what are these Artificial intelligence CPU Instruction sets and explain them in detail each of these entries AVX10.1 AVX10.2 AVX-512 (AVX512F) AVX-512 4x Fused Multiply-Add Single Precision (AVX512\_4FMAPS) AVX-512 4x Neural Network Instructions (AVX512\_4VNNIW) AVX-512 52-bit Integer Multiply-Add Instructions (AVX512\_IFMA) AVX-512 BF16 (AVX512\_BF16) AVX-512 Bit Algorithm (AVX512\_BITALG) AVX-512 Byte and Word Instructions (AVX512BW) AVX-512 Conflict Detection Instructions (AVX512CD) AVX-512 Doubleword and Quadword Instructions (AVX512DQ) AVX-512 Exponential and Reciprocal Instructions (AVX512ER) AVX-512 FP16 (AVX512\_FP16) AVX-512 Intersection (AVX512\_VP2INTERSECT) AVX-512 Neural Network Instructions (AVX512\_VNNI) AVX-512 Prefetch Instructions (AVX512PF) AVX-512 Vector Bit Manipulation Instructions (AVX512\_VBMI) AVX-512 Vector Bit Manipulation Instructions 2 (AVX512\_VBMI2) AVX-512 Vector Length Extensions (AVX512VL) AVX-512 VPOPCNTDQ AVX Vector Neural Network Instructions (AVX-VNNI) AVX-IFMA AVX-NE-CONVERT AVX-VNNI-INT8

**1. AVX10.1 & AVX10.2**

**Purpose**: Intel's **Advanced Vector Extensions 10 (AVX10)** aims to unify vector instruction sets across CPUs, replacing fragmented AVX-512/AVX2 implementations.

**AVX10.1**: Introduces a baseline 256-bit vector width with support for 512-bit registers in future CPUs. Includes core AVX-512 features (e.g., mask registers, embedded rounding).

**AVX10.2**: Adds newer instructions (e.g., FP16, VNNI) for AI/ML workloads.  
**Use Case**: Bridges compatibility between consumer (AVX2) and server (AVX-512) CPUs, enabling scalable vector processing.

**AVX-512 Family**

**Base Foundation**:

**AVX-512F (AVX512 Foundation)**: Mandatory base for all AVX-512 CPUs. Adds 512-bit vectors, mask registers, and expanded opcodes.

**Specialized Extensions:**

**AVX512\_4FMAPS**: Four Fused Multiply-Add (FMA) operations per cycle for single-precision (32-bit float) math.

**Use**: High-performance computing (HPC), physics simulations.

**AVX512\_4VNNIW**: Vectorizes 16-bit integer neural network inference (4 instructions).

**Use**: Accelerates deep learning inference (e.g., image recognition).

**AVX512\_IFMA**: 52-bit integer FMA for cryptography/hashing.

**Use**: RSA, big-number computations.

**AVX512\_BF16**: Brain Float 16 (bfloat16) support for AI training.

**Use**: Reduces precision loss in ML models (e.g., TensorFlow, PyTorch).

**AVX512\_BITALG**: Bit-counting and string processing.

**Use**: Genomics, data compression.

**AVX512BW**: 8/16-bit integer operations (bytes/words).

**Use**: Video encoding, text processing.

**AVX512CD**: Conflict detection for parallel scatter/gather ops.

**Use**: Sparse matrix operations, graph algorithms.

**AVX512DQ**: 32/64-bit integer & float ops (double/quadwords).

**Use**: Cryptography, financial modeling.

**AVX512ER**: Exponential/reciprocal approximations.

**Use**: Scientific simulations (now deprecated).

**AVX512\_FP16**: 16-bit floating-point support.

**Use**: AI training, HPC with reduced memory usage.

**AVX512\_VP2INTERSECT**: Finds intersecting data pairs.

**Use**: Database queries, collision detection.

**AVX512\_VNNI**: Integer ops for neural networks (8/16-bit).

**Use**: Deep learning inference (e.g., ResNet).

**AVX512PF**: Prefetch data to cache for irregular access.

**Use**: Memory-bound workloads (deprecated in consumer CPUs).

**AVX512\_VBMI**: Bit manipulation (e.g., permutations).

**Use**: Cryptography, data shuffling.

**AVX512\_VBMI2**: Expanded bit/string ops (shifts, packing).

**Use**: Compression, parsing.

**AVX512VL**: Enables AVX-512 ops on 128/256-bit vectors.

**Use**: Backward compatibility with AVX/AVX2 code.

**AVX512 VPOPCNTDQ**: Counts set bits in vectors.

**Use**: Error correction, binary ML models.

**AVX Extensions (Non-512)**

**AVX-VNNI**: AVX2-based neural network instructions (16/32-bit int).

**Use**: Accelerates inference on older CPUs (e.g., Ice Lake).

**AVX-IFMA**: 52-bit integer FMA for AVX2 (similar to AVX512\_IFMA).

**Use**: Cryptography on non-AVX-512 CPUs.

**AVX-NE-CONVERT**: Numeric format conversions (e.g., FP8/FP16).

**Use**: AI training with mixed precision.

**AVX-VNNI-INT8**: 8-bit integer neural network ops.

**Use**: Low-precision inference (e.g., edge devices).

**Applications**

**AI/ML**: BF16, VNNI, 4VNNIW, FP16.

**HPC**: 4FMAPS, ER, DQ.

**Data Processing**: BW, BITALG, VBMI2.

**Security**: IFMA, CD.

**CPU Support**

**AVX-512**: Largely restricted to Intel Xeon Scalable and some Core-X CPUs.

**AVX10**: Future Intel CPUs (e.g., Granite Rapids, Lunar Lake).

**AVX-VNNI**: Available in Intel 12th Gen+ and AMD Zen 4.

AVX10 represents Intel’s shift toward a unified vector ISA, while AVX-512 extensions remain critical for specialized workloads in servers. AMD’s Zen 4 supports AVX-512 equivalents (e.g., AVX-VNNI).