

NETWORK **DIGITAL 360**

I NOSTRI SERVIZI

MENU

Agenda  Digitale 

Cittadinanza digitale ▾ Sicurezza Informatica ▾ Sanità digitale

ACCEDI

360On
Scopri il nuovo canale YouTube

COVID-19

Colori delle Regioni, non fidiamoci dell'algoritmo: ecco perché

Home > Cultura Digitale



Ha parametri eccessivi l'algoritmo usato dal Governo per definire le zone di rischio covid, rosse, gialle o arancioni. E nessuno l'ha visto. Trasparenza e oggettività sono dubbie. Ecco perché

17 minuti fa

Federico Fuga

ingegnere elettronico, coordinatore della commissione ICT dell'Ordine degli Ingegneri della provincia di Verona



WEBINAR



Il webcast è disponibile

GUARDA

360On
Scopri il nuovo canale YouTube

Canali

 Cultura digitale

Articoli correlati



LA SOLUZIONE

È lecito avere qualche dubbio sulla **trasparenza e l'oggettività dell'algoritmo con cui il Governo sta decidendo i colori per le Regioni a rischio covid-19; rosso, arancione, giallo, con le note chiusure differenziate.**

Sappiamo che c'è dietro un meccanismo automatico di determinazione dei livelli di allerta e delle conseguenti misure di contenimento.

Il meccanismo è stato inserito all'interno del DPCM del 3 novembre, ed è stato definito fin dalla sua pubblicazione come un **algoritmo**, intendendo probabilmente mettere in risalto la sua assoluta obiettività.

Indice degli argomenti

Il significato del termine algoritmo

Eppure qualcosa non torna. Vale la pena fare una piccola introduzione per chiarire il significato del termine algoritmo. I lettori ci perdoneranno: sembra un excursus, ma andare al passato è utile per capire meglio qualcosa che tocca tutti noi molto (troppo) da vicino.

Nel IX secolo dopo Cristo, il matematico arabo *Muhammad ibn Mūsa* importò in Europa il concetto di procedimento di calcolo. Egli era detto *al-Khwarizmī*, poiché originario di *Khwarizm*, regione dell'Asia centrale. Da ciò nel medioevo si iniziò a utilizzare il termine *algorismus* per identificare ogni procedimento basato sulle cifre arabe che stavano soppiantando la numerazione romana.

La formalizzazione del concetto di algoritmo iniziò molti secoli più tardi, con l'*Entscheidungsproblem* o "*problema della decisione*", proposto da D. Hilbert nel 1928 e poi con il lavoro di A. Turing e A. Church, che nel 1936 indipendentemente definirono il concetto di funzione calcolabile, dando tra l'altro una risposta definitiva al problema di Hilbert, contribuendo alla crisi della matematica della metà del '900.

Un algoritmo è dunque un procedimento per risolvere un problema, costituito da un numero finito di passi, o istruzioni, che gode delle

Tutti i problemi del framework Apple e Google contro il covid

28 Apr 2020

di **Riccardo Berti e Alberto Pelliccione**Condividi 

IL PUNTO

Crisis Management: una "pericolosa opportunità" di innovazione per le imprese

21 Mag 2020

di **Silvia Celani**Condividi 

#DATECIVOCE

Il Covid-19 e la rappresentanza femminile in Italia: tutto quello che deve cambiare

23 Apr 2020

di **Darya Majidi**Condividi 

WHITE PAPER



Scaricalo gratis!

[DOWNLOAD](#)

caratteristiche di essere **finito**, ossia giunga ad una conclusione, **deterministico, non ambiguo** e generale, ossia **che risolva tutti i problemi della stessa classe a cui si applica**.

I due punti fondamentali su cui dobbiamo soffermare la nostra attenzione sono proprio il secondo e il terzo: se il procedimento è veramente un algoritmo esso deve essere deterministico, cioè che dato una certa combinazione di numeri si ottenga sempre lo stesso risultato, e non ambiguo, non deve esserci incertezza sull'interpretazione dei vari passaggi.

Come funziona l'algoritmo del governo che determina i colori delle Regioni

La determinazione del "colore" da attribuire alle regioni e le relative misure da adottare sono, come detto, definite nel DPCM del 3 novembre^[1] tuttavia i criteri per l'attribuzione sono specificati in altri due documenti, il primo, **nel DPCM del 30 aprile** (Circolare della Direzione Generale Prevenzione Sanitaria), il secondo è intitolato "Prevenzione e risposta a COVID-19: evoluzione della strategia e pianificazione nella fase di transizione per il periodo autunno invernale", condiviso dalla Conferenza delle Regioni e Province autonome **l'8 ottobre 2020** e inserito nel DPCM del 3/11 come **allegato 25**.

Il primo documento definisce due concetti basilari: quello di "**indicatore di monitoraggio**", cioè un dato che una volta valutato può descrivere in termini qualitativi la situazione da una certa prospettiva; e il concetto di "**livello di rischio**", che definisce il livello di allerta generale sulla base di una scala a 5 livelli.

Gli indicatori di monitoraggio

L'attività di monitoraggio della zona in oggetto definisce **21 indicatori suddivisi in 3 grandi categorie** che qui sintetizziamo:

- **Capacità di monitoraggio**, corrispondente a sei indicatori tra cui, per esempio, il numero di casi sintomatici notificati per mese, il numero di casi sintomatici con storia di ricovero ospedaliero, il numero di casi con trasferimento e ricovero in terapia intensiva,
- **Capacità di accertamento diagnostico**, indagine e gestione dei contatti, comprendente sei indicatori come il numero di test PCR

effettuato, il tempo medio trascorso tra l'insorgenza dei sintomi e la diagnosi, il numero di addetti al contact tracing, ed altri, e infine

- **Stabilità di trasmissione e tenuta dei servizi sanitari**, che include nove indicatori tra cui il numero di casi riportati dalla protezione civile, il valore stimato dell'Rt, numero di casi di infezione, eccetera.

Questi indicatori, numerici, sono confrontati con dei valori soglia riportati in una tabella allegata al DPCM, in modo da poter attribuire a ciascuno di essi un livello di attenzione, corrispondenti a "normale" oppure "allerta".

Il livello di rischio per l'algorithm covid

Il livello di rischio è come detto **una grandezza sintetica che determina 5 livelli di rischio**, molto basso, basso, moderato, alto e molto alto.

Tale indice viene determinato sulla base di una ulteriore tabella a due entrate, basata su altri due indici, anch'essi sintetizzati su **una scala di 4 livelli** (da molto basso a alto).

Un indice è la "probabilità di infezione/trasmissione" e sintetizza in sostanza quanto il virus si sta diffondendo tra la popolazione, mentre l'altro indice è "l'impatto", ovvero gli effetti della diffusione del virus in relazione sia alle singole persone (ossia la mortalità) sia alla tenuta del Sistema Sanitario nel suo complesso.

Essi vengono valutati con due algoritmi, che il medesimo allegato illustra.

Si tratta in sostanza di una sequenza di tre domande attraverso le quali il valutatore, rispondendo con "sì" o "no", può determinare il valore dell'indice corrispondente sulla scala.

Ad esempio, per quanto riguarda il primo parametro, "probabilità" le domande sono le seguenti:

- Sono stati segnalati dei **nuovi casi negli ultimi cinque giorni** nella regione/PPAA? Se no la probabilità è "molto bassa" e l'algorithm termina, altrimenti si prosegue.
- Vi è evidenza di **aumento di trasmissione** (presenza di almeno due elementi tra trend di casi in aumento, $R_t > 1$ e/o aumento del numero o dimensione dei focolai)? Se no la probabilità è "bassa",

altrimenti prosegui.

- Vi è evidenza di **trasmissione diffusa** nella regione/PPAA non gestibile in modo efficace con misure locali (“zone rosse”)? Se no la probabilità è “moderata”, altrimenti è “alta”.

L’algoritmo per l’impatto è del tutto analogo.

Gli scenari di rischio

Il secondo documento^[2] intitolato “Prevenzione e risposta al Covid-19”, è un documento piuttosto corposo costituito da diversi capitoli riguardanti diversi aspetti relativi alla *preparedness* nella gestione dell’emergenza sanitaria; nel quinto capitolo, intitolato “Approccio alla ri-modulazione delle misure di contenimento/mitigazione a livello regionale/PA in ambito di ipotetici scenari di trasmissione del virus SARS-CoV-2 sul territorio nazionale nel periodo autunno-invernale” vengono definiti **quattro scenari di rischio** a cui corrispondono una serie di misure di controllo e mitigazione che, come spiegato in premessa, sono da intendersi come orientative più che prescrittive.

Il parametro RT

Tali scenari sono basati tutti sul valore del **parametro Rt**.

Ricordiamo che il parametro R_t è un valore che definisce il numero medio di infezioni che una popolazione di individui infetti produce; è un parametro statistico che interpreta l’evoluzione dell’epidemia nel tempo e tiene conto degli effetti degli interventi di sanità pubblica e della situazione contingente locale.

È un parametro utilissimo per stimare l’evoluzione dell’epidemia, ma la cui misura è soggetta inevitabilmente ad effetti statistici; ad esempio, se la qualità della misura peggiora, per esempio perché il sistema di diagnostica e *tracking* è saturo, anche la qualità della stima dell’ R_t peggiora. Al momento, con un tasso di test PCR positivi che si aggira attorno al 25%, non si può considerare una misura sufficientemente buona, e in effetti il *range* per cui l’ R_t è da considerarsi affidabile con probabilità del 95% è tra 1.08 e 1.81. Una settimana fa era tra 1,45 e 1.83^[3].

I 4 scenari, numerati da 1 a 4 per gravità crescente, sono così

definiti: per valori di R_t inferiori a 1 si delinea lo scenario 1 (trasmissione localizzata sostanzialmente invariata rispetto a luglio/agosto), tra 1.0 e 1.25 lo scenario 2 (trasmissibilità sostenuta ma gestibile nel breve periodo), tra 1.25 e 1.5 identifica lo scenario 3 (trasmissibilità sostenuta con rischi di tenuta per il S.S. nel medio periodo) e per valori maggiori a 1.5 si ha il quarto scenario (trasmissibilità sostenuta con criticità nel breve periodo).

Le zone

Il DPCM di novembre all'art.2 definisce dunque quella che informalmente è detta "zona arancione" una regione di rischio alto in scenario 3, mentre l'art. 3 definisce la cosiddetta "zona rossa" una regione a rischio alto in scenario 4.

Sulla base di questi livelli di allerta, informalmente chiamati "colori", si applicano dunque i provvedimenti descritti nei successivi commi.

I problemi dell'algoritmo covid-19 scelto dal Governo

Fin qui tutto sembrerebbe abbastanza deterministico.

Tuttavia, se il passaggio dal livello di rischio e dallo scenario alla zona di rischio sembra privo di ambiguità, **pare invece assai lacunoso il passaggio dagli indicatori di monitoraggio ai parametri di probabilità e impatto.**

Infatti, alla fine della valutazione del livello di allerta di ciascuno dei 21 parametri, otteniamo fino a 21 valori che possono essere "normali" oppure "in allerta". Tuttavia, le "domande" dei due algoritmi sono abbastanza generiche da poter dare spazio ad ambiguità e a discrezionalità. Cosa significa infatti "aumento della dimensione dei focolai" o "evidenza di trasmissione diffusa"?

Non sono certo valutazioni quantitative, ma sono valutazioni del tutto qualitative, ad esempio **l'allocuzione "trasmissione diffusa"** senza indicare un valore di R_t o una qualunque altra soglia quantitativa, calcolabile meccanicamente, o un criterio oggettivo di stima, **lascia molto spazio all'arbitrarietà.**

Questo specifico passaggio soffre dunque di una **grave mancanza di chiarezza** e tale ambiguità è **arbitrarietà** inficia complessivamente

sul requisito di non ambiguità della descrizione dell'algoritmo nel complesso, rendendolo dunque **non deterministico** essendo soggetto al giudizio personale del valutatore.

Senza chiarire dei criteri qualitativi in tale passaggio è, dunque, impossibile codificare l'intera valutazione in un procedimento non ambiguo e deterministico, necessario alla valutazione della bontà e adeguatezza dei risultati.

Troppi indicatori nell'algoritmo

Al di là dei dubbi esposti, un discorso a parte va fatto sull'opportunità, infine, di **avere un così alto numero di indicatori** per determinare gli indici di probabilità e di impatto.

Ogni valutazione matematica di un sistema fisico parte da un modello, per il semplice fatto che i "numeretti" ovvero i dati non "esistono" nella realtà, ma sono una rappresentazione di grandezze che nei processi fisici o, in questo caso, epidemiologici, si combinano, cambiano, si trasformano.

E ogni modellista sa che i parametri del modello, in particolare quelli che servono a valutare lo stato del modello, devono essere stimati ed aggiustati.

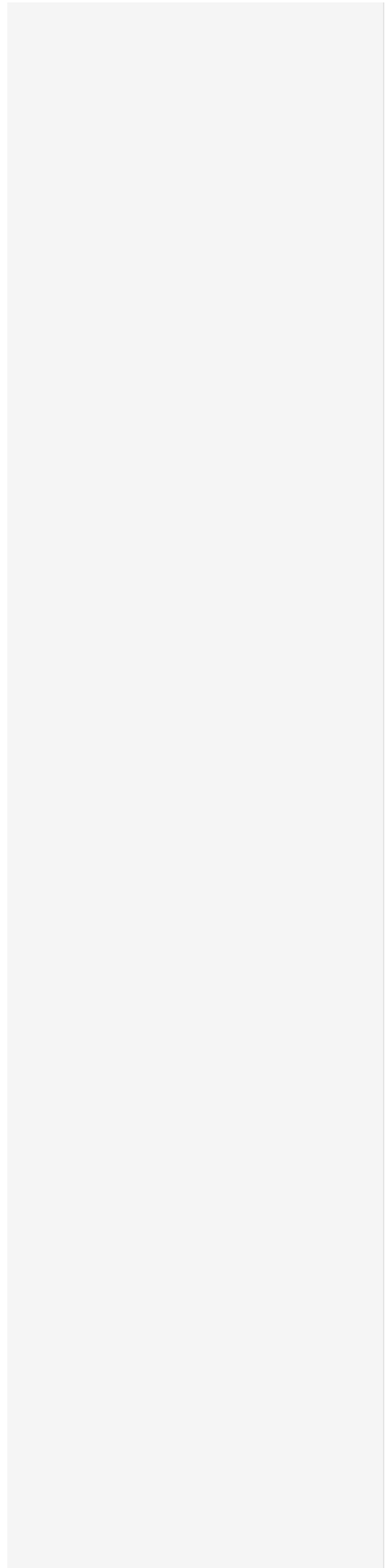
Nei modelli descrittivi, per esempio, essi si aggiustano finché **il risultato del calcolo coincide con quello osservato nella realtà**, mentre in quelli decisionali si devono aggiustare affinché le decisioni siano ragionevoli e coerenti con il contesto.

Si pensi per esempio alla difficoltà che abbiamo sotto la doccia per regolare la temperatura dell'acqua, quando un movimento troppo rapido ci porta rapidamente dal freddo polare alla scottatura.

Dobbiamo "tarare" la velocità con la quale muoviamo il miscelatore, tenendo conto, tra le altre cose, del ritardo nella risposta del sistema, che è sempre fonte di grandissima instabilità.

Dunque, abbiamo un modello con 21 parametri. Sono tanti? Sono pochi?

Talvolta abbiamo l'illusione che è meglio avere tanti dati per avere maggiore informazione. In realtà non è affatto così, è una fallacia logica assai diffusa.



Luca Dellanna^[4], nel suo libro “The Control Heuristic”, usa una parabola molto efficace per spiegare questo concetto: se vogliamo predire che tipo di caffè tra espresso, frappuccino (una bevanda molto in voga quando andai negli USA nel 2006) e caffè turco ordinerà il prossimo cliente della nostra caffetteria, è molto più utile avere pochi parametri come la nazionalità, o i tratti somatici, che tanti altri parametri come, per esempio, l’aver o meno una borsa o il presentarsi ad una certa ora del giorno, anche se essi sembrano mostrare una certa e significativa correlazione con il fenomeno osservato. **Più parametri si usano, meno essi sono significativi** e il loro contenuto informativo (o meglio, la loro significatività) decresce velocemente, mescolandosi sempre più con il **rumore**. **È il problema dell’overfitting**.

Non mi stupisce dunque che il presidente della regione Friuli Venezia Giulia abbia proposto la **riduzione dei 21 parametri di monitoraggio ad un esiguo e più gestibile insieme di 5 o 6**. La trovo personalmente una decisione più che ragionevole.

Conclusioni

Come abbiamo visto, l’ambiguità di alcuni passaggi del cosiddetto “algoritmo” inficiano complessivamente sulla **trasparenza del procedimento di assegnazione delle zone di rischio**, e dunque influiscono pesantemente sull’oggettività delle scelte. Del resto, la sostanziale incoerenza con cui alcune regioni sono state assegnate alle zone di più stretta limitazione mentre altre, in situazione apparentemente simile, sono rimaste più libere, fa pensare che ci fosse già in origine un **qualcosa di rotto nel meccanismo**.

Di certo, un più **trasparente** e semplice meccanismo di calcolo, magari basato su un numero ridotto di parametri essenziali, avrebbe permesso **una minore discrezionalità** e una più efficace e tempestiva azione che sarebbe forse stata accettata di buon grado sia dai presidenti di regione che dai cittadini stessi.


Del resto, **in condizioni di incertezza i meccanismi inutilmente complicati introducono soltanto maggiore complessità senza essere più efficaci**, e ritardando l’azione aggravano una situazione particolarmente critica quando obiettivamente nessuno ha il lusso di sprecare giorni in inazione e in discussione.

1. <http://www.governo.it/it/articolo/coronavirus-il-presidente-conte-firma-il-dpcm-del-3-novembre-2020/15617> ↑
2. http://www.salute.gov.it/imgs/C_17_pagineAree_5373_16_file.pdf
↑
3. https://www.huffingtonpost.it/entry/covid-indice-rt_it_5fb05a2ec5b6c5f3d2f7f156 ↑
4. <https://www.luca-dellanna.com/> ↑

■
@RIPRODUZIONE RISERVATA

Articolo 1 di 4

Agenda  Digitale 

Seguici 





About

Autori

Tags

Rss Feed

Privacy

Cookie

Cookie Center

NETWORK **DIGITAL** 360

NetworkDigital360 è il più grande network in Italia di testate e portali B2B dedicati ai temi della Trasformazione Digitale e dell'Innovazione Imprenditoriale. Ha la missione di diffondere la cultura digitale e imprenditoriale nelle imprese e pubbliche amministrazioni italiane.

TUTTE LE TESTATE

Applicazioni e Tecnologie

AI4BUSINESS
BIGDATA4INNOVATION
BIG DATA & ANALYTICS ZEROUNO
BLOCKCHAIN4INNOVATION
CLOUD COMPUTING ZEROUNO
CYBERSECURITY CORCOM
CYBERSECURITY360
DOCUMENTI AGENDADIGITALE.EU
ECOMMERCE AGENDADIGITALE.EU
FATTURAZIONE AGENDADIGITALE.EU
INDUSTRIA 4.0 CORCOM
INDUSTRY 4.0 AGENDADIGITALE.EU
INFRASTRUTTURE AGENDADIGITALE.EU
INDUSTRY4BUSINESS
INTERNET4THINGS
PAGAMENTIDIGITALI
RISKMANAGEMENT360
SEARCHDATACENTER ZEROUNO
SEARCHSECURITY ZEROUNO
SICUREZZA AGENDADIGITALE.EU
SMART CITY AGENDADIGITALE.EU
SMART MOBILITY ECONOMYUP

PA

CITTADINANZA AGENDADIGITALE.EU
FORUMPA
PA DIGITALE CORCOM
SANITÀ AGENDADIGITALE.EU

Digital Transformation

AGENDADIGITALE.EU
CORCOM
DIGITAL4EXECUTIVE
DIGITAL4PMI
TECHCOMPANY360
ZEROUNO

Funzioni di Business

DIGITAL4FINANCE
DIGITAL4HR
DIGITAL4LEGAL
DIGITAL4MARKETING
DIGITAL4PROCUREMENT
DIGITAL4SUPPLYCHAIN
PROCUREMENT AGENDADIGITALE.EU

Startup e Open Innovation

ECONOMYUP
STARTUP AGENDADIGITALE.EU
STARTUPBUSINESS

Industry

AGRIFOOD.TECH
AUTOMOTIVEUP
BANKINGUP
ENERGYUP
INDUSTRY4BUSINESS
INSURANCEUP
MEDIA CORCOM
RETAILUP
SANITÀ AGENDADIGITALE.EU
SCUOLA AGENDADIGITALE.EU
TELCO CORCOM
TURISMO AGENDADIGITALE.EU

Studenti

UNIVERSITY2BUSINESS