Mesures de sécurité appliquées - Penpal Al

Ce document présente l'ensemble des mesures de sécurité implémentées dans la plateforme Penpal AI, en conformité avec les standards OWASP et les meilleures pratiques de sécurité applicative.

Vue d'ensemble de la sécurité

Architecture de sécurité multicouche

La plateforme Penpal Al implémente une approche **"Defence in Depth"** avec plusieurs couches de sécurité :

- 1. Sécurité réseau : Firewall, SSL/TLS, isolation containers
- 2. Authentification et autorisation : JWT, OAuth 2.0 PKCE, RBAC
- 3. Sécurité applicative : Validation input, protection injection, rate limiting
- 4. Sécurité des données : Chiffrement, pseudonymisation, audit logs
- 5. Monitoring sécuritaire: Détection anomalies, alerting, incidents

Référentiels de sécurité couverts

- **WASP Top 10** (2021) Vulnérabilités web critiques
- **WASP API Security Top 10** (2023) Sécurité APIs
- V NIST Cybersecurity Framework Gouvernance sécurité
- **RGPD** Protection données personnelles
- **V** OAuth 2.0 Security Best Practices Authentification moderne

OWASP Top 10 - Couverture détaillée

A01:2021 - Broken Access Control

Risque: Contrôles d'accès défaillants permettant actions non autorisées.

Mesures implémentées :

Authentification JWT robuste

```
// Extraction sécurisée du token (cookie + header)
function extractJwtFromCookieOrHeader(req: Request) {
   // Priorité au cookie HttpOnly sécurisé
   const tokenFromCookie = req.cookies?.auth_token;
   if (tokenFromCookie) {
      return tokenFromCookie;
   }

   // Fallback sur Authorization header
   return ExtractJwt.fromAuthHeaderAsBearerToken()(req);
}
```

Guards d'autorisation inter-services

```
// Protection communication service-to-service
@Injectable()
export class ServiceAuthGuard implements CanActivate {
  async canActivate(context: ExecutionContext): Promise<boolean> {
    const apiKey = request.headers["x-api-key"] as string;
    const serviceName = request.headers["x-service-name"] as string;
    // Validation service autorisé
    if (!serviceName || !this.allowedServices.includes(serviceName)) {
     throw new UnauthorizedException("Service not authorized");
    }
    // Validation clé API
    if (apiKey !== this.apiKey) {
      throw new UnauthorizedException("Invalid API key");
    }
   return true;
  }
}
```

RBAC (Role-Based Access Control)

• Rôles utilisateur : user, premium, admin

Permissions granulaires : Accès API basé sur rôles

• Session management : Validation token à chaque requête

Configuration production:

```
# JWT sécurisé
JWT_SECRET=complex_secret_256_bits_minimum
JWT_EXPIRATION=1h
JWT_REFRESH_EXPIRATION=7d

# Services autorisés
ALLOWED_SERVICES=auth-service,ai-service,payment-service
INTERNAL_API_KEY=service_api_key_complex_256_bits
```

A02:2021 – Cryptographic Failures

Risque: Données sensibles exposées par chiffrement faible ou absent.

Mesures implémentées :

Chiffrement des mots de passe

```
// Hashage bcrypt avec salt robuste
@Injectable()
export class AuthService {
  private readonly saltRounds = 12; // Configuration production

  async hashPassword(password: string): Promise<string> {
    return bcrypt.hash(password, this.saltRounds);
  }

  async validatePassword(password: string, hash: string): Promise<boolean> {
    return bcrypt.compare(password, hash);
   }
}
```

Configuration SSL/TLS

- Certificats Let's Encrypt : Renouvellement automatique
- TLS 1.2+: Protocoles sécurisés uniquement
- HSTS Headers : Forcer HTTPS navigation
- Secure cookies: HttpOnly, Secure, SameSite strict

Chiffrement au repos

```
# MongoDB avec chiffrement
MONGODB_URI=mongodb://user:password@mongodb-prod:27017/penpal-ai?ssl=true
# Redis avec authentification
REDIS_PASSWORD=complex_redis_password_production
```

Variables sensibles sécurisées

- Coolify Secrets: Variables chiffrées au repos
- Rotation périodique : Clés API renouvelées tous les 90 jours
- Séparation environnements : Secrets différents prod/staging

A03:2021 - Injection

Risque: Injection de code malveillant (SQL, NoSQL, XSS, etc.).

Mesures implémentées :

Validation et sanitisation des inputs

```
// Validation stricte avec class-validator
export class CreateUserDto {
  @IsString()
  @MinLength(2)
  @MaxLength(50)
  QMatches(/^[a-zA-ZA-y\s'-]+$/, {
    message: 'Name must contain only letters, spaces, hyphens and
apostrophes'
  })
  name: string;
  @IsEmail()
  @Transform(({ value }) => value.toLowerCase().trim())
  email: string;
  @IsString()
  @MinLength(8)
  \alpha_{A-Z}(-x^2)(?=.*[A-Z])(?=.*Ad)(?=.*[0$!%*?\&])[A-Za-z\d0$!%*?
&]/, {
    message: 'Password must contain uppercase, lowercase, number and
special character'
  })
  password: string;
}
```

Protection contre injection paramètres

```
// SecurityService - Détection patterns suspects
private validateParameters(ip: string, userAgent: string, params:
Record<string, any>): void {
  const suspiciousPatterns = [
    /<script/i,
                        // XSS
    /javascript:/i, // XSS
    /onclick/i,  // XSS Event handlers
/onerror/i,  // XSS Event handlers
/eval\(/i,  // Code injection
/expression\(/i,  // CSS expression injection
                 // CSS URL injection
    /url\(/i,
    /import\(/i, // ES6 import injection
  1:
  for (const [key, value] of Object.entries(params)) {
    if (typeof value === "string") {
      if (suspiciousPatterns.some(pattern => pattern.test(value))) {
        this.logSecurityEvent(ip, userAgent, "parameter_injection",
           `Suspicious parameter detected: ${key}=${value}`);
        throw new UnauthorizedException("Invalid request parameters");
    }
```

```
}
```

Validation globale NestJS

Protection MongoDB (NoSQL injection)

• Mongoose : Schémas typés stricts

• Validation côté serveur : Tous inputs validés

• Queries paramétrées : Pas de concaténation directe

A04:2021 - Insecure Design

Risque: Failles de conception architecturale.

Mesures implémentées :

Architecture microservices sécurisée

- Principe moindre privilège : Chaque service accès minimal requis
- Isolation réseau : Communication inter-services chiffrée
- API Gateway pattern : Point d'entrée unique avec contrôles
- Circuit breaker: Protection contre cascades de pannes

OAuth 2.0 avec PKCE

```
// Protection contre interception codes d'autorisation
@Injectable()
export class OAuthService {
   // PKCE (Proof Key for Code Exchange)
   private generatePKCE() {
     const codeVerifier = this.generateRandomString(128);
     const codeChallenge = crypto
        .createHash('sha256')
```

```
.update(codeVerifier)
   .digest('base64url');

return { codeVerifier, codeChallenge };
}

// State parameter pour CSRF protection
private generateState() {
   return crypto.randomBytes(32).toString('hex');
}
```

Gestion des erreurs sécurisée

- Logs détaillés côté serveur : Debug complet
- Messages génériques côté client : Pas de leak d'information
- Monitoring erreurs : Alerting sur patterns suspects

A05:2021 – Security Misconfiguration

Risque: Configurations par défaut ou mal sécurisées.

Mesures implémentées :

Configuration sécurisée production

```
// Headers de sécurité avec Helmet
app.use(helmet({
 contentSecurityPolicy: {
    directives: {
      defaultSrc: ["'self'"],
      scriptSrc: ["'self'", "'unsafe-inline'", "https://js.stripe.com"],
      styleSrc: ["'self'", "'unsafe-inline'",
"https://fonts.googleapis.com"],
      fontSrc: ["'self'", "https://fonts.gstatic.com"],
      imgSrc: ["'self'", "data:", "https:"],
      connectSrc: ["'self'", "https://api.openai.com",
"https://api.anthropic.com"]
   }
 },
 hsts: {
   maxAge: 31536000, // 1 an
    includeSubDomains: true,
   preload: true
 }
}));
```

CORS strictement configuré

```
// Configuration CORS restrictive
app.enableCors({
  origin: process.env.NODE_ENV === 'production'
     ? ['https://app.penpal-ai.maksou.dev', 'https://staging.app.penpal-ai.maksou.dev']
     : ['http://localhost:3000'],
     methods: ['GET', 'POST', 'PUT', 'DELETE', 'PATCH'],
     credentials: true,
     optionsSuccessStatus: 200
});
```

Désactivation fonctionnalités développement

```
# Production uniquement
SWAGGER_ENABLED=false
DEBUG_MODE=false
VERBOSE_LOGGING=false
NODE_ENV=production
```

Configuration cookies sécurisée

A06:2021 – Vulnerable and Outdated Components

Risque: Dépendances avec vulnérabilités connues.

Mesures implémentées :

Audit automatique dépendances

```
# Dans CI/CD pipeline
npm audit --audit-level=moderate
npm audit fix
# Outils de scan sécurité
```

```
npm install -g yarn-audit-fix
yarn audit --level moderate
```

Politique de mise à jour

- Scan hebdomadaire : Vulnerabilités nouvelles
- Mise à jour critique : < 48h pour vulnérabilités critiques
- Mise à jour majeure : Testing complet avant déploiement
- Dependencies pinning: Versions exactes en production

Stack moderne et maintenue

A07:2021 - Identification and Authentication Failures

Risque: Authentification faible ou défaillante.

Mesures implémentées :

Politique de mots de passe robuste

```
// Validation frontend + backend
const passwordSchema = z
   .string()
   .min(8, 'Minimum 8 caractères')
   .regex(/[A-Z]/, 'Au moins 1 majuscule')
   .regex(/[a-z]/, 'Au moins 1 minuscule')
   .regex(/[0-9]/, 'Au moins 1 chiffre')
   .regex(/[^A-Za-z0-9]/, 'Au moins 1 caractère spécial');
```

Rate limiting intelligent

```
// Protection brute force
@Injectable()
export class SecurityService {
```

```
private readonly maxAttempts = 5;
private readonly lockoutDuration = 15 * 60 * 1000; // 15 minutes

async checkRateLimit(ip: string, userAgent: string): Promise<void> {
   const key = `rate_limit:${ip}`;
   const attempts = await this.redis.get(key);

if (attempts && parseInt(attempts) >= this.maxAttempts) {
   const ttl = await this.redis.ttl(key);
   throw new TooManyRequestsException(
        `Too many attempts. Try again in ${Math.ceil(ttl / 60)} minutes`
      );
   }
}
```

Session management sécurisé

```
// Gestion tokens JWT
export class JwtStrategy extends PassportStrategy(Strategy) {
  async validate(payload: JwtPayload) {
    // Validation utilisateur encore actif
    const user = await this.authService.validateUserById(payload.sub);
    if (!user) {
     throw new UnauthorizedException("User no longer exists");
    }
    // Validation expiration
    if (payload.exp < Date.now() / 1000) {</pre>
      throw new UnauthorizedException("Token expired");
    }
    return {
      id: payload.sub,
      email: payload.email,
      roles: payload.roles,
    };
 }
}
```

OAuth 2.0 sécurisé

- PKCE obligatoire : Protection mobile/SPA
- State parameter : Protection CSRF
- ID Token validation : Signature + expiration + audience
- Scope minimal : Permissions strictement nécessaires

Risque: Intégrité code et données compromise.

Mesures implémentées :

CI/CD Pipeline sécurisé

```
# GitHub Actions avec vérifications sécurité
name: Security Pipeline
on: [push, pull_request]
jobs:
  security-scan:
    runs-on: ubuntu-latest
    steps:
      - name: Checkout code
        uses: actions/checkout@v4
      - name: Dependency vulnerability scan
        run: npm audit --audit-level=moderate
      - name: SAST scan
        uses: github/super-linter@v4
      - name: Container security scan
        uses: aquasecurity/trivy-action@master
        with:
          image-ref: 'ghcr.io/${{ github.repository }}:latest'
```

Intégrité des builds

- Multi-stage Dockerfile: Build reproductible
- Image signing : Vérification intégrité containers
- Lock files : Dependencies versions exactes
- Checksum verification: Validation packages

Audit trails complets

```
// Logging sécuritaire structuré
@Injectable()
export class SecurityService {
  private logSecurityEvent(ip: string, userAgent: string, event: string,
  details: string): void {
    const securityLog = {
        timestamp: new Date().toISOString(),
        type: 'security_event',
        event,
        details,
        ip,
        userAgent,
```

```
service: 'auth-service',
    environment: process.env.NODE_ENV,
    sessionId: this.generateSessionId(),
    correlationId: this.generateCorrelationId()
};

this.logger.warn(JSON.stringify(securityLog));
}
```

A09:2021 - Security Logging and Monitoring Failures

Risque: Détection incidents manquée par logging insuffisant.

Mesures implémentées :

Logging sécuritaire centralisé

```
// Events sécurité trackés
type SecurityEvent =
  | 'login_success' | 'login_failed' | 'password_reset'
  | 'suspicious_activity' | 'rate_limit_exceeded'
  | 'parameter_injection' | 'invalid_token';
@Injectable()
export class SecurityService {
  // Logging succès authentification
  logSuccessfulAuth(ip: string, userAgent: string, email: string): void {
    this.logSecurityEvent(ip, userAgent, 'login_success', `User ${email}
authenticated`);
  }
  // Logging tentatives suspectes
  logSuspiciousActivity(ip: string, userAgent: string, reason: string):
void {
   this.logSecurityEvent(ip, userAgent, 'suspicious_activity', reason);
}
```

Monitoring temps réel

```
# Métriques sécurité Prometheus
# - Taux d'erreurs authentification
# - Tentatives brute force par IP
# - Patterns d'injection détectés
# - Durée sessions anormales
# - Accès APIs non autorisés
```

```
# Alerting Grafana
# - > 10 échecs auth/minute => Alert immediate
# - Injection pattern détecté => Alert critique
# - Service auth down => Alert critique
# - Trafic suspect => Investigation
```

SIEM Integration (future)

• Export logs : Format SIEM standard

• Correlation rules : Détection patterns complexes

• Incident response : Procédures automatisées

A10:2021 - Server-Side Request Forgery (SSRF)

Risque: Requêtes serveur manipulées vers resources internes.

Mesures implémentées :

Validation URLs externes

```
// Whitelist domaines autorisés
@Injectable()
export class HttpClientService {
  private readonly allowedDomains = [
    'api.openai.com',
    'api.anthropic.com',
    'api.stripe.com',
    'api.sendgrid.com'
  ];
  async makeRequest(url: string): Promise<any> {
    const urlObj = new URL(url);
    // Validation domaine autorisé
    if (!this.allowedDomains.includes(urlObj.hostname)) {
     throw new ForbiddenException(`Domain not allowed:
${urlObj.hostname}`);
    }
    // Validation protocole sécurisé
    if (!['https:'].includes(urlObj.protocol)) {
     throw new ForbiddenException(`Protocol not allowed:
${url0bj.protocol}`);
   return this.httpService.get(url).toPromise();
  }
}
```

Isolation réseau

- Network policies : Communication inter-services restreinte
- Egress filtering : Sortie internet contrôlée
- DNS filtering : Résolution domaines malveillants bloquée

OWASP API Security Top 10 - Couverture

API1:2023 – Broken Object Level Authorization

Protection implémentée :

```
// Validation ownership ressources
@Get(':id')
async getUserData(@Param('id') id: string, @CurrentUser() user: User) {
   // Vérification utilisateur peut accéder à cette ressource
   if (id !== user.id && !user.roles.includes('admin')) {
      throw new ForbiddenException('Access denied to this resource');
   }
   return this.userService.findOne(id);
}
```

API2:2023 - Broken Authentication

Protection implémentée :

- JWT avec expiration courte (1h)
- · Refresh tokens avec rotation
- Rate limiting agressif sur endpoints auth
- MFA preparation (hooks pour future implémentation)

API3:2023 – Broken Object Property Level Authorization

Protection implémentée :

```
// DTOs avec propriétés limitées par rôle
export class UserResponseDto {
  @Expose()
  id: string;

  @Expose()
  email: string;

  @Expose()
  @Transform(({ obj, key }) => obj.user?.roles?.includes('admin') ?
  obj[key] : undefined)
```

```
adminOnlyField?: string;
}
```

API4:2023 – Unrestricted Resource Consumption

Protection implémentée :

```
// Rate limiting par endpoint
@UseGuards(ThrottlerGuard)
@Throttle(10, 60) // 10 requests per minute
@Post('generate-text')
async generateText(@Body() dto: GenerateTextDto) {
    // Validation taille input
    if (dto.prompt.length > 4000) {
        throw new BadRequestException('Prompt too long');
    }
    return this.aiService.generateText(dto);
}
```

API5:2023 - Broken Function Level Authorization

Protection implémentée :

```
// Guards de rôles granulaires
@UseGuards(JwtAuthGuard, RolesGuard)
@Roles('admin')
@Delete('users/:id')
async deleteUser(@Param('id') id: string) {
   return this.userService.delete(id);
}
```

Sécurité Infrastructure

Sécurité réseau

Firewall Hetzner Cloud

```
# Ports ouverts production
22/tcp # SSH (IP whitelistée admin)
80/tcp # HTTP (redirect vers HTTPS)
443/tcp # HTTPS
8080/tcp # Coolify (IP whitelistée admin)
# Tous autres ports fermés par défaut
# DDoS protection Hetzner activée
```

SSL/TLS Configuration

```
# Let's Encrypt via Coolify
# - TLS 1.2+ uniquement
# - Perfect Forward Secrecy
# - HSTS headers
# - OCSP Stapling
# - Certificate Transparency

# Test sécurité SSL
# A+ rating sur SSL Labs
```

Conteneurisation sécurisée

Docker Security

```
# Multi-stage builds pour images minimales
FROM node:20-alpine AS production

# Non-root user
RUN addgroup -S appgroup && adduser -S appuser -G appgroup
USER appuser

# Health checks
HEALTHCHECK --interval=30s --timeout=3s --start-period=5s --retries=3 \
        CMD wget --no-verbose --tries=1 --spider
http://localhost:${PORT}/health || exit 1
```

Secrets Management

• Coolify Secrets : Variables chiffrées

Rotation automatique : Scripts dédiés
 Accès contrôlé : RBAC Coolify

• Audit secrets : Logs accès

Monitoring sécuritaire

Métriques sécurité

```
# Prometheus metrics
penpal_auth_failed_attempts_total: # Tentatives échouées
penpal_suspicious_activity_total: # Activités suspectes
penpal_rate_limit_exceeded_total: # Rate limiting
penpal_security_events_total: # Events sécurité globaux
```

Alerting

```
# Grafana alerts
- name: "Authentication Failures"
  condition: "penpal_auth_failed_attempts_total > 50 in 5m"
  action: "Email admin + Block IP"

- name: "Injection Attempts"
  condition: "penpal_security_events_total{type='injection'} > 0"
  action: "Immediate alert + Investigation"
```

Conformité et gouvernance

RGPD Compliance

Protection données personnelles

```
// Pseudonymisation utilisateurs
export class User {
   @Prop({ required: true, unique: true })
   id: string; // UUID v4 non réversible

   @Prop({ required: true })
   email: string;

   @Prop()
   @Transform(({ value }) => this.hashPersonalData(value))
   personalData?: string; // Hashé si sensible
}
```

Droits utilisateurs

- Droit d'accès : API /api/v1/users/me/data
- Droit rectification : API /api/v1/users/me/update
- Droit effacement : API /api/v1/users/me/delete
- Portabilité: Export JSON complet

Audit et compliance

Logs d'audit

```
// Traçabilité complète actions utilisateur
@Injectable()
export class AuditService {
  async logUserAction(userId: string, action: string, resource: string,
```

```
details?: any): Promise<void> {
    const auditLog = {
        timestamp: new Date().toISOString(),
        userId,
        action,
        resource,
        details,
        ip: this.request.ip,
        userAgent: this.request.headers['user-agent'],
        sessionId: this.extractSessionId()
    };
    await this.auditRepository.create(auditLog);
}
```

Rapports compliance

• Mensuel: Rapport sécurité complet

• Incident: Post-mortem obligatoire

• Audit externe : Preparation documentation

• Certification: Roadmap ISO 27001

Gestion d'incidents sécurité

Détection et classification

Niveaux de criticité

- 1. **Critique** : Accès non autorisé données, service compromis
- 2. Élevé: Tentative intrusion, vulnérabilité exploitable
- 3. Moyen: Anomalie suspecte, configuration non optimale
- 4. Faible: Violation politique, alerte préventive

Procédures de réponse

Incident critique (< 15 minutes)

```
# 1. Containment immédiat
# - Isolation service compromis
# - Blocage IP/utilisateur suspect
# - Préservation preuves (logs, snapshots)

# 2. Assessment impact
# - Données compromises ?
# - Services affectés ?
# - Utilisateurs impactés ?

# 3. Eradication
```

```
# - Correction vulnérabilité
# - Changement secrets compromis
# - Mise à jour systèmes

# 4. Recovery
# - Restauration service
# - Validation intégrité données
# - Monitoring renforcé

# 5. Lessons learned
# - Post-mortem incident
# - Amélioration processus
# - Formation équipe
```

Communication incident

• Interne : Équipe technique + management

• Externe : Utilisateurs si données compromises

• Légal : CNIL si requis (< 72h)

• Assurance : Déclaration si dommages

Tests de sécurité

Pentesting régulier

```
# Tests automatisés
npm run security:scan  # SAST daily
npm run deps:audit  # Dependencies weekly
npm run container:scan  # Container security

# Tests manuels
# - Pentest externe : Semestriel
# - Code review sécurité : Chaque PR
# - Configuration review : Mensuel
# - Social engineering : Annuel
```

Bug bounty program (futur)

Scope : APIs publiques, frontend
Rewards : Selon criticité OWASP
Disclosure : 90 jours responsible

• Hall of fame: Reconnaissance chercheurs

Roadmap sécurité

Court terme (Q1 2025)

• V MFA implementation : 2FA TOTP/SMS

- **WAF deployment** : Cloudflare protection
- **SIEM integration**: Logs centralisés
- **V** Backup encryption : Chiffrement backups

Moyen terme (Q2-Q3 2025)

- **Zero Trust Architecture** : Micro-segmentation
- **API Gateway** : Centralisation sécurité APIs
- Secrets automation : Vault HashiCorp
- Compliance certification : SOC 2 Type II

Long terme (Q4 2025+)

- ML-based anomaly detection : IA sécurité
- 📋 Homomorphic encryption : Chiffrement avancé
- 📋 Quantum-resistant crypto : Préparation post-quantique
- ISO 27001 certification : Standard international

Métriques et KPIs sécurité

Indicateurs techniques

```
# Disponibilité sécurisée
Security_Uptime: 99.9%
                                           # Services sécurité
opérationnels
Failed Auth Rate: < 1%
                                           # Taux échecs authentification
Injection_Detection_Rate: 100%
                                          # Détection tentatives injection
Vulnerability_MTTR: < 24h</pre>
                                           # Temps correction vulnérabilité
# Performance sécurité
Auth_Response_Time: < 200ms</pre>
                                          # Latence authentification
Security_Scan_Coverage: > 95%
                                          # Couverture scans sécurité
Incident_Response_Time: < 15min</pre>
                                         # Temps réponse incident
critique
```

Indicateurs métier

```
# Confiance utilisateur
User_Security_Satisfaction: > 90%
                                       # Satisfaction sécurité
                                       # Incidents fuite données
Data_Breach_Incidents: 0
                                       # Score conformité RGPD
Compliance_Score: > 95%
Security_Training_Completion: 100%
                                        # Formation équipe
# Coût sécurité
Security_Investment_ROI: 300%
                                       # ROI investissement sécurité
Incident_Cost_Reduction: 80%
                                        # Réduction coût incidents
                                       # Réduction prime assurance
Insurance_Premium_Reduction: 20%
```

Conclusion

La plateforme Penpal Al implémente une **stratégie de sécurité multicouche** alignée sur les standards OWASP et les meilleures pratiques industry. L'approche **"Security by Design"** garantit protection robuste contre les menaces actuelles tout en préparant l'évolution vers des standards futurs.

Points forts

- **Couverture OWASP Top 10 complète** avec contrôles détaillés
- **V** Architecture microservices sécurisée avec isolation réseau
- **Authentification moderne** JWT + OAuth 2.0 PKCE
- **Monitoring sécuritaire temps réel** avec alerting automatique
- Conformité RGPD avec audit trails complets