

13 Giugno 2018

Data analytics per l'industria italiana

David Vannozzi - Direttore Generale



Le attività di Cineca

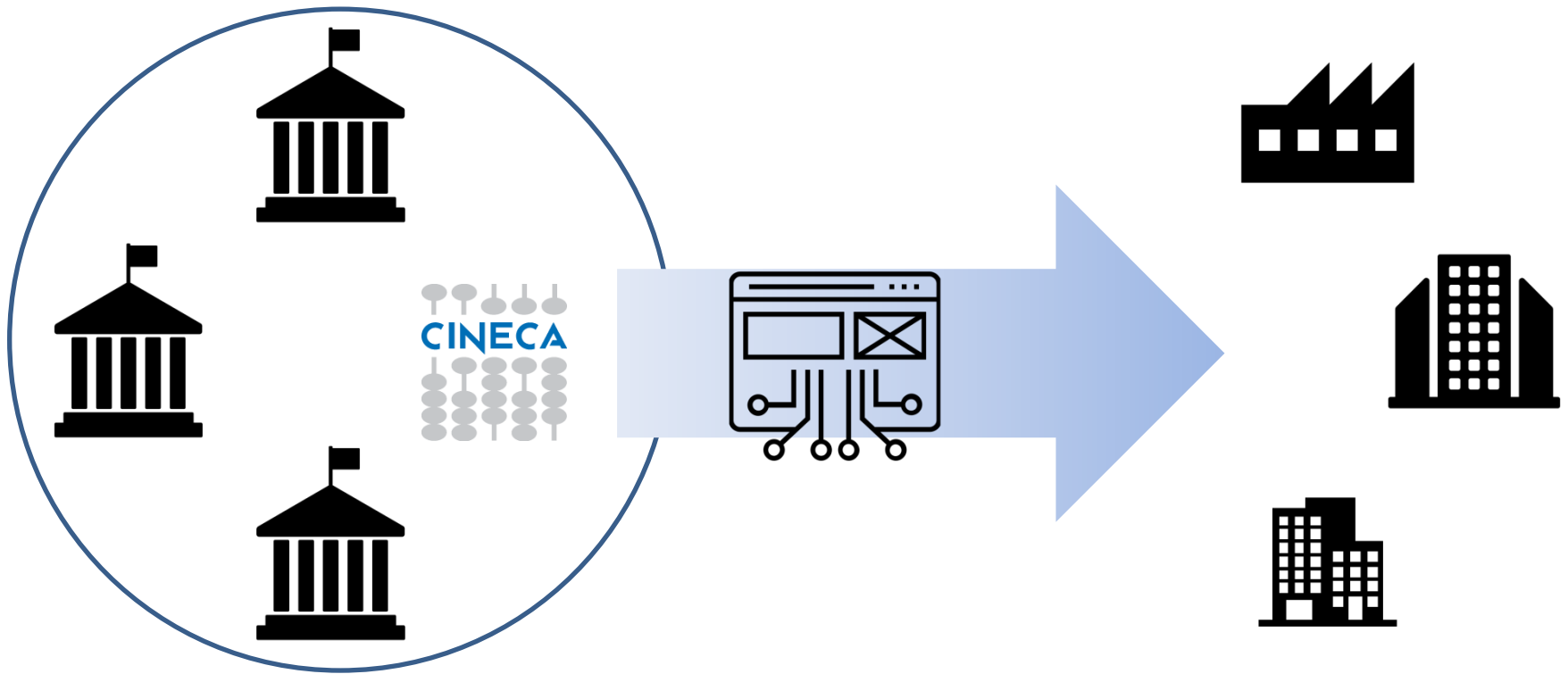
Supercalcolo per la comunità scientifica

Sistemi gestionali e servizi a sostegno delle Università e del MIUR

Trasferimento tecnologico alle imprese e alla pubblica amministrazione



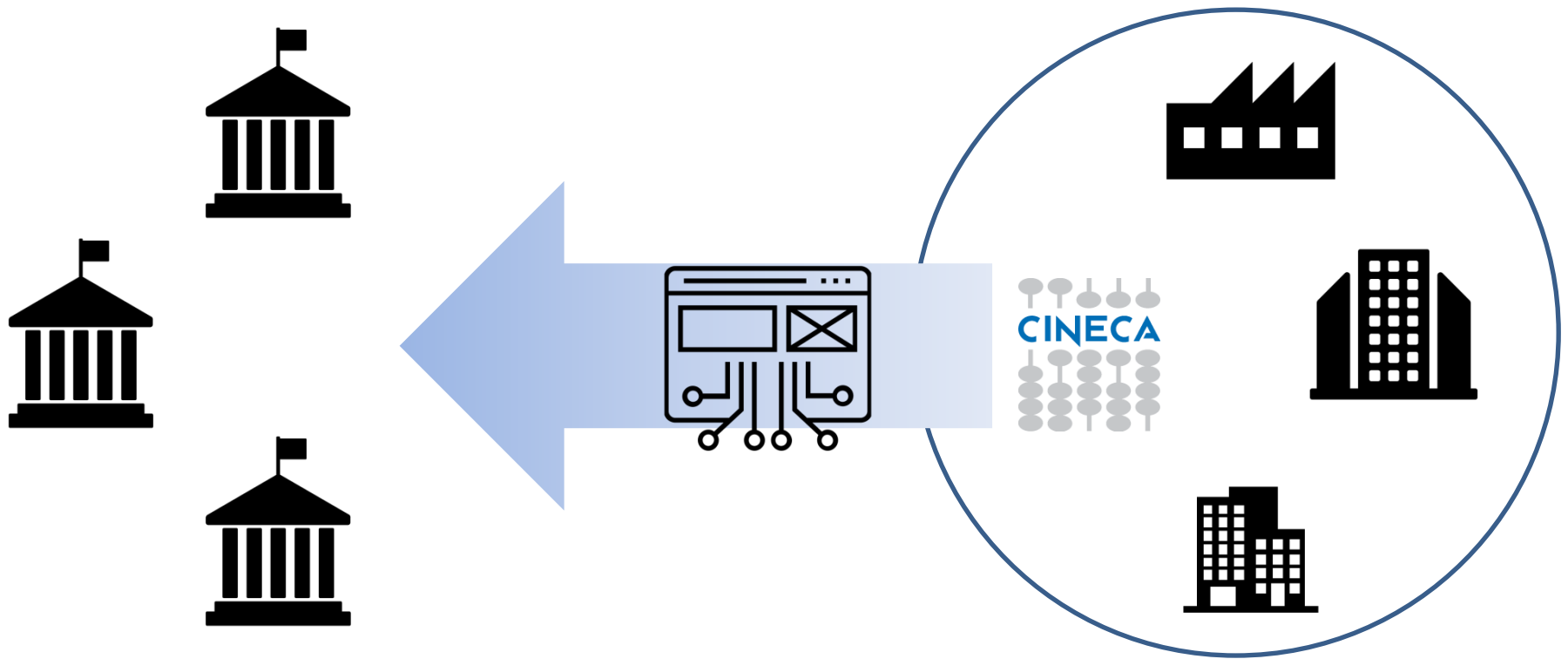
Trasferimento tecnologico



PAI è trasferimento tecnologico

La tecnologia sviluppata con i consorziati viene messa al servizio delle altre Pubbliche Amministrazioni e del tessuto produttivo del paese

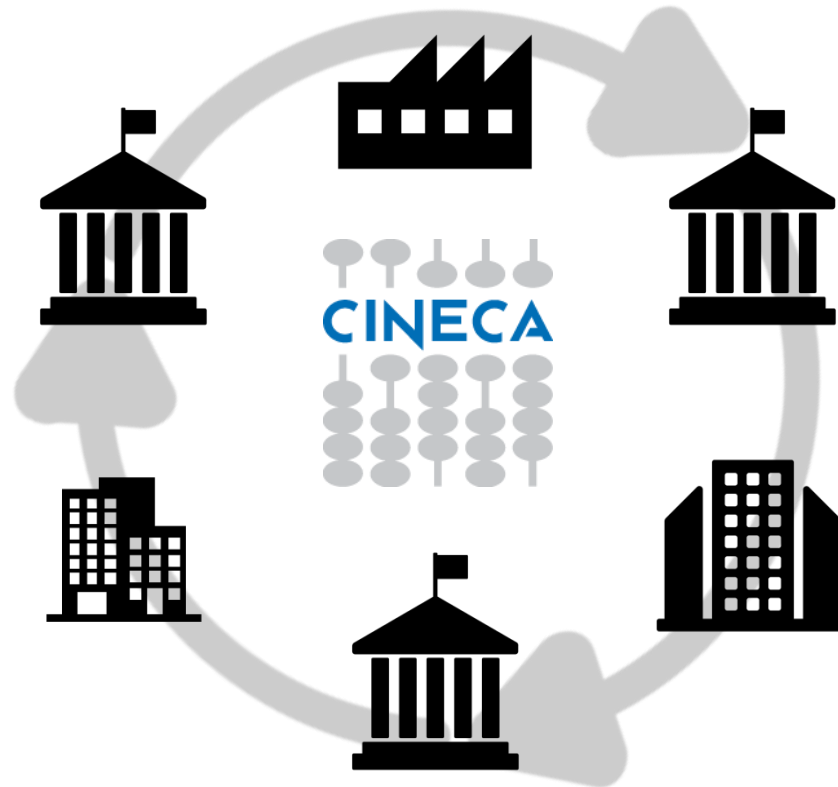
Sviluppo innovativo



PAI è sviluppo innovativo

L'interazione con Pubbliche Amministrazioni e Industria porta Cineca a esplorare nuove tecnologie e applicazioni, che possono poi essere portate a vantaggio dei consorziati.

Interazione virtuosa



PAI è interazione virtuosa

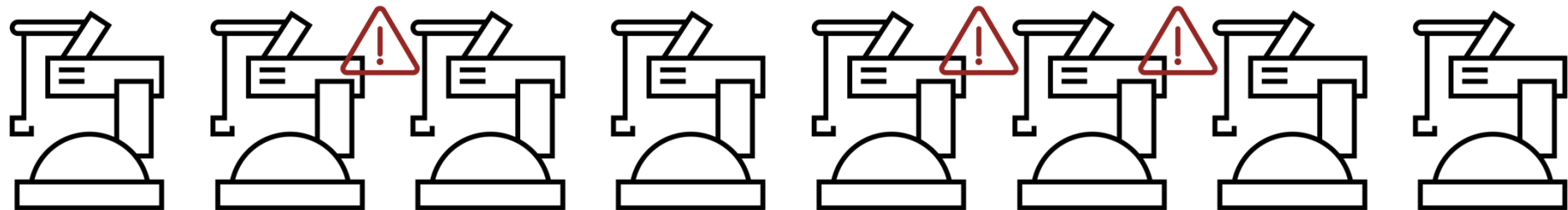
PAI funge da snodo abilitante di un'interazione virtuosa tra sistema accademico amministrativo e produttivo per il rinnovamento tecnologico del Paese

Progetto di ricerca su Data Analytics

- Avviato nel quarto quadrimestre del 2016
- Con una grande impresa italiana
- Un progetto di **ricerca congiunta**: per costruire una soluzione a un problema industriale con i metodi di Data Analytics
 - Non esiste una soluzione nota
 - L'unione delle specifiche expertise porta una crescita sia per il Cineca che per l'Azienda

Il problema

insert 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735
76536735	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345
stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356
insert 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735	insert 6735	remove 6735
76536735	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345	warning 76536735	43472345
stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356	stop alert 76356
43472345	warning 43472345	43472345	warning 43472345	43472345	warning 43472345	43472345	warning 43472345	43472345	warning 43472345	43472345	warning 43472345	43472345	warning 43472345	43472345	warning 43472345	43472345



Macchine industriali producono log non modificabili,
la cui correlazione con eventi di system failure nella catena di produzione non è
sempre nota

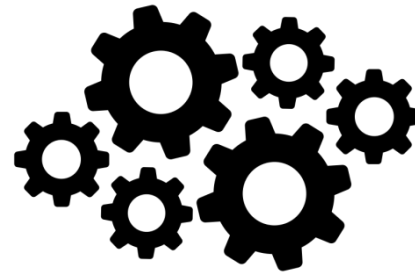
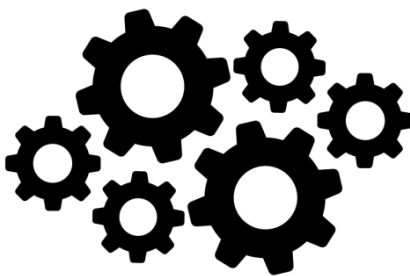
L'ipotesi



Individuare pattern di log associati ad un system failure su un determinato set di macchine, e un metodo replicabile su altri set di macchine e altri log

Il metodo

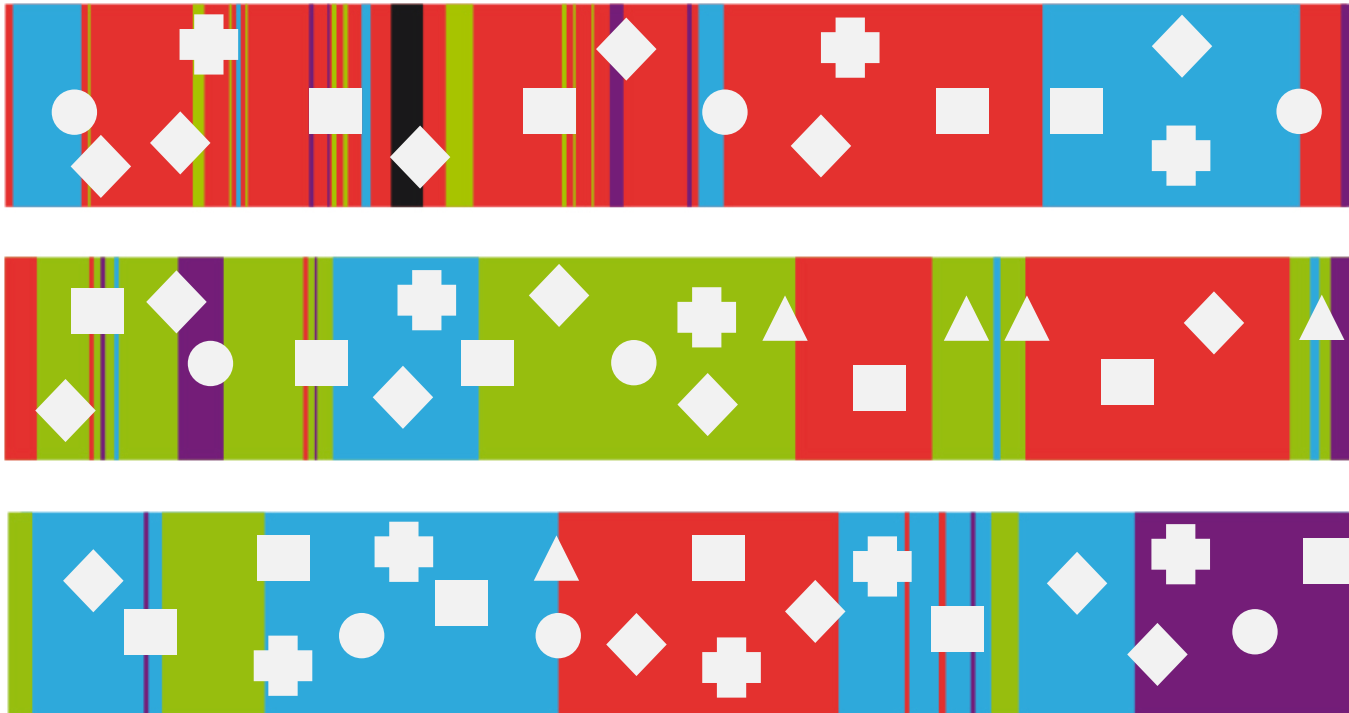
insert	6735	insertremove6735	insert	6735	insertremove6735	insert	6735	insertremove6735	insert	6735	insert	6735	insert	6735	insert	6735
76536735	warning	76536735	76536735	warning	76536735	76536735	warning	76536735	76536735	warning	76536735	76536735	warning	76536735	76536735	warning
stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert
insert	6735	insertremove6735	insert	6735	insertremove6735	insert	6735	insertremove6735	insert	6735	insert	6735	insert	6735	insert	6735
76536735	warning	76536735	76536735	warning	76536735	76536735	warning	76536735	76536735	warning	76536735	76536735	warning	76536735	76536735	warning
stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert	76356	stop alert
43472345	warning	43472345	43472345	warning	43472345	43472345	warning	43472345	43472345	warning	43472345	43472345	warning	43472345	43472345	warning



Lavorando in modalità **blind** applichiamo metodi di Data Analytics ai log delle macchine cercando correlazioni con gli stati di system failure.

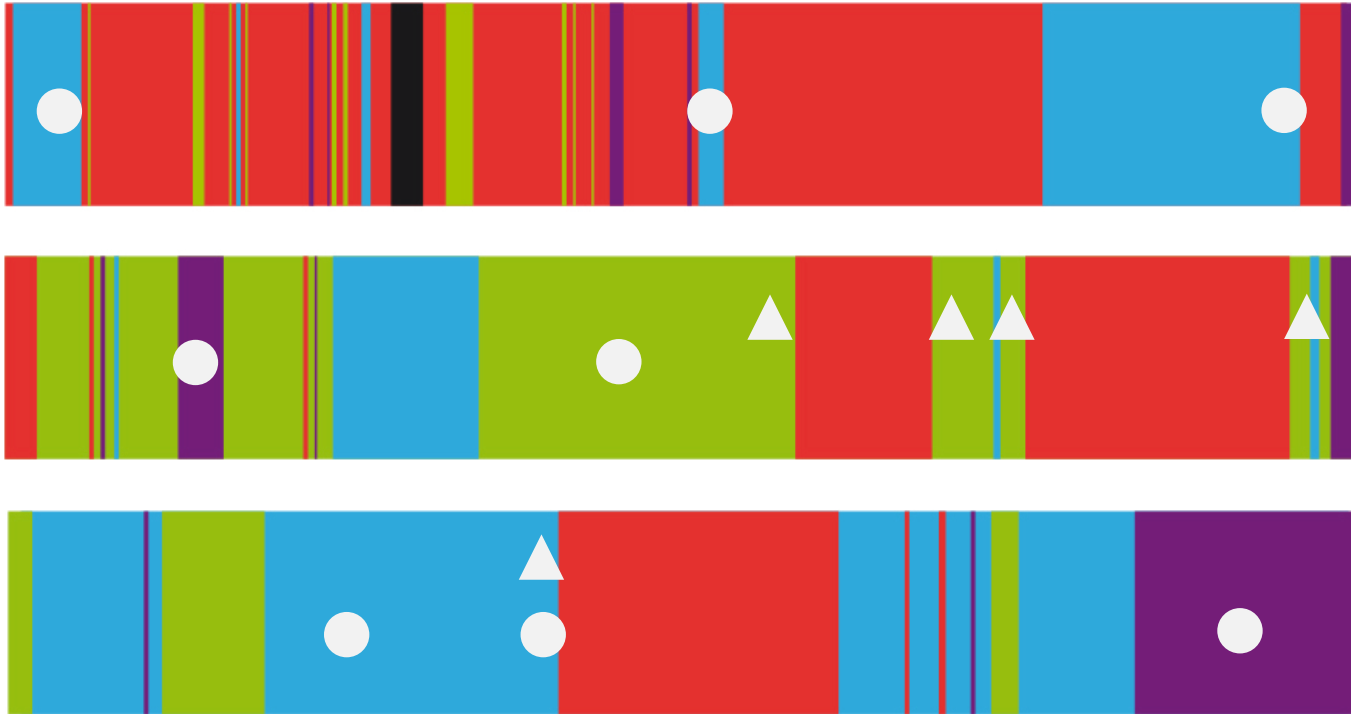
I Data Scientist Cineca **non hanno nessuna informazione** sulle macchine, sulla catena produttiva e sui messaggi di log. Si lavora in maniera puramente statistica, al fine di trovare un metodo che sia applicabile a prescindere dagli specifici messaggi, macchine e configurazioni.

Ricerca di pattern



Date informazioni sullo stato delle macchine ed i log, abbiamo ricercato pattern che permettessero di **associare determinati messaggi o sequenza di messaggi all'avvicinarsi di un evento di system failure**, in termini puramente statistici - senza entrare nella semantica dei messaggi o nel funzionamento e manutenzione della macchina.

Apprendimento



Non sempre lo stesso messaggio è ugualmente significativo di un errore in arrivo: la sua significatività rispetto allo stato di system failure è influenzato dalla macchina su cui si realizza

Astrazione e replicabilità



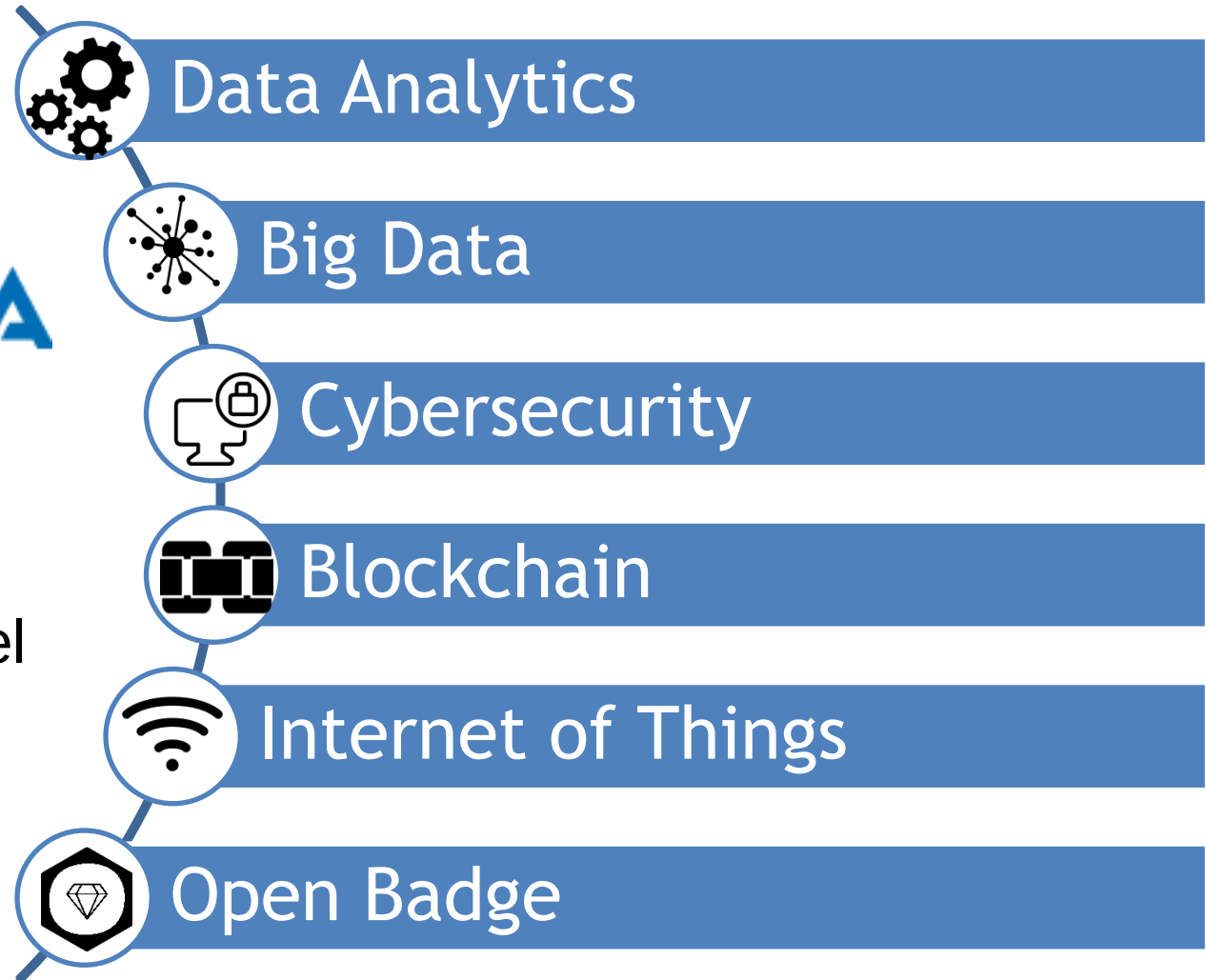
Il fatto che il ragionamento sia interamente basato sui dati di log, senza nessun parametro riguardante il funzionamento delle macchine o la struttura dell'impianto, ci permette di ottenere **un algoritmo potenzialmente riutilizzabile** per qualunque set di log, con **configurazioni minime** riguardanti ad esempio la lingua in cui i log sono scritti.

L'algoritmo ottenuto può essere eseguito in tempo reale per prevedere e anticipare gli eventi di system failure in arrivo.

Conclusioni

CINECA

Per il
trasferimento
tecnologico
e la crescita del
sistema
produttivo



Grazie

David Vannozzi
Direttore Generale

www.cineca.it