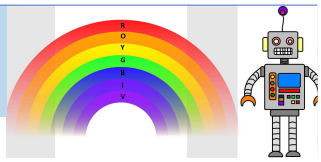
The united RAINBOW colors of bots

Detecting and filtering out machine-generated statistics in the
World Bank's Open Knowledge Repository

Introducing: The rainbow process of detecting bot statistics



From the **EASIEST** to detect to the **MOST DIFFICULT** to detect:

RED: Brut force DDOS (and similar) attacks on a server

ORANGE: Crawler HTTP agents: Usually declare themselves

YELLOW: Online lists of known bot IP addresses

Note: List update is too infrequent to serve as the **only** filter

GREEN: Newly found bot IP addresses

Note: Run suspicious IP addresses through online meta blacklists like www.anti-abuse.org, www.abuseipdb.com.

BLUE: Non-blacklisted IP addresses behaving suspiciously
e.g.: Zero to large number of views/downloads;
same download patters every day, etc.

INDIGO: Bots disguised as VPNs: Multiple HTTP agents going through one IP address simulate VPN legitimate activity

VIOLET: Bots set in arrays of IP addresses, avoiding detection

The "**Heisen-BOT principle**": Machine-generated statistics detection Is not an exact science and some bot activity is bound to filter in undetected

Future Trends:

- Botnet detection tool market to grow 37.6% by 2026
- Bots will become better in disguising their identity and mimicking human behavior
- Developments in machine learning improve automated bot detection

Tal Ayalon, Open Knowledge Repository project manager,
World Bank Group: tayalon@worldbankgroup.org