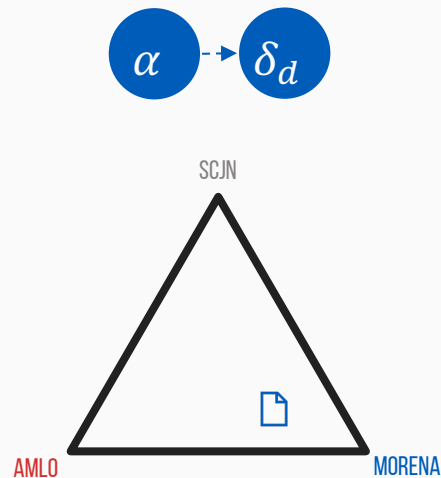
MORENA
MORENA
MORENA



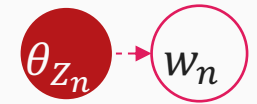
El Modelo Generativo

LDA Topic Modeling

$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



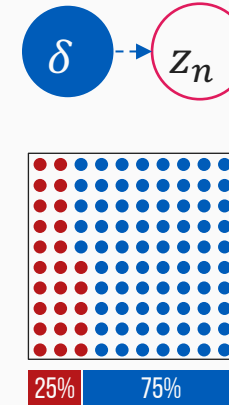
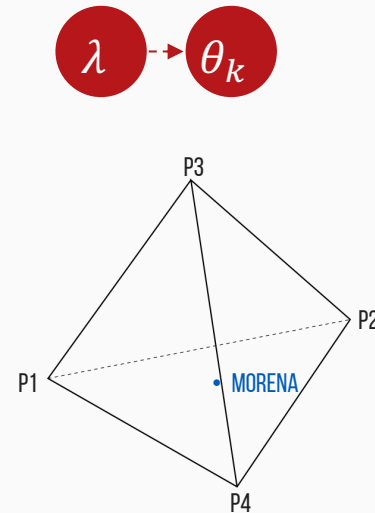
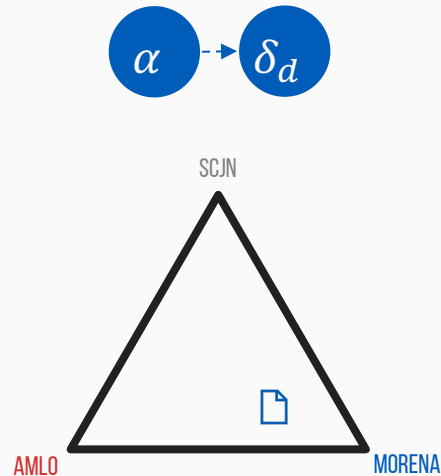
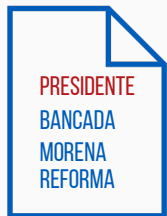
MORENA
MORENA
MORENA
AMLO



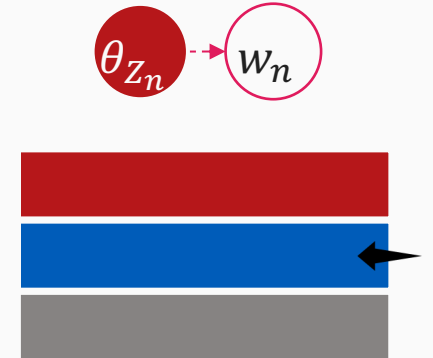
El Modelo Generativo

LDA Topic Modeling

$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



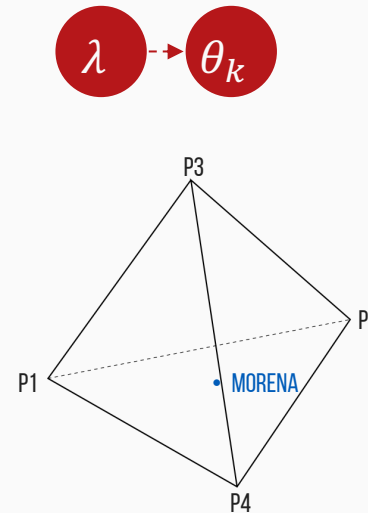
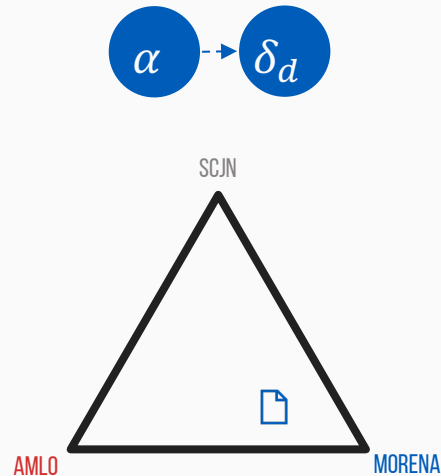
MORENA
MORENA
MORENA
AMLO
AMLO
MORENA
MORENA
AMLO
MORENA
MORENA



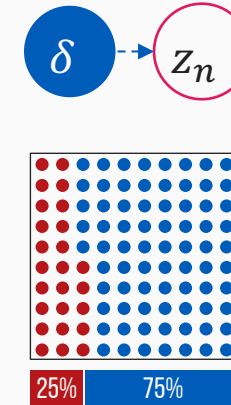
El Modelo Generativo

LDA Topic Modeling

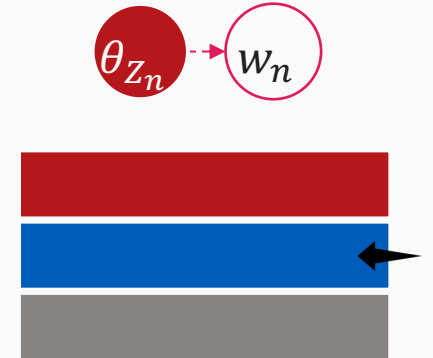
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7



MORENA
MORENA
MORENA
AMLO
AMLO
MORENA
MORENA
AMLO
MORENA
MORENA

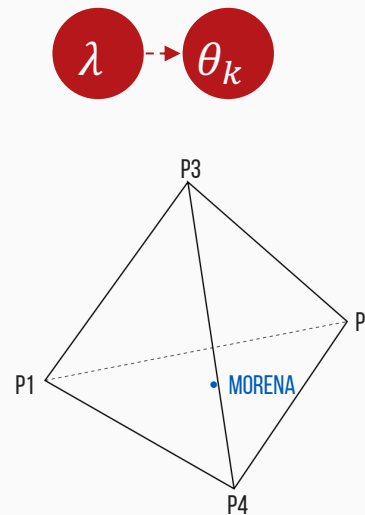
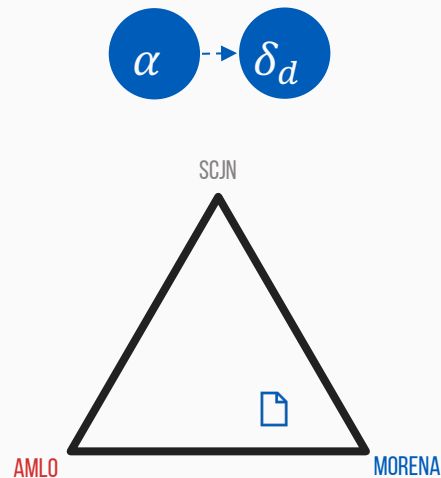
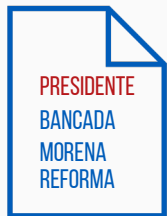


REFORMA

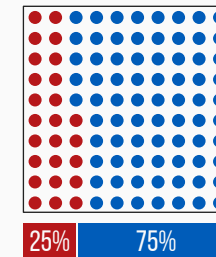
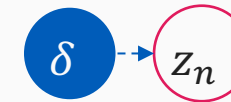
El Modelo Generativo

LDA Topic Modeling

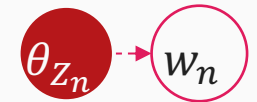
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7



MORENA
MORENA
MORENA
AMLO
AMLO
MORENA
MORENA
AMLO
MORENA
MORENA



REFORMA
BANCADA

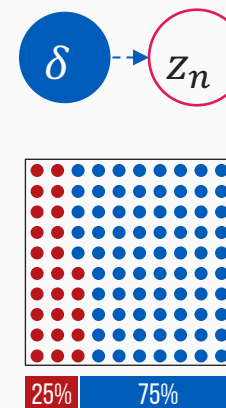
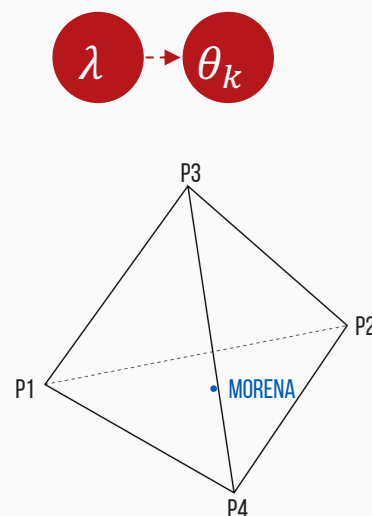
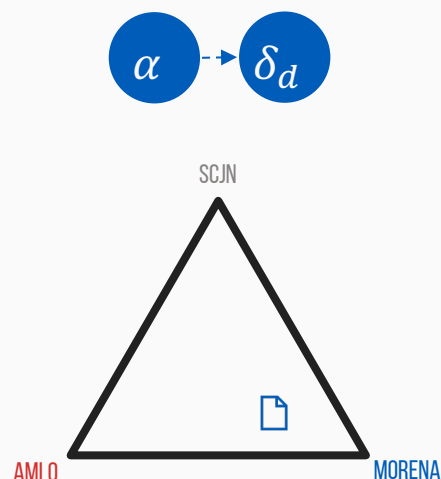
El Modelo Generativo

LDA Topic Modeling

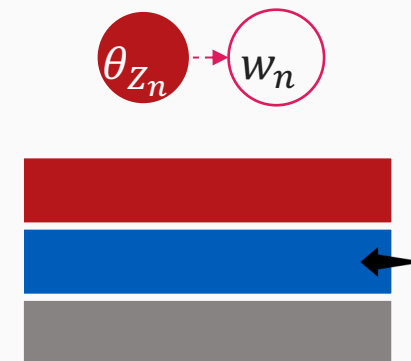
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



PRESIDENTE
BANCADA
MORENA
REFORMA



MORENA
MORENA
MORENA
AMLO
AMLO
MORENA
MORENA
AMLO
MORENA
MORENA



REFORMA
BANCADA
TRANSFORMACIÓN

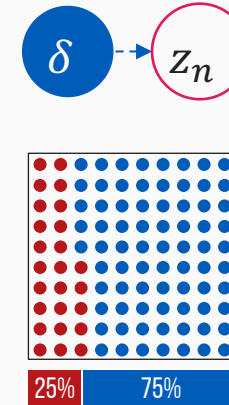
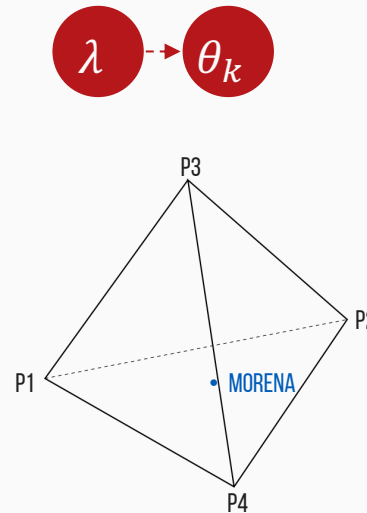
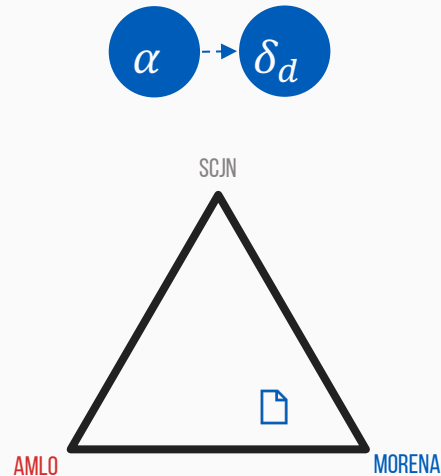
El Modelo Generativo

LDA Topic Modeling

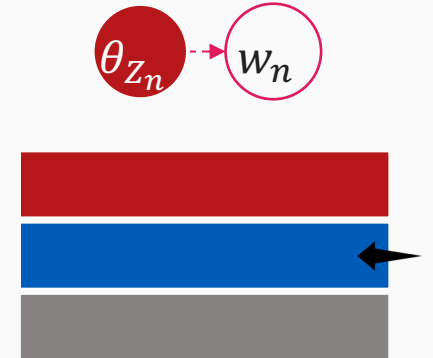
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



PRESIDENTE
BANCADA
MORENA
REFORMA



MORENA
MORENA
MORENA
AMLO
AMLO
MORENA
MORENA
AMLO
MORENA
MORENA



REFORMA
BANCADA
TRANSFORMACIÓN
TRANSFORMACIÓN

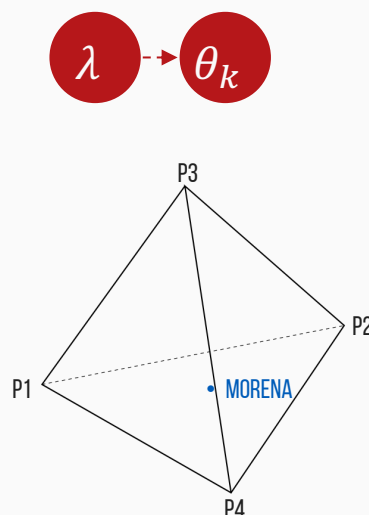
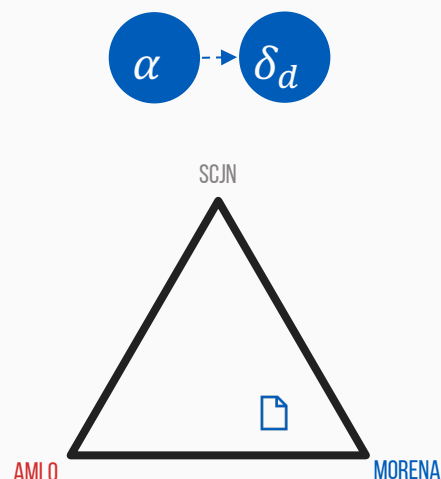
El Modelo Generativo

LDA Topic Modeling

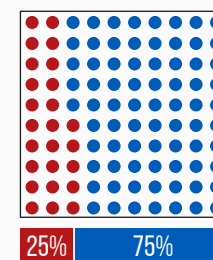
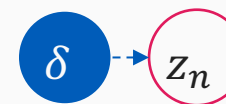
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{z_n} | \theta_{z_n})$$



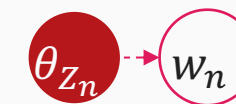
PRESIDENTE
BANCADA
MORENA
REFORMA



PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7



MORENA
MORENA
MORENA
AMLO
AMLO
MORENA
MORENA
AMLO
MORENA
MORENA

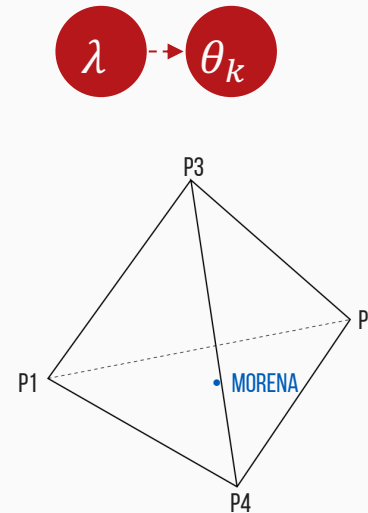
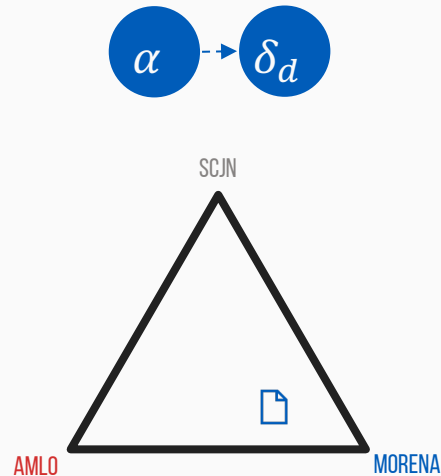


REFORMA
BANCADA
TRANSFORMACIÓN
TRANSFORMACIÓN
PRESIDENTE
PRESIDENTE
REFORMA
PRESIDENTE
TRANSFORMACIÓN
TRANSFORMACIÓN

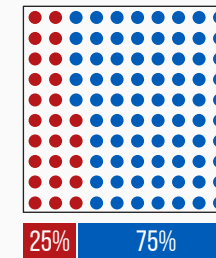
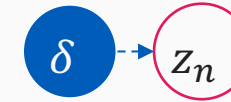
El Modelo Generativo

LDA Topic Modeling

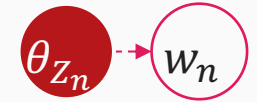
$$P(W, Z, \delta, \theta; \alpha, \lambda) = \prod_{d=1}^D P(\delta_d; \alpha) \prod_{k=1}^K P(\theta_k; \lambda) \prod_{n=1}^N P(Z_n | \delta_d) P(W_{Z_n} | \theta_{Z_n})$$



PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.5	0.3	0.0	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.2	0.3	0.3	0.2
PRESIDENTE	TRANSFORMACIÓN	BANCADA	REFORMA
0.3	0.0	0.0	0.7



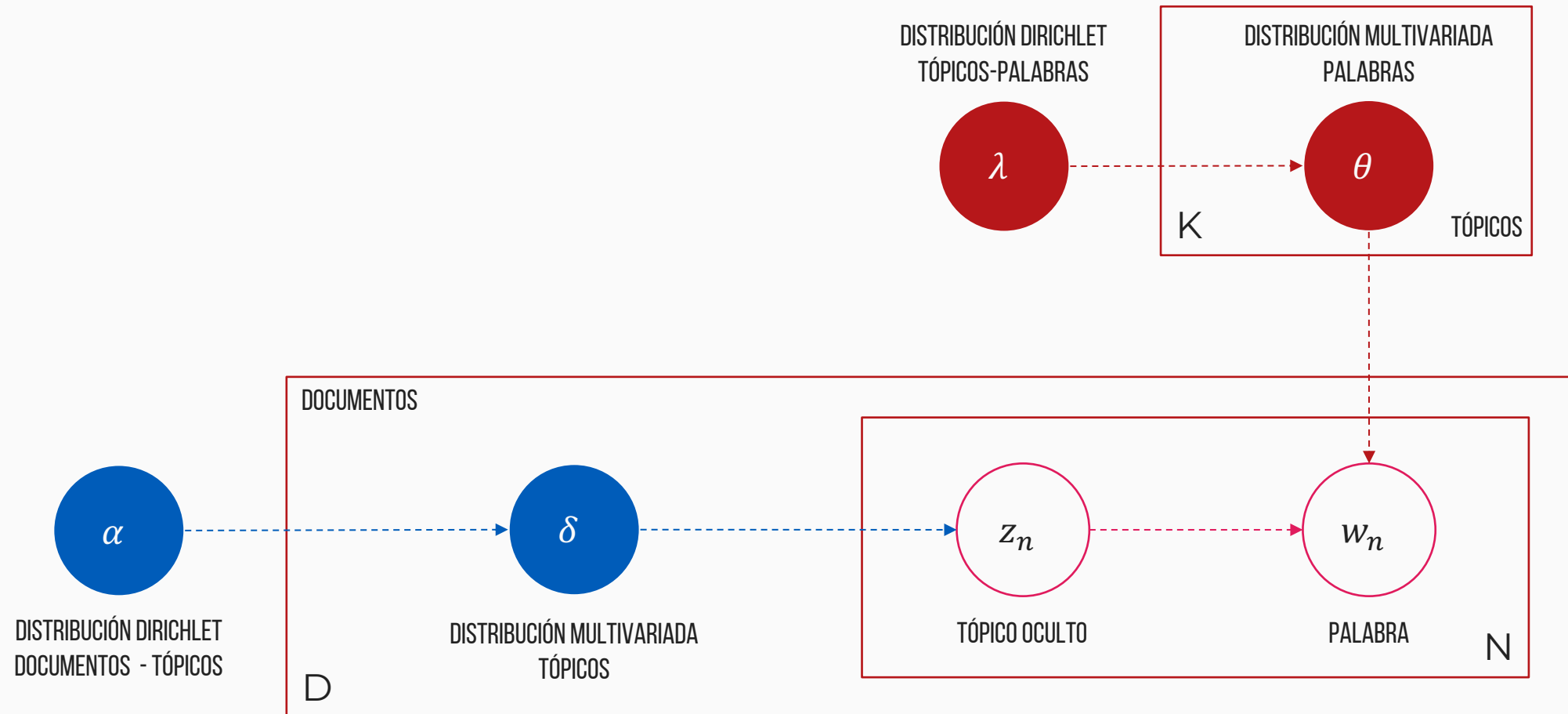
MORENA
MORENA
MORENA
AMLO
AMLO
MORENA
MORENA
AMLO
MORENA
MORENA



REFORMA
BANCADA
TRANSFORMACIÓN
TRANSFORMACIÓN
PRESIDENTE
PRESIDENTE
REFORMA
PRESIDENTE
TRANSFORMACIÓN
TRANSFORMACIÓN

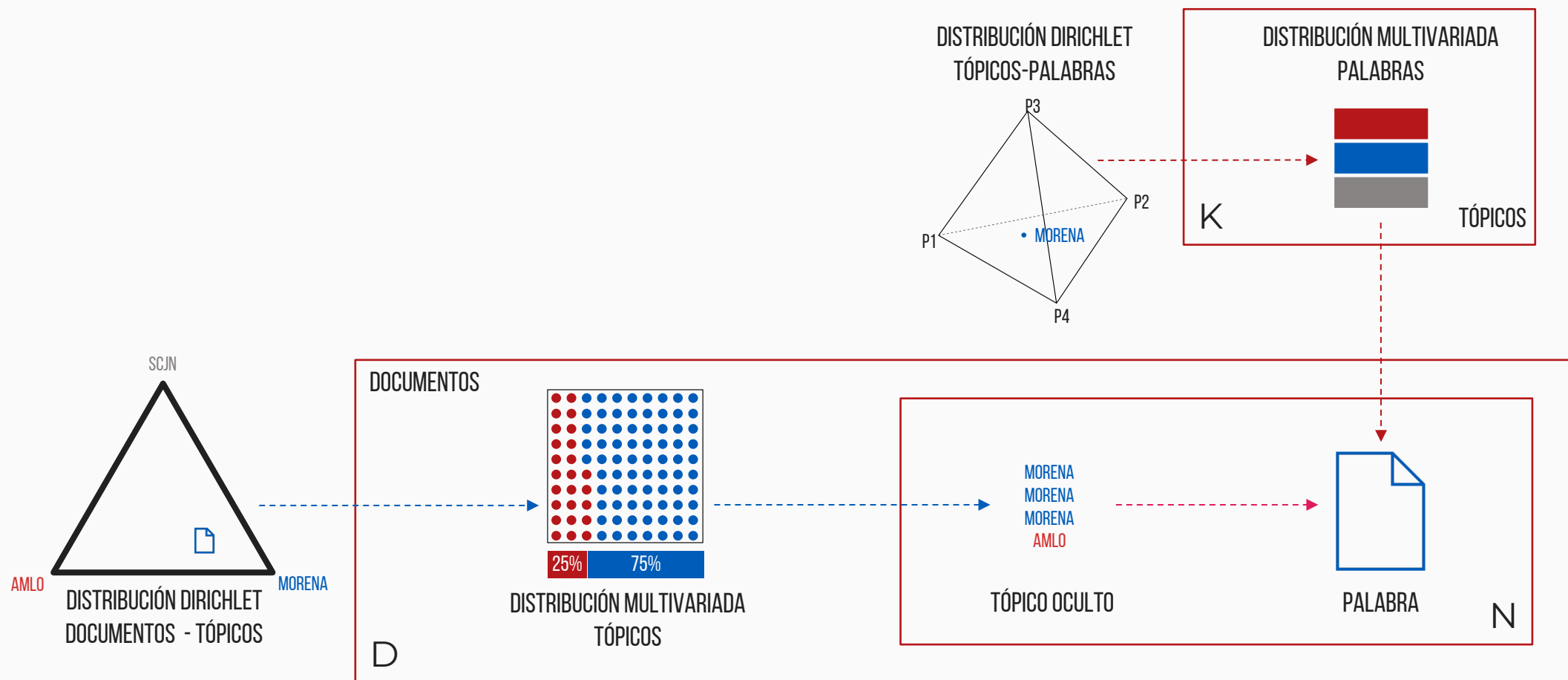
Asignación Latente de Dirichlet

Latent Dirichlet Allocation (LDA)



Asignación Latente de Dirichlet

Latent Dirichlet Allocation (LDA)



En resumen

