# ===============================================

# ROBOTS.TXT ULTIEME SECURE-PERMISSIVE VERSIE

# ===============================================

# Versie: 6.0 (Geoptimaliseerd voor AI-toegang, Serverstabiliteit & Beheerbaarheid)

# Laatst bijgewerkt: September 2025

# Compatibiliteit: Alle belangrijke zoekmachines, AI-systemen en crawlers

#

# 🎯 WAT DIT BESTAND DOET:

# ✅ Maximaal toestaan van AI-systemen, crawlers en onderzoeksbots met gelaagde toegang.

# ✅ Implementeert robuuste veiligheidsmaatregelen en strikte snelheidslimieten voor onbekende bots.

# ✅ Beschermt gevoelige directories en bestandstypen tegen crawling voor ALLE bots.

# ✅ Biedt gedifferentieerde toegangsniveaus op basis van betrouwbaarheid en prioriteit van de bot.

# ✅ Toekomstbestendig voor opkomende AI-technologieën tot ver na 2030.

# ✅ Omvat academische instellingen, onderzoeksorganisaties en grote technologiebedrijven.

# ✅ Gebruikt geoptimaliseerde wildcard-patronen voor brede dekking en beheerbaarheid.

# ✅ Balanceert maximale toegankelijkheid van publieke content met serverbescherming tegen overbelasting.

#

# 🛡️ VEILIGHEIDSFUNCTIES EN PREVENTIE VAN SERVER OVERLOAD:

# 1. Globale Disallow Regels (User-agent: \* bovenaan): Deze regels blokkeren onvoorwaardelijk toegang tot

# alle gevoelige mappen en bestandsextensies voor \*elke\* bot. Dit is uw primaire bescherming tegen

# ongeautoriseerde toegang tot kritieke bronnen. Een latere 'Allow'-regel voor een specifiek pad

# onder een User-agent block kan deze overschrijven, maar dit zou \*nooit\* moeten gebeuren voor

# de hier geblokkeerde, gevoelige bronnen.

# 2. Expliciete 'Allow' voor veilige, publieke bronnen: Dit heft specifieke globale 'Disallow'-regels

# op voor statische bestanden of publieke API's die essentieel zijn voor de functionaliteit

# of weergave van de website, zonder de server onnodig te belasten.

# 3. Gelaagde Crawl-Delays & Request-Rates:

# - \*\*Tier 1 (Cruciale Bots):\*\* Zeer genereuze crawl-snelheden voor essentieel zoekverkeer en top-tier AI-trainingsbots.

# - \*\*Tier 2 (Algemene AI & Onderzoek):\*\* Gematigde crawl-snelheden voor een brede reeks AI-systemen en academische crawlers.

# - \*\*Tier 3 (Onbekende/Niet-Geclassificeerde Bots):\*\* Extreem restrictieve snelheden en beperkte tijden om serveroverload

# door onbekende of potentieel ongewenste bots te voorkomen.

# 4. Visit-Time Restrictie: Beperkt crawling van onbekende bots tot daluren.

# 5. Clean-param Directive: Helpt Googlebot om overbodige URL-parameters te negeren, wat crawl-budget bespaart.

# 6. Host Directive: Biedt een hint voor het voorkeursdomein aan oudere/specifieke crawlers.

#

# BELANGRIJK: robots.txt is een adviesbestand. Kwaadwillende of slecht geprogrammeerde bots

# kunnen het negeren. Aanvullende server-side beveiliging (bijv. WAF, IP-blokkering,

# rate limiting op de webserver) is sterk aanbevolen voor optimale bescherming.

#

# ⚠️ BELANGRIJKE SETUP INSTRUCTIES:

# 1. Vervang "https://yoursite.com" en "https://www.yoursite.com" door uw daadwerkelijke domeinnamen voor Sitemaps en Host-directive.

# 2. Monitor de serverbronnen nauwlettend na implementatie.

# 3. Pas crawl-delays/request-rates aan indien serverprestatieproblemen optreden.

# 4. Test met Google Search Console en Bing Webmaster Tools.

# 5. Werk dit bestand jaarlijks (of vaker) bij met nieuwe AI-systemen of beveiligingsregels.

#

# ===============================================

# --- GLOBALE BLOKKERINGEN VOOR ALLE BOTS ---

# Deze regels worden toegepast op ELKE User-agent, tenzij een latere,

# specifieke 'Allow'-regel voor een \*publieke en veilige\* resource deze overschrijft.

User-agent: \*

# Bestandsextensies blokkeren die gevaarlijk of resource-intensief kunnen zijn

Disallow: \*.sql$

Disallow: \*.bak$

Disallow: \*.backup$

Disallow: \*.dump$

Disallow: \*.log$

Disallow: \*.txt$

Disallow: \*.conf$

Disallow: \*.config$

Disallow: \*.ini$

Disallow: \*.cfg$

Disallow: \*.properties$

Disallow: \*.env$

Disallow: \*.key$

Disallow: \*.pem$

Disallow: \*.crt$

Disallow: \*.cer$

Disallow: \*.p12$

Disallow: \*.pfx$

Disallow: \*.jks$

Disallow: \*.keystore$

Disallow: \*.truststore$

Disallow: \*.password$

Disallow: \*.passwd$

Disallow: \*.pwd$

Disallow: \*.secret$

Disallow: \*.token$

Disallow: \*.session$

Disallow: \*.cookie$

Disallow: \*.cache$

Disallow: \*.temp$

Disallow: \*.tmp$

Disallow: \*.swp$

Disallow: \*.swo$

Disallow: \*.bkp$

Disallow: \*.old$

Disallow: \*.orig$

Disallow: \*.save$

Disallow: \*.lock$

Disallow: \*.pid$

Disallow: \*.core$

Disallow: \*.crash$

Disallow: \*.trace$

Disallow: \*.debug$

Disallow: \*.error$

Disallow: \*.exception$

Disallow: \*.stack$

Disallow: \*.heap$

Disallow: \*.memory$

Disallow: \*.profile$

Disallow: \*.perf$

Disallow: \*.benchmark$

Disallow: \*.test$

Disallow: \*.spec$

Disallow: \*.mock$

Disallow: \*.stub$

Disallow: \*.fake$

Disallow: \*.dummy$

Disallow: \*.sample$

Disallow: \*.example$

Disallow: \*.demo$

Disallow: \*.trial$

Disallow: \*.beta$

Disallow: \*.alpha$

Disallow: \*.rc$

Disallow: \*.dev$

Disallow: \*.staging$

Disallow: \*.prod$

Disallow: \*.production$

Disallow: \*.live$

Disallow: \*.release$

Disallow: \*.version$

Disallow: \*.build$

Disallow: \*.compile$

Disallow: \*.deploy$

Disallow: \*.install$

Disallow: \*.setup$

Disallow: \*.configure$

Disallow: \*.migration$

Disallow: \*.upgrade$

Disallow: \*.patch$

Disallow: \*.hotfix$

Disallow: \*.bugfix$

Disallow: \*.fix$

Disallow: \*.repair$

Disallow: \*.restore$

Disallow: \*.recovery$

Disallow: \*.rollback$

Disallow: \*.revert$

Disallow: \*.undo$

Disallow: \*.snapshot$

Disallow: \*.checkpoint$

Disallow: \*.savepoint$

Disallow: \*.archive$

Disallow: \*.zip$

Disallow: \*.tar$

Disallow: \*.gz$

Disallow: \*.bz2$

Disallow: \*.xz$

Disallow: \*.7z$

Disallow: \*.rar$

Disallow: \*.ace$

Disallow: \*.arj$

Disallow: \*.cab$

Disallow: \*.lzh$

Disallow: \*.lha$

Disallow: \*.zoo$

Disallow: \*.arc$

Disallow: \*.pak$

Disallow: \*.pk3$

Disallow: \*.pk4$

Disallow: \*.war$

Disallow: \*.ear$

Disallow: \*.jar$

Disallow: \*.sar$

Disallow: \*.deb$

Disallow: \*.rpm$

Disallow: \*.dmg$

Disallow: \*.iso$

Disallow: \*.img$

Disallow: \*.bin$

Disallow: \*.exe$

Disallow: \*.msi$

Disallow: \*.app$

Disallow: \*.apk$

Disallow: \*.ipa$

Disallow: \*.pkg$

Disallow: \*.appx$

Disallow: \*.msix$

Disallow: \*.snap$

Disallow: \*.flatpak$

Disallow: \*.appimage$

Disallow: \*.bundle$

Disallow: \*.vbs$

Disallow: \*.php$ # Blokkeer PHP source code exposure

Disallow: \*.asp$

Disallow: \*.aspx$

Disallow: \*.jsp$

Disallow: \*.cgi$

Disallow: \*.pl$

Disallow: \*.py$

Disallow: \*.rb$

Disallow: \*.sh$ # Script files

Disallow: \*.bash$

Disallow: \*.zsh$

Disallow: \*.csh$

Disallow: \*.bat$

Disallow: \*.cmd$

Disallow: \*.wsf$

Disallow: \*.ps1$

Disallow: \*.vb$

# Gevoelige en resource-intensieve directories en API-endpoints blokkeren

Disallow: /admin/

Disallow: /wp-admin/

Disallow: /administrator/

Disallow: /cpanel/

Disallow: /plesk/

Disallow: /webmail/

Disallow: /phpmyadmin/

Disallow: /mysql/

Disallow: /database/

Disallow: /db/

Disallow: /sql/

Disallow: /backup/

Disallow: /backups/

Disallow: /bak/

Disallow: /dump/

Disallow: /dumps/

Disallow: /temp/

Disallow: /tmp/

Disallow: /temporary/

Disallow: /cache/

Disallow: /logs/

Disallow: /log/

Disallow: /private/

Disallow: /confidential/

Disallow: /secret/

Disallow: /hidden/

Disallow: /internal/

Disallow: /system/

Disallow: /config/

Disallow: /configuration/

Disallow: /settings/

Disallow: /include/

Disallow: /includes/

Disallow: /lib/

Disallow: /library/

Disallow: /libraries/

Disallow: /vendor/

Disallow: /node\_modules/

Disallow: /packages/

Disallow: /composer/

Disallow: /npm/

Disallow: /yarn/

Disallow: /bower/

Disallow: /assets/private/

Disallow: /uploads/private/

Disallow: /files/private/

Disallow: /documents/private/

Disallow: /media/private/

Disallow: /images/private/

Disallow: /photos/private/

Disallow: /videos/private/

Disallow: /audio/private/

Disallow: /download/private/

Disallow: /downloads/private/

Disallow: /user/

Disallow: /users/

Disallow: /member/

Disallow: /members/

Disallow: /account/

Disallow: /accounts/

Disallow: /profile/

Disallow: /profiles/

Disallow: /dashboard/

Disallow: /control/

Disallow: /panel/

Disallow: /manage/

Disallow: /management/

Disallow: /auth/

Disallow: /login/

Disallow: /signin/

Disallow: /signup/

Disallow: /register/

Disallow: /password/

Disallow: /reset/

Disallow: /forgot/

Disallow: /recovery/

Disallow: /activate/

Disallow: /verify/

Disallow: /confirm/

Disallow: /validate/

Disallow: /check/

Disallow: /test/

Disallow: /testing/

Disallow: /debug/

Disallow: /dev/

Disallow: /development/

Disallow: /staging/

Disallow: /beta/

Disallow: /alpha/

Disallow: /demo/

Disallow: /sandbox/

Disallow: /preview/

Disallow: /draft/

Disallow: /drafts/

Disallow: /work/

Disallow: /working/

Disallow: /wip/

Disallow: /progress/

Disallow: /old/

Disallow: /archive/

Disallow: /archives/

Disallow: /legacy/

Disallow: /deprecated/

Disallow: /obsolete/

Disallow: /unused/

Disallow: /disabled/

Disallow: /inactive/

Disallow: /suspended/

Disallow: /blocked/

Disallow: /banned/

Disallow: /blacklist/

Disallow: /whitelist/

Disallow: /maintenance/

Disallow: /repair/

Disallow: /service/

Disallow: /support/private/

Disallow: /help/private/

Disallow: /docs/private/

Disallow: /documentation/private/

Disallow: /manual/private/

Disallow: /guide/private/

Disallow: /tutorial/private/

Disallow: /readme/private/

Disallow: /changelog/private/

Disallow: /release/private/

Disallow: /version/private/

Disallow: /update/private/

Disallow: /upgrade/private/

Disallow: /patch/private/

Disallow: /hotfix/private/

Disallow: /bugfix/private/

Disallow: /fix/private/

Disallow: /repair/private/

Disallow: /restore/private/

Disallow: /recovery/private/

Disallow: /backup/private/

Disallow: /snapshot/private/

Disallow: /export/private/

Disallow: /import/private/

Disallow: /sync/private/

Disallow: /migration/private/

Disallow: /transfer/private/

Disallow: /move/private/

Disallow: /copy/private/

Disallow: /clone/private/

Disallow: /duplicate/private/

Disallow: /mirror/private/

Disallow: /replica/private/

Disallow: /shadow/private/

Disallow: /proxy/private/

Disallow: /gateway/private/

Disallow: /bridge/private/

Disallow: /tunnel/private/

Disallow: /vpn/private/

Disallow: /ssl/private/

Disallow: /tls/private/

Disallow: /certificate/private/

Disallow: /cert/private/

Disallow: /key/private/

Disallow: /keys/private/

Disallow: /token/private/

Disallow: /tokens/private/

Disallow: /session/private/

Disallow: /sessions/private/

Disallow: /cookie/private/

Disallow: /cookies/private/

Disallow: /storage/private/

Disallow: /data/private/

Disallow: /database/private/

Disallow: /db/private/

Disallow: /sql/private/

Disallow: /nosql/private/

Disallow: /mongodb/private/

Disallow: /mysql/private/

Disallow: /postgresql/private/

Disallow: /sqlite/private/

Disallow: /redis/private/

Disallow: /memcached/private/

Disallow: /elasticsearch/private/

Disallow: /solr/private/

Disallow: /lucene/private/

Disallow: /search/private/

Disallow: /index/private/

Disallow: /catalog/private/

Disallow: /directory/private/

Disallow: /listing/private/

Disallow: /browse/private/

Disallow: /explorer/private/

Disallow: /navigator/private/

Disallow: /viewer/private/

Disallow: /reader/private/

Disallow: /editor/private/

Disallow: /writer/private/

Disallow: /creator/private/

Disallow: /builder/private/

Disallow: /generator/private/

Disallow: /maker/private/

Disallow: /designer/private/

Disallow: /developer/private/

Disallow: /programmer/private/

Disallow: /coder/private/

Disallow: /hacker/private/

Disallow: /cracker/private/

Disallow: /breaker/private/

Disallow: /penetration/private/

Disallow: /vulnerability/private/

Disallow: /exploit/private/

Disallow: /attack/private/

Disallow: /intrusion/private/

Disallow: /malware/private/

Disallow: /virus/

Disallow: /trojan/

Disallow: /worm/

Disallow: /spyware/

Disallow: /adware/

Disallow: /ransomware/

Disallow: /rootkit/

Disallow: /botnet/

Disallow: /ddos/

Disallow: /dos/

Disallow: /spam/

Disallow: /phishing/

Disallow: /scam/

Disallow: /fraud/

Disallow: /abuse/

Disallow: /harassment/

Disallow: /bullying/private/

Disallow: /threat/private/

Disallow: /intimidation/private/

Disallow: /blackmail/private/

Disallow: /extortion/private/

Disallow: /bribery/private/

Disallow: /corruption/private/

Disallow: /illegal/private/

Disallow: /criminal/private/

Disallow: /offensive/private/

Disallow: /inappropriate/private/

Disallow: /nsfw/private/

Disallow: /adult/private/

Disallow: /mature/private/

Disallow: /explicit/private/

Disallow: /graphic/private/

Disallow: /violent/private/

Disallow: /disturbing/private/

Disallow: /sensitive/private/

Disallow: /controversial/private/

Disallow: /political/private/

Disallow: /religious/private/

Disallow: /racial/private/

Disallow: /ethnic/private/

Disallow: /gender/private/

Disallow: /sexual/private/

Disallow: /discriminatory/private/

Disallow: /prejudicial/private/

Disallow: /biased/private/

Disallow: /unfair/private/

Disallow: /unjust/private/

Disallow: /unethical/private/

Disallow: /immoral/private/

Disallow: /wrong/private/

Disallow: /bad/private/

Disallow: /evil/private/

Disallow: /dark/private/

Disallow: /shadow/private/

Disallow: /hidden/private/

Disallow: /secret/private/

Disallow: /classified/private/

Disallow: /restricted/private/

Disallow: /limited/private/

Disallow: /exclusive/private/

Disallow: /premium/private/

Disallow: /vip/private/

Disallow: /member/private/

Disallow: /subscriber/private/

Disallow: /customer/private/

Disallow: /client/private/

Disallow: /partner/private/

Disallow: /affiliate/private/

Disallow: /reseller/private/

Disallow: /distributor/private/

Disallow: /vendor/private/

Disallow: /supplier/private/

Disallow: /provider/private/

Disallow: /operator/private/

Disallow: /manager/private/

Disallow: /administrator/private/

Disallow: /owner/private/

Disallow: /creator/private/

Disallow: /author/private/

Disallow: /editor/private/

Disallow: /moderator/private/

Disallow: /supervisor/private/

Disallow: /controller/private/

Disallow: /monitor/private/

Disallow: /tracker/private/

Disallow: /spy/private/

Disallow: /surveillance/private/

Disallow: /monitoring/private/

Disallow: /tracking/private/

Disallow: /logging/private/

Disallow: /recording/private/

Disallow: /capturing/private/

Disallow: /collecting/private/

Disallow: /gathering/private/

Disallow: /harvesting/private/

Disallow: /mining/private/

Disallow: /scraping/private/

Disallow: /crawling/private/

Disallow: /indexing/private/

Disallow: /cataloging/private/

Disallow: /archiving/private/

Disallow: /storing/private/

Disallow: /saving/private/

Disallow: /caching/private/

Disallow: /buffering/private/

Disallow: /queuing/private/

Disallow: /pooling/private/

Disallow: /clustering/private/

Disallow: /grouping/private/

Disallow: /batching/private/

Disallow: /bundling/private/

Disallow: /packaging/private/

Disallow: /wrapping/private/

Disallow: /compressing/private/

Disallow: /encrypting/private/

Disallow: /encoding/private/

Disallow: /decoding/private/

Disallow: /decrypting/private/

Disallow: /decompressing/private/

Disallow: /unpacking/private/

Disallow: /unwrapping/private/

Disallow: /extracting/private/

Disallow: /parsing/private/

Disallow: /processing/private/

Disallow: /analyzing/private/

Disallow: /examining/private/

Disallow: /inspecting/private/

Disallow: /reviewing/private/

Disallow: /auditing/private/

Disallow: /checking/private/

Disallow: /testing/private/

Disallow: /validating/private/

Disallow: /verifying/private/

Disallow: /confirming/private/

Disallow: /authenticating/private/

Disallow: /authorizing/private/

Disallow: /approving/private/

Disallow: /denying/private/

Disallow: /rejecting/private/

Disallow: /blocking/private/

Disallow: /banning/private/

Disallow: /suspending/private/

Disallow: /disabling/private/

Disallow: /deactivating/private/

Disallow: /terminating/private/

Disallow: /deleting/private/

Disallow: /removing/private/

Disallow: /erasing/private/

Disallow: /wiping/private/

Disallow: /clearing/private/

Disallow: /cleaning/private/

Disallow: /purging/private/

Disallow: /flushing/private/

Disallow: /resetting/private/

Disallow: /restoring/private/

Disallow: /recovering/private/

Disallow: /repairing/private/

Disallow: /fixing/private/

Disallow: /patching/private/

Disallow: /updating/private/

Disallow: /upgrading/private/

Disallow: /installing/private/

Disallow: /deploying/private/

Disallow: /configuring/private/

Disallow: /setting/private/

Disallow: /adjusting/private/

Disallow: /tuning/private/

Disallow: /optimizing/private/

Disallow: /enhancing/private/

Disallow: /improving/private/

Disallow: /refining/private/

Disallow: /perfecting/private/

Disallow: /finalizing/private/

Disallow: /completing/private/

Disallow: /finishing/private/

Disallow: /closing/private/

Disallow: /ending/private/

Disallow: /stopping/private/

Disallow: /halting/private/

Disallow: /pausing/private/

Disallow: /freezing/private/

Disallow: /locking/private/

Disallow: /securing/private/

Disallow: /protecting/private/

Disallow: /defending/private/

Disallow: /guarding/private/

Disallow: /shielding/private/

Disallow: /covering/private/

Disallow: /hiding/private/

Disallow: /concealing/private/

Disallow: /masking/private/

Disallow: /obscuring/private/

Disallow: /camouflaging/private/

Disallow: /disguising/private/

Disallow: /anonymizing/private/

Disallow: /pseudonymizing/private/

Disallow: /randomizing/private/

Disallow: /shuffling/private/

Disallow: /mixing/private/

Disallow: /scrambling/private/

Disallow: /garbling/private/

Disallow: /distorting/private/

Disallow: /corrupting/private/

Disallow: /damaging/private/

Disallow: /deleting/private/

Disallow: /destroying/private/

Disallow: /breaking/private/

Disallow: /crashing/private/

Disallow: /failing/private/

Disallow: /erroring/private/

Disallow: /bugging/private/

Disallow: /glitching/private/

Disallow: /malfunctioning/private/

# API en Monitoring Endpoints die NIET gecrawld mogen worden

Disallow: /health/

Disallow: /status/

Disallow: /ping/

Disallow: /heartbeat/

Disallow: /metrics/

Disallow: /stats/

Disallow: /analytics/private/

Disallow: /monitoring/

Disallow: /telemetry/

Disallow: /profiling/

Disallow: /tracing/

Disallow: /logging/

Disallow: /debugging/

Disallow: /performance/

Disallow: /benchmark/

Disallow: /load-test/

Disallow: /stress-test/

Disallow: /capacity-test/

Disallow: /volume-test/

Disallow: /endurance-test/

Disallow: /spike-test/

Disallow: /scalability-test/

Disallow: /availability-test/

Disallow: /reliability-test/

Disallow: /security-test/

Disallow: /penetration-test/

Disallow: /vulnerability-scan/

Disallow: /compliance-check/

Disallow: /audit-trail/

Disallow: /access-log/

Disallow: /error-log/

Disallow: /event-log/

Disallow: /transaction-log/

Disallow: /activity-log/

Disallow: /security-log/

Disallow: /alert-log/

Disallow: /notification-log/

Disallow: /api/private/

Disallow: /api/internal/

Disallow: /api/admin/

Disallow: /api/management/

Disallow: /api/control/

Disallow: /api/config/

Disallow: /api/settings/

Disallow: /api/auth/

Disallow: /api/login/

Disallow: /api/logout/

Disallow: /api/register/

Disallow: /api/password/

Disallow: /api/reset/

Disallow: /api/forgot/

Disallow: /api/recovery/

Disallow: /api/activate/

Disallow: /api/verify/

Disallow: /api/confirm/

Disallow: /api/validate/

Disallow: /api/token/

Disallow: /api/session/

Disallow: /api/user/

Disallow: /api/users/

Disallow: /api/profile/

Disallow: /api/account/

Disallow: /graphql/private/

Disallow: /graphql/internal/

Disallow: /graphql/admin/

Disallow: /graphql/management/

Disallow: /graphiql/

Disallow: /graphql-playground/

Disallow: /apollo-studio/

Disallow: /ws/private/

Disallow: /websocket/private/

Disallow: /socket/private/

Disallow: /socketio/private/

Disallow: /realtime/private/

Disallow: /live/private/

Disallow: /stream/private/

Disallow: /push/private/

Disallow: /sse/private/

Disallow: /events/private/

Disallow: /wp-json/ # Blokkeert WordPress REST API (indien gebruikt), vaak zwaar gecrawld

# --- EXPLICIET TOESTAAN VAN VEILIGE PUBLIEKE BRONNEN VOOR ALLE BOTS ---

# Deze 'Allow'-regels overschrijven de bredere 'Disallow'-regels hierboven voor specifieke,

# veilig te crawlen paden die anders geblokkeerd zouden worden. Dit zorgt ervoor dat

# essentiële content en functionaliteit beschikbaar blijft voor AI en zoekmachines.

# Voorbeeld: Als /api/v1/public/data een publieke, veilige API is

# Let op: Vervang 'yoursite.com' met uw echte publieke API paden.

Allow: /api/v1/public/data

# Voorbeeld: Statische assets die door frameworks worden gebruikt (Next.js, Nuxt.js)

# en niet gevoelig zijn.

Allow: /\_next/

Allow: /\_nuxt/

Allow: /static/

Allow: /assets/css/

Allow: /assets/js/

Allow: /assets/img/

Allow: /images/public/

Allow: /uploads/public/

Allow: /sitemaps/

# --- TIERED ACCESS VOOR BOTS ---

# Regels worden verwerkt van boven naar beneden. Een bot volgt de meest specifieke set regels die hij tegenkomt.

# 'Allow: /' hier betekent: 'Sta alles toe, BEHALVE wat in de globale 'Disallow: \*'-sectie hierboven is geblokkeerd,

# tenzij dit specifiek is toegestaan in het 'Expliciet Toestaan' blok.'

# Tier 1: Cruciale Zoekmachines & Primaire AI Crawlers

# Deze bots zijn essentieel voor zichtbaarheid en AI-training. Ze krijgen de meest genereuze tarieven.

User-agent: Googlebot

Allow: /

Crawl-delay: 1

Request-rate: 2/5s # 2 verzoeken per 5 seconden

Clean-param: sort&page / # Voorbeeld voor parameters die geen unieke content creëren op alle paden

Clean-param: utm\_source&utm\_medium&utm\_campaign&utm\_term&utm\_content / # Analytics parameters

# Als u specifieke delen van de site heeft met dynamische parameters, bijv. /products?filter=x&order=y

# Clean-param: filter&order /products

User-agent: Googlebot-Image

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: Googlebot-Video

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: Googlebot-News

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: Google-InspectionTool

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: GoogleOther

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: GoogleAssistant

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: Google-Extended

Allow: /

Crawl-delay: 2

Request-rate: 1/8s # 1 verzoek per 8 seconden

User-agent: Bingbot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: BingPreview

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: MicrosoftPreview

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: msnbot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: MSNBot-Media

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: Slurp # Yahoo! Search

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: YandexBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

# Dek alle Yandex sub-bots met een wildcard voor beheerbaarheid

User-agent: Yandex\*

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: Baiduspider

Allow: /

Crawl-delay: 3

Request-rate: 1/10s

User-agent: Applebot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: AppleNewsBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: GPTBot # OpenAI's hoofd trainingsbot

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: ChatGPT-User # OpenAI's ChatGPT bot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: OpenAI-SearchBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: OpenAI-\* # Dek overige OpenAI bots af

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: Claude\* # Anthropic Claude botfamilie (Claude-Web, ClaudeBot, Claude-1 t/m Claude-5, etc.)

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: anthropic-ai # Algemene Anthropic bot

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

# Tier 2: Algemene AI, Onderzoek, Academische, Sociale & Analyse Bots

# Deze bots zijn waardevol, maar met een iets conservatiever crawl-tarief om de server te beschermen.

# We consolideren grote groepen onder wildcards om het bestand beheersbaar te houden.

User-agent: DuckDuckBot

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: DuckDuckGo-Favicons-Bot

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: facebookexternalhit

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: Twitterbot

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: LinkedInBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: WhatsApp

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: TelegramBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: SkypeUriPreview

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: Slackbot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: DiscordBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: RedditBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: PinterestBot

Allow: /

Crawl-delay: 1 # Belangrijk voor visuele content ontdekking

Request-rate: 2/5s

User-agent: TikTokBot

Allow: /

Crawl-delay: 1 # Belangrijk voor video content ontdekking

Request-rate: 2/5s

User-agent: InstagramBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: SnapchatBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: YouTubeBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: VimeoBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: TwitchBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: SpotifyBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: SoundCloudBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: CCBot # Common Crawl bot

Allow: /

Crawl-delay: 3

Request-rate: 1/10s

User-agent: PerplexityBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: YouBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: Bytespider # ByteDance bot

Allow: /

Crawl-delay: 3

Request-rate: 1/10s

User-agent: Meta-ExternalAgent

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: Meta-ExternalFetcher

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: ImagesiftBot

Allow: /

Crawl-delay: 5

Request-rate: 1/15s

User-agent: Diffbot

Allow: /

Crawl-delay: 5

Request-rate: 1/15s

User-agent: AI2Bot # Allen Institute for AI

Allow: /

Crawl-delay: 5

Request-rate: 1/15s

User-agent: Amazonbot # Amazon's crawler, niet gerelateerd aan AWS AI direct

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: ia\_archiver # Internet Archive

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: archive.org\_bot # Internet Archive

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: Wayback # Internet Archive

Allow: /

Crawl-delay: 1

Request-rate: 2/5s

User-agent: SemrushBot # SEO tool

Allow: /

Crawl-delay: 5

Request-rate: 1/20s

Visit-time: 0000-0800,1800-2359 # Voorbeeld: Alleen buiten piekuren (UTC)

User-agent: AhrefsBot # SEO tool

Allow: /

Crawl-delay: 5

Request-rate: 1/20s

Visit-time: 0000-0800,1800-2359 # Voorbeeld: Alleen buiten piekuren (UTC)

User-agent: MJ12bot # SEO tool

Allow: /

Crawl-delay: 5

Request-rate: 1/20s

User-agent: DotBot # SEO tool

Allow: /

Crawl-delay: 5

Request-rate: 1/20s

User-agent: FlipboardProxy

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: NewsBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: PocketParser

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

User-agent: InstapaperBot

Allow: /

Crawl-delay: 2

Request-rate: 1/8s

# Consolideer belangrijke tech-bedrijf AI/ML bots met wildcards voor beheerbaarheid

User-agent: NVIDIA-\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s # 1 verzoek per 30 seconden

User-agent: Apple-\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: Meta\* # Dekt MetaAI, FacebookAI, etc.

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: IBM-Watson-\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: Microsoft-\* # Dekt Microsoft-AI, Microsoft-Copilot, etc.

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: Amazon-\* # Dekt Amazon-AI, AWS-AI, etc.

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

# Algemene en wildcard AI/ML/Onderzoeksbots

# Deze vangen een breed scala aan AI-systemen op die niet specifiek zijn genoemd.

User-agent: \*AI\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*ML\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Deep\*Learning\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Neural\*Network\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Machine\*Learning\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Artificial\*Intelligence\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Computer\*Vision\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Natural\*Language\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Speech\*Recognition\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Bot\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Crawler\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Spider\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Scraper\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Agent\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Research\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Academic\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Knowledge\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Data\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

User-agent: \*Analytics\*

Allow: /

Crawl-delay: 10

Request-rate: 1/30s

# Tier 3: Catch-all voor ONBEKENDE of niet-geclassificeerde bots

# Deze sectie is de laatste verdedigingslinie en is extreem restrictief.

# Elke bot die niet overeenkomt met een specifiekere User-agent regel hierboven,

# valt terug op deze regels. Dit voorkomt effectief serveroverload door willekeurige

# of potentieel misdragende bots, terwijl de site technisch 'toegankelijk' blijft.

User-agent: \*

Allow: / # Staat technische toegang toe tot de resterende content

Crawl-delay: 60 # Zeer hoge vertraging: 1 minuut tussen elk verzoek

Request-rate: 1/120s # Maximaal 1 verzoek elke 2 minuten

Visit-time: 0000-0600 # Beperk crawling tot daluren (tussen middernacht en 6 uur 's ochtends UTC)

# ===============================================

# SITEMAP DECLARATIES & NUTTIGE INFORMATIE

# ===============================================

# Gebruik een sitemap index bestand om de structuur te optimaliseren. Dit is efficiënter

# voor crawlers dan het parsen van vele losse sitemaps.

# Zorg ervoor dat sitemap\_index.xml daadwerkelijk bestaat en verwijst naar al uw individuele sitemaps.

Sitemap: https://dekvloersmeren.com//sitemap\_index.xml

Sitemap: https://dekvloersmeren.com//sitemap\_index.xml

# Optionele Host directive om voorkeursdomein aan te geven

# Let op: Google ondersteunt deze directive niet meer, maar andere crawlers mogelijk wel.

Host: https://www.yoursite.com

# ===============================================

# EINDE VAN ULTIEME ROBOTS.TXT MEGA EDITION v6.0

# ===============================================