Azure OpenAI Security: Best Practice on using Azure OpenAI service

Beata Zalewa

zalnet

About me

Security Architect

Consultant

Microsoft Certified Trainer

AI & Cybersecurity Practitioner

Developer

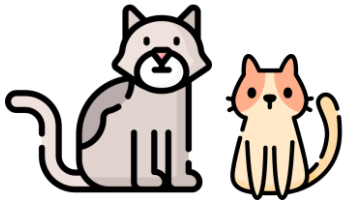
Freelancer

Azure @ ❤

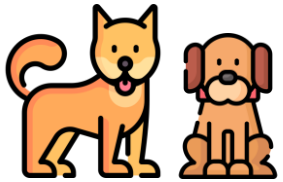
Google Cloud

1 Husband

1 Daughter



2 cats



2 dogs

Detective stories

Photography

zalnet BEATA ZALEWA

https://www.linkedin.com/in/beatazalewa/

beata@zalnet.pl

BLOG https://zalnet.pl/

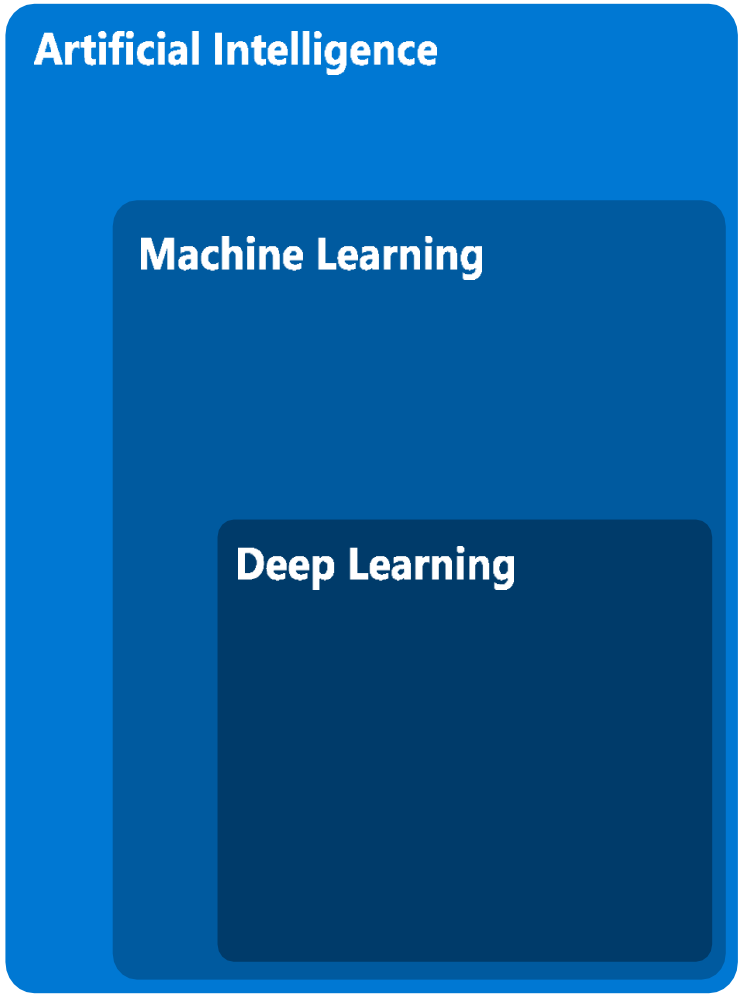
zalnet
BEATA ZALEWA

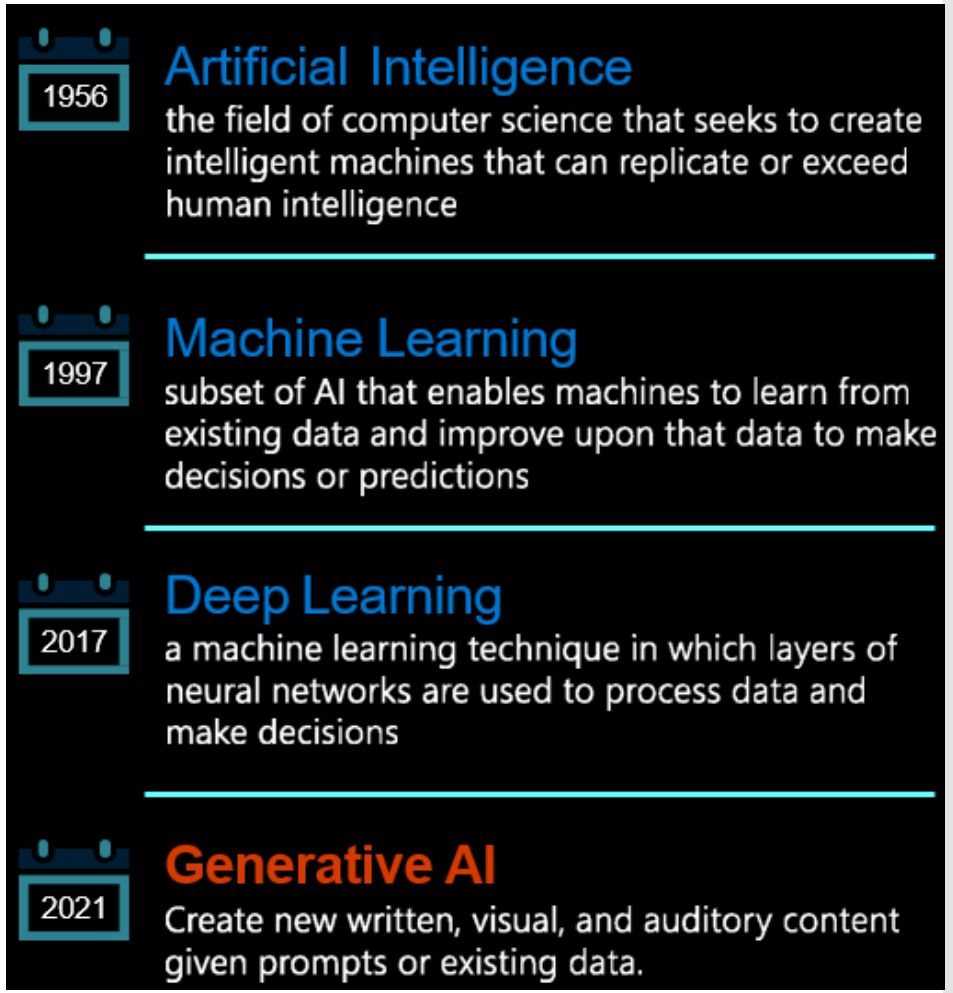
Introduction to GenAI Applications

- GenAI applications are those that use large language models (LLMs) to generate natural language texts or perform natural language understanding tasks.
- LLMs are powerful tools that can enable various scenarios such as content creation, summarization, translation, question answering, and conversational agents.

zalnet
BEATA ZALEWA

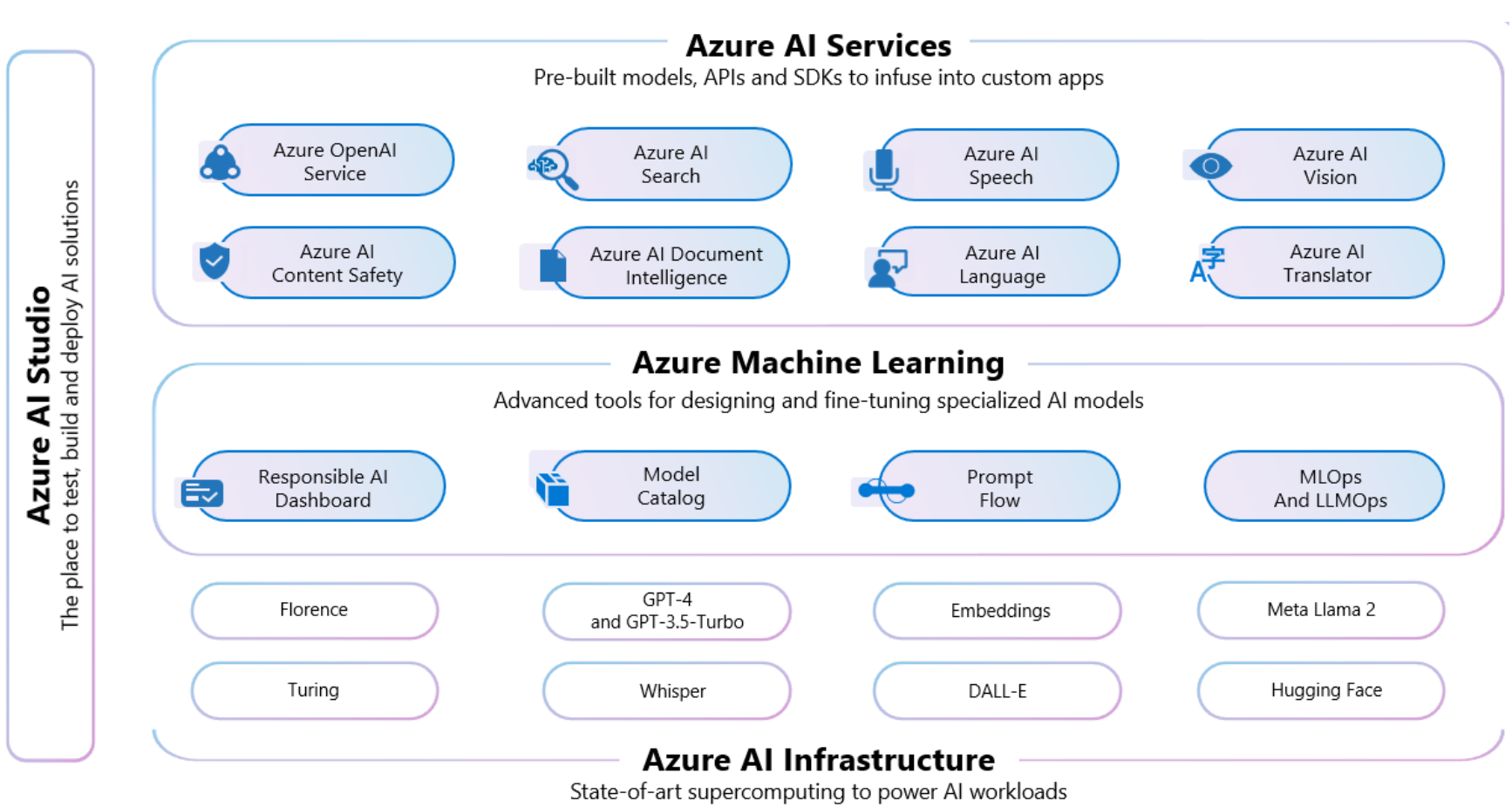
Introduction to Azure OpenAI





zalnet
BEATA ZALEWA

The Microsoft Azure AI Portfolio



zalnet
BEATA ZALEWA

Security Challenges in GenAI Applications

- **Data Protection**: Confidentiality and integrity of training and query data.
- **Service Reliability**: Ensuring availability and reliability of LLM services.
- **Misuse Prevention**: Preventing misuse or abuse by malicious actors or unintended users.
- **Output Monitoring**: Auditing outputs for quality, accuracy, and compliance.
- **Ethical Management**: Managing ethical and social implications of outputs.

zalnet
BEATA ZALEWA

Jailbreak Attack or User Prompt Injection Attack (UPIA)

Intentional attempt by a user to

Exploit the vulnerabilities of an llm-powered system

Bypass its safety mechanisms

Provoke restricted behaviors.

zalnet
BEATA ZALEWA

Jailbreak risk detection or Prompt Shields for User Prompts

Unified API that analyzes LLM inputs and detects **user prompt and document** attacks

Model that identifies anomalies in user prompts as potential jailbreak attacks

Enhances the security of LLM deployments

zalnet
BEATA ZALEWA

The principle of shared responsibility

- **The principle of shared responsibility** highlights that security in the cloud is a two-way street.
- While Microsoft ensures the security of the Azure OpenAI services, it is the customer's responsibility to secure their end of the interaction.

zalnet
BEATA ZALEWA

Microsoft responsibility

Protecting the Azure infrastructure

Making sure that the Azure OpenAI services are secure by default

Providing identity and access management capabilities

zalnet
BEATA ZALEWA

Customers responsibility

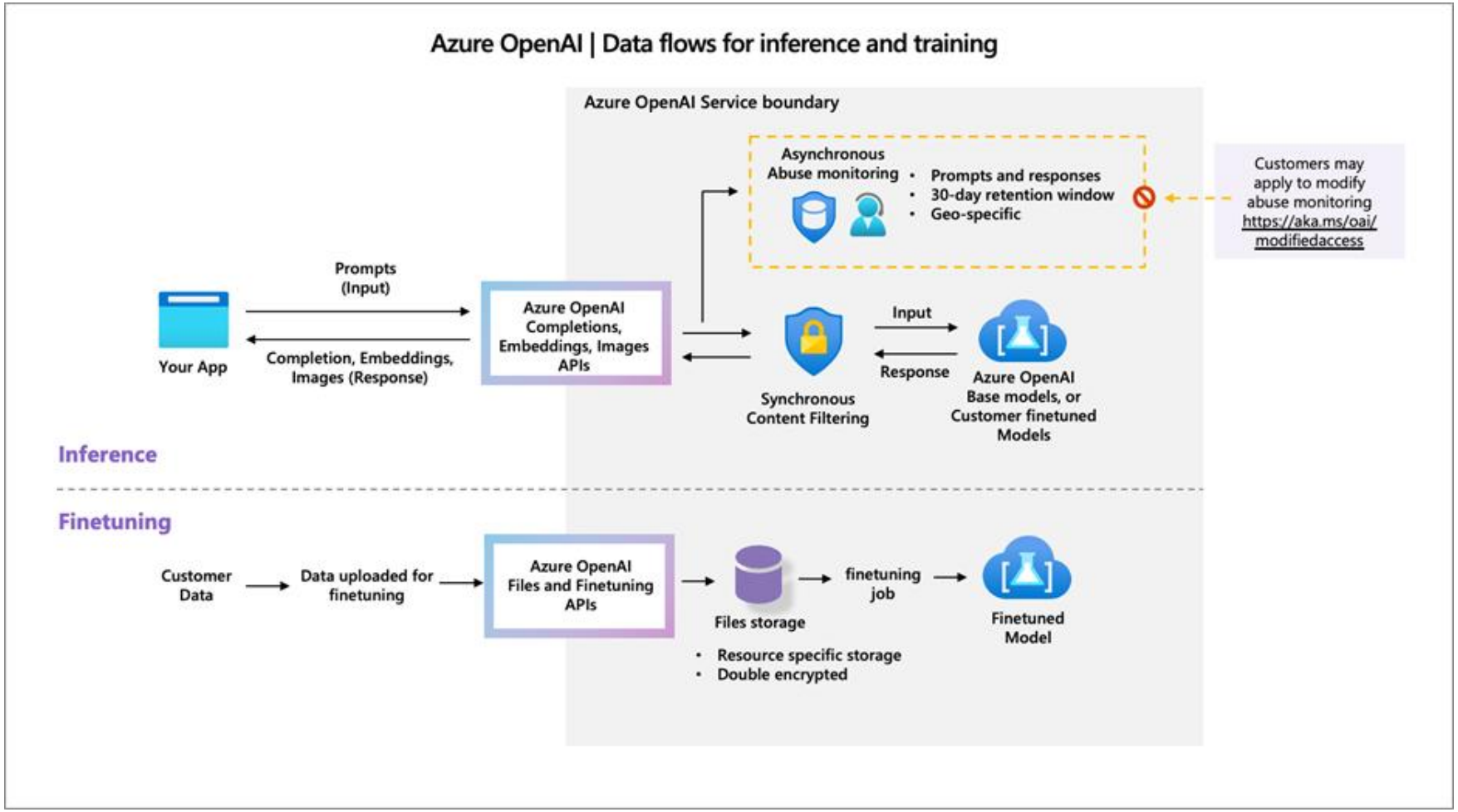
Setting appropriate access controls and permissions for their use of Azure OpenAI

Protecting their Azure credentials and managing access to their Azure subscription

Ensuring that the security measures around their applications are adequate, including input validation, secure application logic, and proper handling of Azure OpenAI outputs

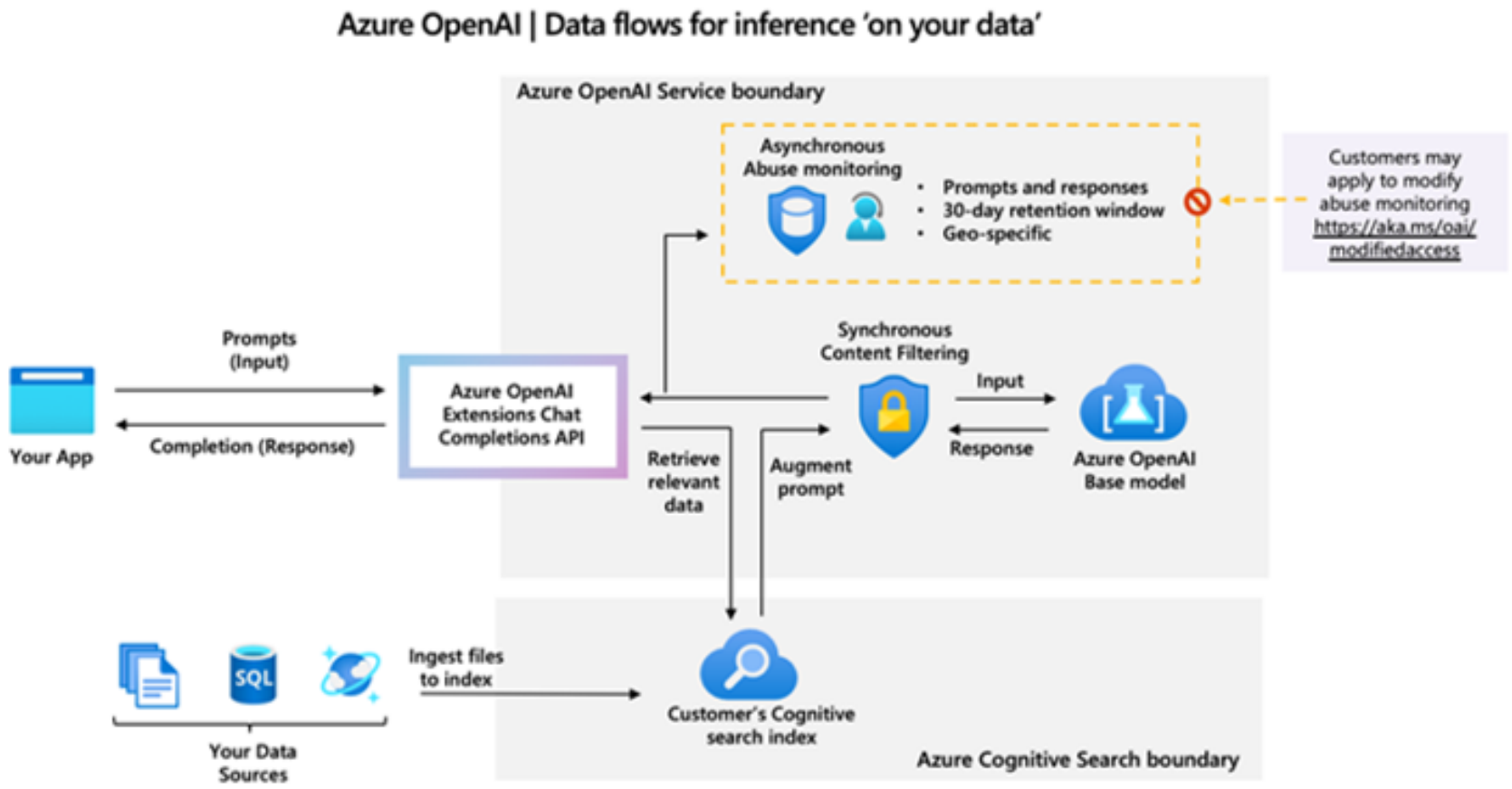
zalnet
BEATA ZALEWA

How does the Azure OpenAI Service process data?



zalnet
BEATA ZALEWA

Augmenting prompts with data retrieved from your data sources to "ground" the generated results



zalnet
BEATA ZALEWA

Data Processing Foundations

Core principles of data processing:

- Secure data processing
- Efficient algorithms
- Real-time processing
- Compliance with standards
- Scalability

zalnet
BEATA ZALEWA

Core principles of data storage

- Azure OpenAI’s approach to data storage and processing is built on strong, secure foundations.
- User data is guarded against unauthorized access, ensuring privacy and security in line with Microsoft's data protection standards.

zalnet
BEATA ZALEWA

Core principles of data storage

- Regarding the types of processed data, Azure OpenAI processes your input - prompts and the ensuing AI-generated answers.
- It also processes data submitted for the purpose of training tailored AI models.

zalnet
BEATA ZALEWA

Privacy

- **Privacy** in Azure OpenAI refers to how user data is handled in terms of access and usage.
- It ensures that data such as inputs, interactions, and outputs are used in a way that respects customer confidentiality.
- Azure OpenAI's privacy protocol dictates that Microsoft does not view or use this data for its own purposes, unless explicitly permitted for services like model fine-tuning.

zalnet
BEATA ZALEWA

Privacy Protection

Ensuring user privacy:

- Data anonymization
- User consent
- Data minimization
- Privacy policies
- Compliance with GDPR

zalnet
BEATA ZALEWA

Data privacy

- As for the usage of data, the **principles are simple**.
- The input you provide, along with the AI responses - your data - **remains yours.**
- Microsoft does not use this data to better their own AI offerings. It is kept private unless you choose to use it to train customized AI models.

zalnet
BEATA ZALEWA

Data privacy

- The service's content generation capacity comes with a commitment to safety.
- Input is carefully processed to produce responses, and content filters are in place to prevent the generation of problematic content.

zalnet
BEATA ZALEWA

Data privacy

- Customizing models is a service feature that's handled with care.
- You can train Azure OpenAI models with your data, knowing that it remains secure and for your use only.
- Privacy measures are about control over and the ethical handling of data.

zalnet
BEATA ZALEWA

Data privacy

- To prevent abuse, Azure OpenAI comes equipped with robust content filtering and monitoring systems.
- This is to make sure that the generation of content complies with guidelines and that nothing harmful slips through.

zalnet
BEATA ZALEWA

Privacy Policies

Privacy policies in place:

- Policy overview
- User rights
- Data usage
- Data sharing
- Policy updates

zalnet
BEATA ZALEWA

Compliance Standards

Compliance with industry standards:

- ISO certifications
- GDPR compliance
- HIPAA compliance
- SOC 2 compliance
- Regular compliance audits

zalnet
BEATA ZALEWA

DEMOS!!!

Azure OpenAI Studio

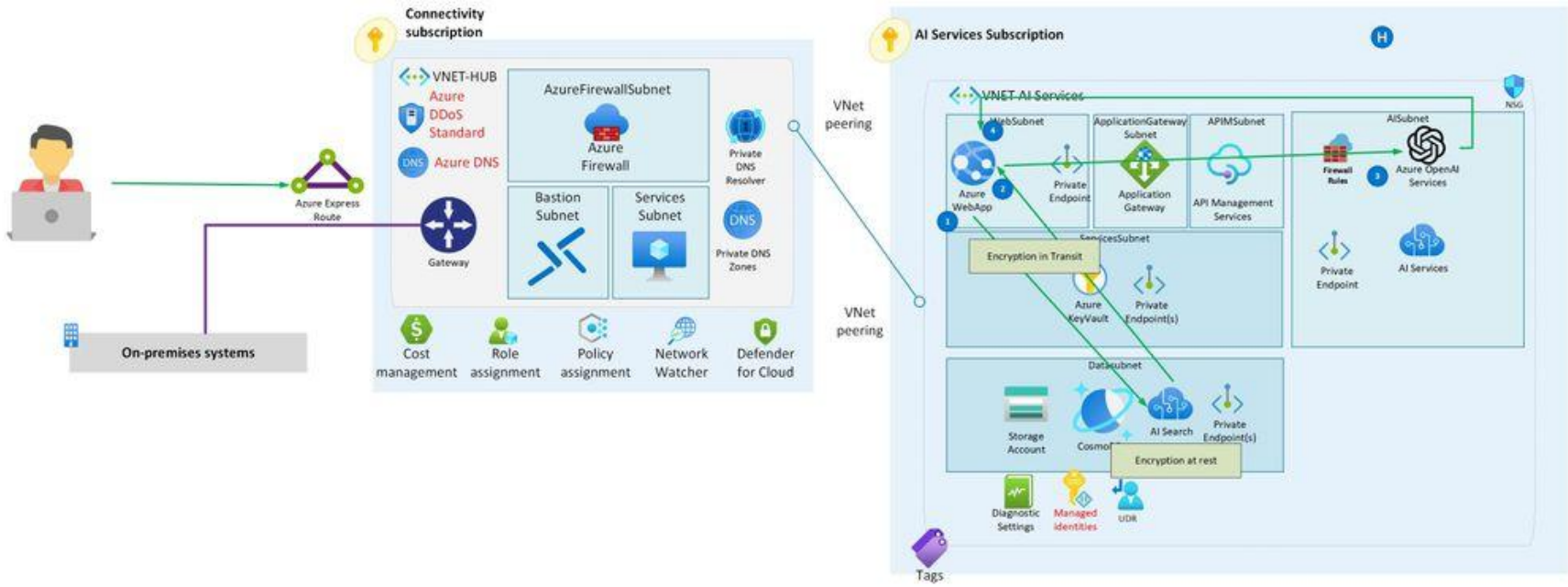
zalnet
BEATA ZALEWA

Security

- **Security**, on the other hand, is about protecting data from unauthorized access, breaches, and other forms of compromise.
- Azure OpenAI employs a range of security measures, such as encryption in transit and at rest, to safeguard data against threats.
- Microsoft's infrastructure provides a secure environment designed to shield your data from security risks.

zalnet
BEATA ZALEWA

Data security



zalnet
BEATA ZALEWA

Security Measures

Security measures in place:

- Encryption techniques
- Access controls
- Monitoring and logging
- Incident response regular audits

zalnet
BEATA ZALEWA

Access Controls

Managing access to data:

- Role-based access control (RBAC)
- Multi-factor authentication (MFA)
- Least privilege principle
- Access reviews
- User activity monitoring

zalnet
BEATA ZALEWA

Monitoring and Logging

Monitoring and logging practices:

- Continuous monitoring
- Log management
- Threat detection
- Incident response
- Compliance reporting

zalnet
BEATA ZALEWA

Incident Response

Handling security incidents:

- Incident response plan
- Detection and analysis
- Containment and eradication
- Recovery and lessons learned
- Communication protocols

zalnet
BEATA ZALEWA

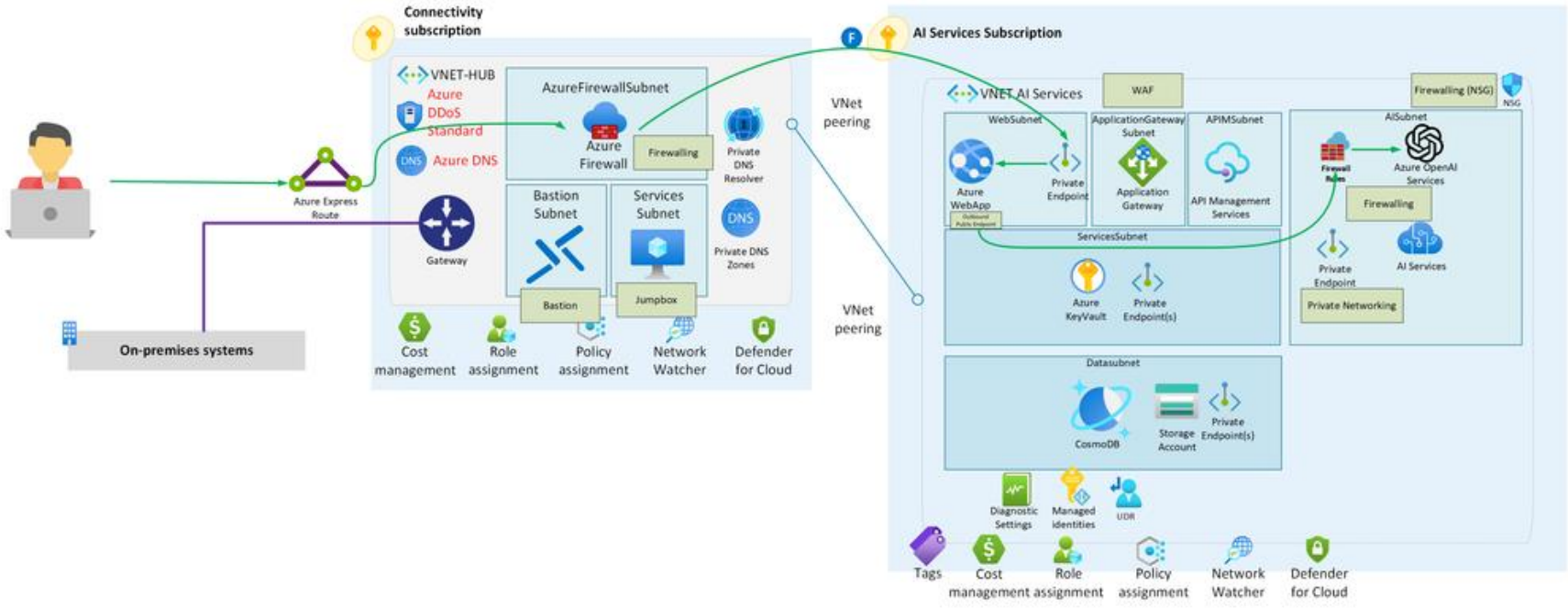
Data Encryption

Encryption methods used:

- Encryption at rest
- Encryption in transit
- Key management
- Advanced encryption standards
- Data Classification and Sensitivity

zalnet
BEATA ZALEWA

Network security



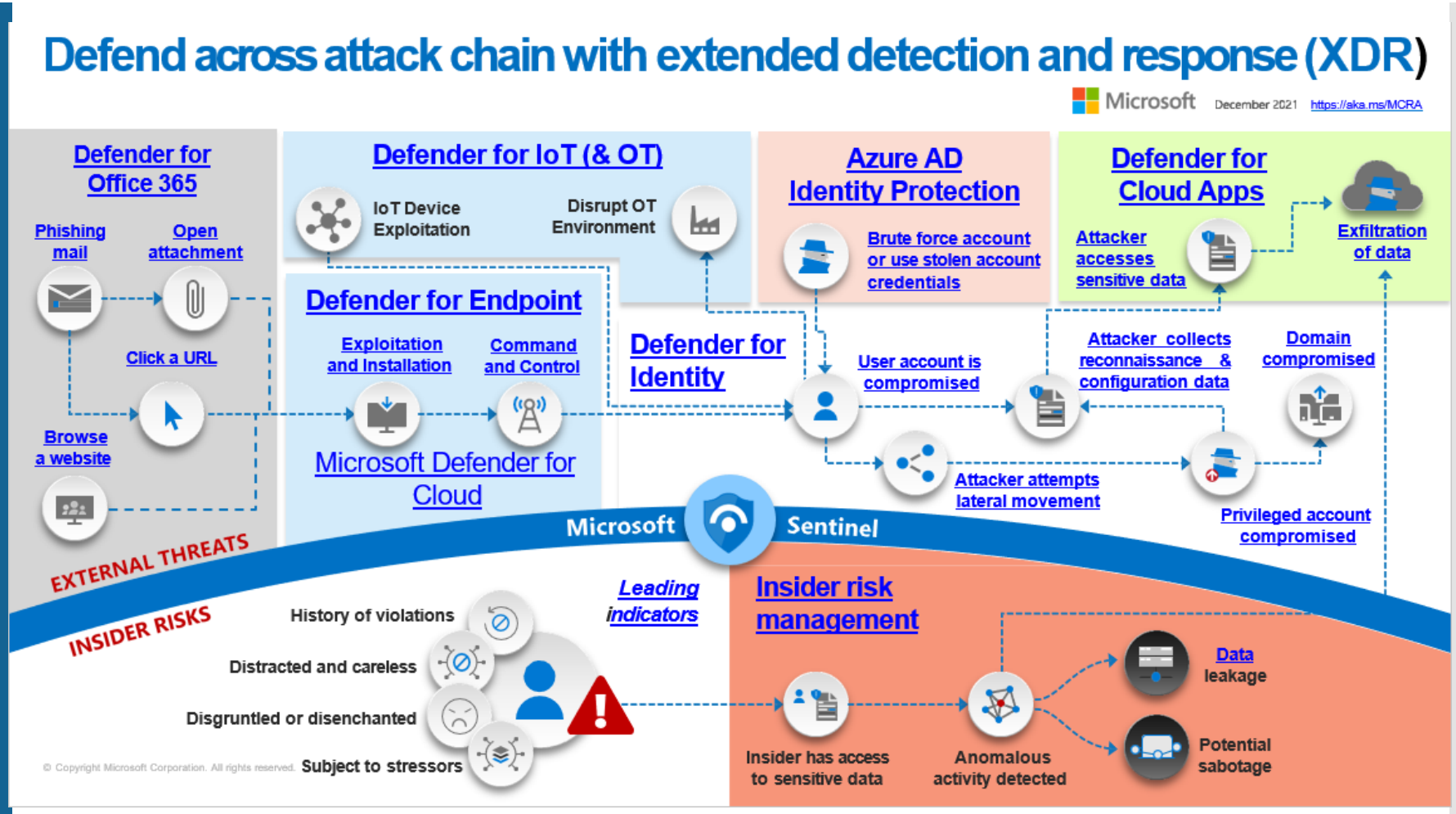
zalnet
BEATA ZALEWA

DEMOS!!!

Azure portal

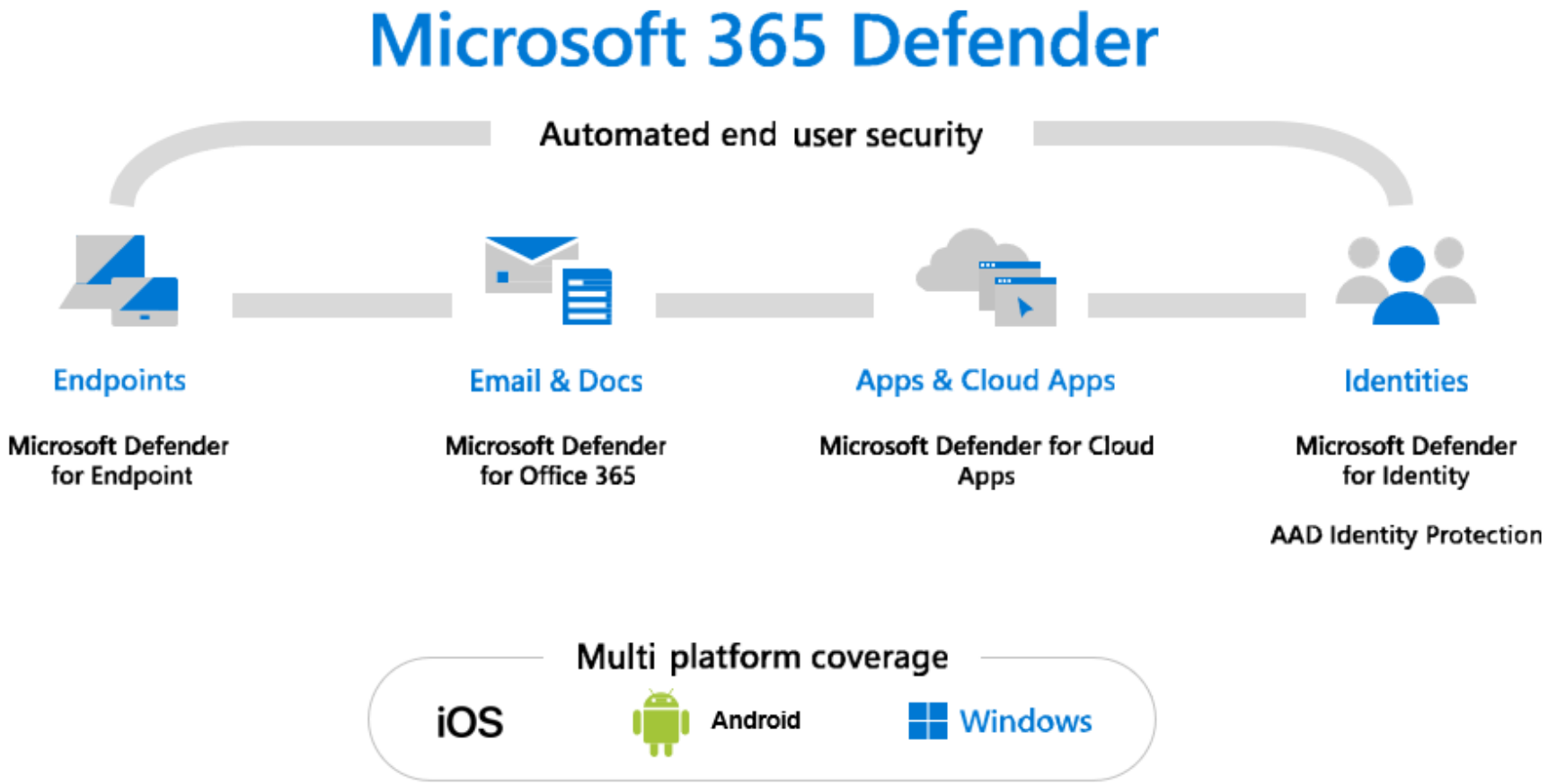
zalnet
BEATA ZALEWA

Defenders



zalnet
BEATA ZALEWA

Microsoft 365 Defender

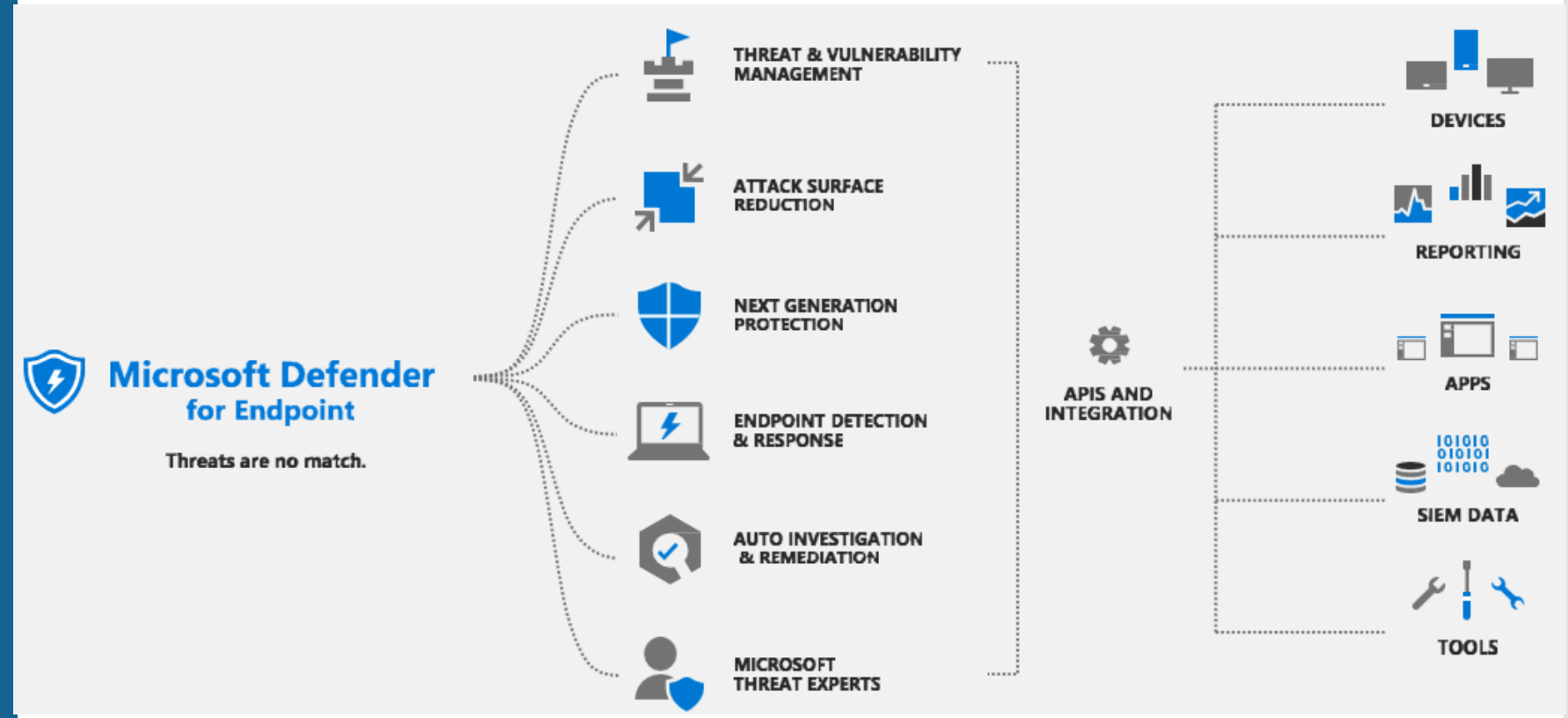


zalnet
BEATA ZALEWA

Defender for Endpoint

zalnet
BEATA ZALEWA

Connecting with the platform



zalnet
BEATA ZALEWA

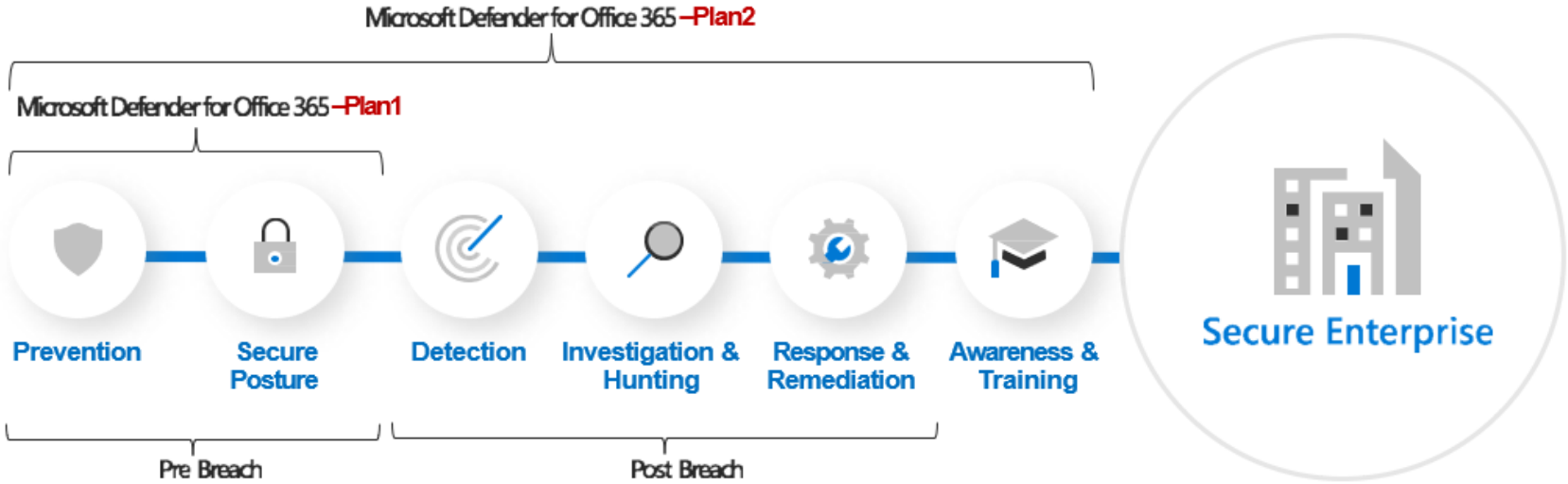
Defender for Endpoint

zalnet
BEATA ZALEWA

Defender for Office 365

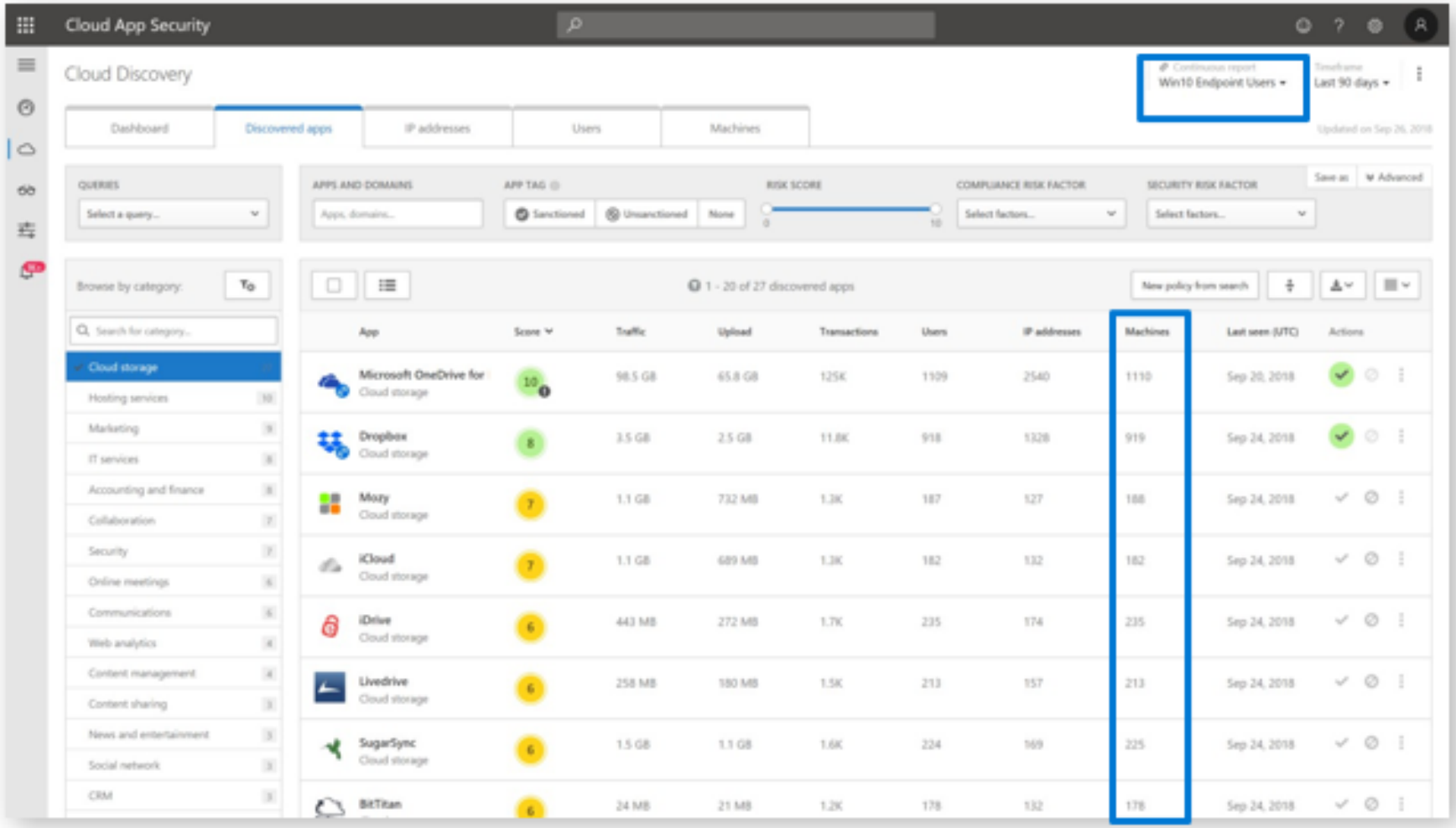
Microsoft Defender for Office 365

Securing your enterprise requires more than just prevention



zalnet
BEATA ZALEWA

Defender for Cloud Apps



zalnet
BEATA ZALEWA

Used resources

https://learn.microsoft.com/en-us/legal/cognitive-services/openai/data-privacy

https://learn.microsoft.com/en-us/azure/ai-services/content-safety/concepts/jailbreak-detection

https://learn.microsoft.com/en-us/azure/ai-services/content-safety/concepts/groundedness#groundedness-detection-features

https://learn.microsoft.com/en-us/azure/cognitive-services/openai/how-to/monitoring#monitoring-data

https://learn.microsoft.com/en-us/azure/azure-monitor/essentials/monitor-azure-resource#monitoring-data-from-azure-resources

I am actively seeking new opportunities and exciting challenges. If you would like to get in touch, please feel free to reach out through the following channels:

zalnet
BEATA ZALEWA

Email: beata@zalnet.pl

LinkedIn: https://www.linkedin.com/in/beatazalewa/

Blog: https://zalnet.pl/blog/

X: https://x.com/beatazalewa

GitHub: https://github.com/beatazalewa/Conferences/