

마이크로소프트
(MSFT)

투자 의견
(팩트셋 평균)

BUY
*(1.10)

목표주가(\$)
(팩트셋 평균)

610.34

현재주가(\$)
(26.08.26)

531.01

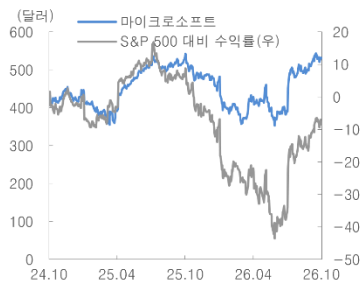
IT



Microsoft

기업명	Microsoft Corporation
한글명	마이크로소프트
홈페이지	https://www.microsoft.com
결산월	6월
거래소 국가	US
상장 거래소	NASDAQ
시가총액(십억 달러)	3,943.51
시가총액(조원)	5,266.51
발행주식수(백만주)	7,426.43
52주최고가/최저가	553.72 / 349.20
주요주주	The Vanguard, 7.48%

주가수익률(%)	1M	3M	6M	12M
절대수익률	4.3	35.3	24.2	-1.2
상대수익률	3.4	30.7	15.1	-14.9



Azure 성장 재가속,
AI 투자 효율 개선 본격화

- Azure 매출 48% 성장 AI 컴퓨팅 수요와 효율 개선 확인
- Copilot 사용 확대 기업용 AI 에이전트 상용화 가속
- Azure 성장 추가 가속 전망 CapEx와 수익성이 관건

FY 1Q27 Review: Azure 성장 가속과 현금흐름 개선

FY 1Q27 매출액은 922.8억 달러(+19% YoY)로 컨센서스 약 906.6억 달러를 1.8% 상회했다. 영업이익은 459.2억 달러(+21%), 조정 EPS는 4.89달러(+18%)로 모두 시장 예상을 웃돌았다. 특히 Azure 매출은 331.7억 달러(+48%)로 전분기 성장률 42%에서 재가속됐으며, 마이크로소프트 클라우드 매출도 639억 달러(+30%)로 성장했다. CapEx는 452억 달러로 예상보다 적었고, 영업현금흐름 580억 달러와 잉여현금흐름 186억 달러를 기록하면서 대규모 AI 투자에 따른 현금흐름 우려도 일부 완화됐다.

AI 컴퓨팅 효율 개선과 에이전트 상용화

이번 어닝콜에서 가장 중요한 변화는 AI 인프라의 매출 성장이 단순한 GPU 증설뿐 아니라 기존 설비의 효율 개선에서도 발생하고 있다는 점이다. 마이크로소프트는 기존 GPU의 처리 효율을 높이고 활용하지 못했던 CPU 자원을 매출로 연결하면서 파운드리 매출이 전년 대비 3배 이상 증가했다고 설명했다. 코파일럿(Copilot) 일간 사용자도 전분기 대비 약 50% 증가했고, 깃허브(Github)에서 AI 에이전트가 생성하는 코드 변경 요청은 전체의 3분의 1을 넘어서었다. 사티아 나델라 CEO는 개별 AI 모델보다 모델 에이전트·기업 데이터·인프라를 결합해 실제 업무를 가장 낮은 비용으로 수행하는 프런티어 생태계를 핵심 전략으로 강조했다.

AI 투자 확대 속 수익성 검증 필요

향후 관건은 이러한 효율 개선이 지속적인 매출 성장과 현금흐름으로 이어질 수 있느냐다. 마이크로소프트는 FY 2Q27 Azure 성장률을 49~50%로 제시했지만, CapEx도 500억 달러 이상으로 증가할 전망이다. 이번 분기 투자액이 예상보다 적었던 것은 일부 외부 데이터센터 용량 인도가 지연된 영향으로 투자 축소를 의미하지는 않는다. 따라서 향후 코파일럿과 기업용 에이전트의 실제 사용량 증가, 자체 MAI 모델과 Maia 반도체의 원가 절감 효과, 클라우드 수익성과 잉여현금흐름의 개선 여부를 함께 확인할 필요가 있다.

* 투자 의견: 애널리스트의 컨센서스를 구간별로(1- Strong Buy, 1.5- Buy, 2- Hold, 2.5- Sell, 3.0- Strong Sell) 수치화. 10에 근접할수록 긍정적인 투자 의견

FY 1Q27 실적 지표 (GAAP 기준)

(단위: 십억 달러, 달러 %)

구분	FY 1Q26	FY 4Q26	발표치	FY 1Q27				FY 2027
				컨센	서프	YoY	QoQ	컨센
매출액	77.7	90.0	92.3	90.6	1.8	18.8	2.5	95.6
영업이익	38.0	40.6	45.9	44.1	4.2	21.0	13.1	-
순이익	27.7	35.8	36.6	35.1	4.4	32.0	2.4	-
EPS	3.72	4.81	4.92	4.72	3.6	32.3	2.3	4.82

자료: Factset, 대신증권 Research Center

표 1. 세부 사업부문 매출액

세부 사업부문 (십억 달러)	FY 1Q26	FY 2Q26	FY 3Q26	FY 4Q26	FY 1Q27
매출액	77.7	81.3	82.9	90.0	92.3
Agents and Infra	33.0	34.1	35.0	37.8	76.8
Devices and Consumer	30.9	32.9	34.7	39.3	15.5

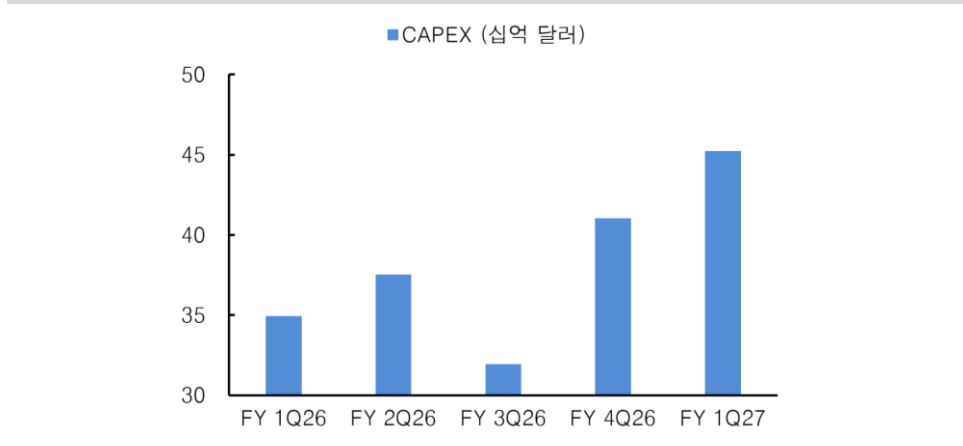
자료: Microsoft, 대신증권 Research Center

표 2. 세부 사업부문 전년동기대비 성장률

세부 사업부문 (% YoY)	FY 1Q26	FY 2Q26	FY 3Q26	FY 4Q26	FY 1Q27
매출액	18.4	16.7	18.3	17.7	18.8
Agents and Infra	16.6	15.9	16.9	14.3	24.5
Devices and Consumer	28.2	28.8	29.6	31.6	(3.2)

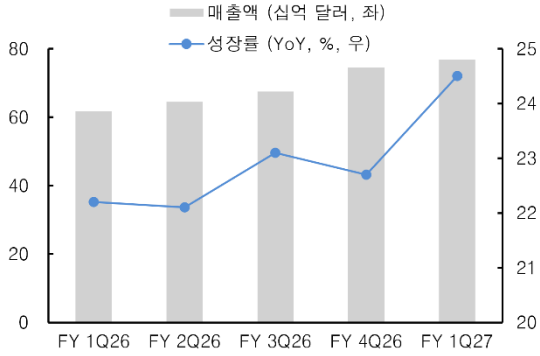
자료: Microsoft, 대신증권 Research Center

그림 1. 분기별 CAPEX



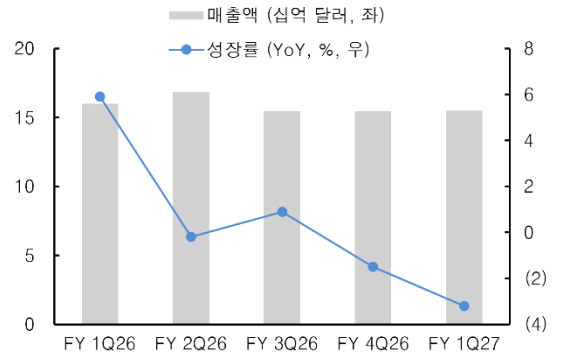
자료: Microsoft, 대신증권 Research Center

그림 2. Agents and Infra 매출액 및 성장률



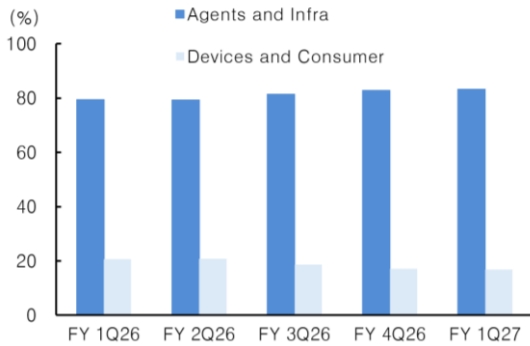
자료: Microsoft, 대신증권 Research Center

그림 3. Devices and Consumer 매출액 및 성장률



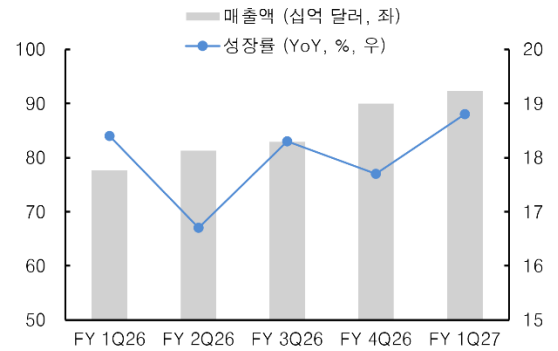
자료: Microsoft, 대신증권 Research Center

그림 4. 주요 사업부문별 매출액 비중



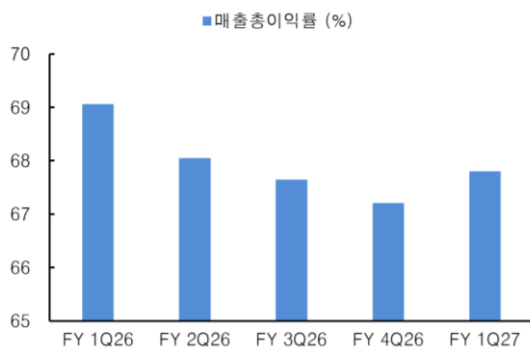
자료: Microsoft, 대신증권 Research Center

그림 5. 분기별 매출액 및 성장률



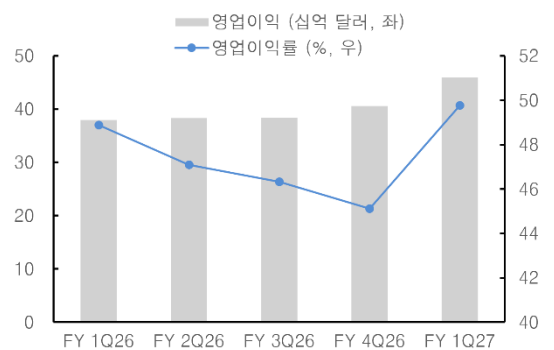
자료: Microsoft, 대신증권 Research Center

그림 6. 매출총이익률



자료: Microsoft, 대신증권 Research Center

그림 7. 영업이익 및 영업이익률



자료: Microsoft, 대신증권 Research Center

표 3. 분기 실적 추이

구분 (십억 달러)	FY 1Q25	FY 2Q25	FY 3Q25	FY 4Q25	FY 1Q26	FY 2Q26	FY 3Q26	FY 4Q26	FY 1Q27
매출액	65.6	69.6	70.1	76.4	77.7	81.3	82.9	90.0	92.3
Agents and Infra	28.3	29.4	29.9	33.1	33.0	34.1	35.0	37.8	76.8
Devices and Consumer	24.1	25.5	26.8	29.9	30.9	32.9	34.7	39.3	15.5
성장률 (전년대비, %)									
매출액	16.0	12.3	13.3	18.1	18.4	16.7	18.3	17.7	18.8
Agents and Infra	12.3	13.9	10.4	15.7	16.6	15.9	16.9	14.3	24.5
Devices and Consumer	20.4	18.7	20.8	25.6	28.2	28.8	29.6	31.6	(3.2)
비중 (%)									
Agents and Infra	43.2	42.3	42.7	43.3	42.5	42.0	42.2	42.0	83.2
Devices and Consumer	36.7	36.7	38.2	39.1	39.8	40.5	41.8	43.7	16.8
매출원가	20.1	21.8	21.9	24.0	24.0	26.0	26.8	29.5	29.7
매출총이익	45.5	47.8	48.1	52.4	53.6	55.3	56.1	60.5	62.6
매출총이익률 (%)	69.4	68.7	68.7	68.6	69.0	68.0	67.6	67.2	67.8
영업비용	14.9	16.2	16.1	18.1	15.7	17.0	17.7	19.9	16.6
연구개발비	7.5	7.9	8.2	8.8	8.1	8.5	8.9	10.0	9.0
판매비 및 일반관리비	7.4	8.3	7.9	9.3	7.5	8.5	8.7	9.9	7.6
영업이익	30.6	31.7	32.0	34.3	38.0	38.3	38.4	40.6	45.9
영업이익률 (%)	46.6	45.5	45.7	44.9	48.9	47.1	46.3	45.1	49.8
당기순이익	24.7	24.1	25.8	27.2	27.7	38.5	31.8	35.8	36.6
순이익률 (%)	37.6	34.6	36.9	35.6	35.7	47.3	38.3	39.7	39.7
EPS (GAAP, 달러)	3.30	3.23	3.46	3.65	3.72	5.16	4.27	4.81	4.9
성장률 (% YoY)	10.4	10.2	17.7	23.7	12.7	59.8	23.4	31.8	32.3

자료: Microsoft, 대신증권 Research Center

Azure 성장 가속, AI 투자 효율 개선과 수익성 검증

마이크로소프트의 FY27 1Q(2026년 7~9월) 매출액은 922.8억 달러(+19% YoY), 영업이익은 459.2억 달러(+21%)로 모두 시장 예상을 상회했다. 영업이익률은 49.8%로 전년 동기 48.9%에서 상승했다. 순이익은 366.3억 달러(+32%), GAAP EPS는 4.92달러(+32%)였으며, 오픈AI(OpenAI) 투자에 따른 회계상 영향을 제외한 조정 순이익과 EPS는 각각 364.2억 달러(+18%), 4.89달러(+18%)를 기록했다. 매출총이익률은 67.8%로 소폭 하락했지만, 매출 확대와 운영 효율 개선으로 영업이익률은 오히려 상승했다.

이번 실적에서 가장 중요한 것은 애저(Azure)의 성장 가속과 AI 컴퓨팅 효율 개선이다. 애저 매출은 331.7억 달러(+48%)로 전분기 성장률 42%를 크게 웃돌았으며, 마이크로소프트 클라우드 전체 매출도 639억 달러(+30%)로 증가했다. 특히 이번 성장은 신규 데이터센터와 GPU 증설뿐 아니라 기존 인프라의 효율 개선에서도 발생했다. 마이크로소프트는 기존 GPU가 같은 시간에 더 많은 AI 연산을 수행하도록 최적화하고, 충분히 사용되지 않던 CPU 자원도 추가 매출로 연결했다. 여기에 고성능 AI 모델의 사용 확대와 토큰당 매출 증가가 더해지면서 파운드리(Foundry) 매출은 전년 대비 3배 이상 성장했다. 즉 AI 인프라 투자를 계속 확대하는 동시에 기존 설비로도 더 많은 매출을 창출하기 시작했다는 점이 긍정적이다.

이번 분기부터 사업부 공시도 기존 3개에서 2개로 개편됐다. 애저와 마이크로소프트 365, 기업용 에이전트 등을 포함하는 에이전트·인프라(Agents and Infra) 매출은 768억 달러(+25%)로 전체의 약 83%를 차지한 반면, 디바이스·소비자(Devices and Consumer) 매출은 155억 달러(-3%)로 감소했다. AI 인프라와 기업용 소프트웨어 중심으로 사업 구조가 더욱 뚜렷해지고 있는 것이다.

FY27 2Q 매출 가이드는 960억~971억 달러(+18~19%)이며, 애저 성장률은 49~50%로 추가 가속될 전망이다. 다만 AI 컴퓨팅 수요가 여전히 공급을 초과하고 있어 실제 성장 속도는 신규 데이터센터의 가동 시점과 기존 설비의 효율 개선에 영향을 받을 것으로 예상된다.

수익성 측면에서는 막대한 AI 투자 부담이 지속될 전망이다. 이번 분기 CapEx는 452억 달러로 예상보다 적었지만 이는 일부 외부 데이터센터 용량 인도가 지연된 영향이며, FY 2Q27에는 500억 달러를 넘어설 것으로 예상된다. 마이크로소프트 클라우드 매출총이익률도 이번 분기 66%에서 다음 분기 약 65%로 낮아질 전망이다. 다만 영업현금흐름 580억 달러와 잉여현금흐름 186억 달러를 기록했고, FY27 연간 잉여현금흐름도 흑자를 유지할 것으로 예상된다. 결국 향후 핵심은 AI 사용량 증가 자체보다 동일한 투자비와 전력으로 얼마나 많은 업무를 처리하고 실제 매출로 연결할 수 있는냐다. 코파일럿과 기업용 에이전트의 상용화, 자체 MAI 모델과 Maia 반도체의 완가 절감, 기업 데이터와 여러 AI 모델을 결합한 프런티어 생태계 구축이 장기적인 성장성과 수익성을 결정할 것으로 판단한다.

CapEx 452억 달러, 현금흐름은 예상보다 양호

이번 분기 CapEx는 금융리스를 포함해 452억 달러로 전년 동기 349억 달러 대비 약 30% 증가했다. 직전 분기 410억 달러와 비교해도 증가했지만, 회사가 예상했던 약 500억 달러보다는 적었다. 다만 이는 투자 계획을 축소했기 때문이 아니라 외부 데이터센터 사업자로부터 공급받을 예정이었던 일부 컴퓨팅 용량의 인도가 지연됐기 때문이다. 실제 CapEx의 약 75%는 GPU와 CPU, 메모리, 스토리지 등 상대적으로 교체 주기가 짧은 장비에 투입됐다. 나머지는 데이터센터 건물과 장기 인프라에 사용됐다. 메모리와 스토리지 가격 상승 역시 투자비 증가 요인으로 언급됐다.

반면 영업현금흐름은 580억 달러로 전년 대비 29% 증가했다. 실제 유형자산 취득을 위한 현금지출은 393.7억 달러였으며, 이를 차감한 잉여현금흐름은 186.3억 달러를 기록했다. 클라우드 고객의 청구와 대금 회수가 늘어나면서 현금창출력이 개선됐고, 시장이 우려했던 대규모 잉여현금흐름 적자는 발생하지 않았다. 다만 현금 및 단기투자자산은 전분기 말 768억 달러에서 645억 달러로 감소했다.

향후 투자 부담은 여전히 크다. 마이크로소프트는 FY 2Q27 CapEx가 500억 달러를 넘어설 것으로 예상하고 있으며, 2026년 연간 CapEx는 약 1,700억 달러를 전망했다. 이전보다 투자 시점이 일부 뒤로 밀렸지만 FY27 하반기에는 지연된 외부 컴퓨팅 용량 인도와 GPU·CPU 도입이 겹치면서 투자액이 다시 확대될 것으로 예상된다. 회사는 FY27 전체로는 잉여현금흐름 흑자를 유지할 것으로 전망하면서도 필요할 경우 리스와 채권 발행 등 추가 자금조달 수단을 활용할 수 있다고 설명했다.

그림 8. 마이크로소프트의 AI 투자 및 현금흐름 현황 (단위: 억 달러)

구분	FY 1Q27 (2026년 7~9월)	주요 내용
CapEx (금융리스 포함)	452 (+29.5% YoY, +10.2% QoQ)	• 전년 동기 대비 약 30% 증가, 직전 분기 대비 증가 (외부 데이터센터 인도 지연 영향)
단기 장비 투자 비중	약 75%	• GPU·CPU·메모리·스토리지 등 교체 주기가 짧은 장비 중심
장기 인프라 투자 비중	약 25%	• 데이터센터 건물, 전력·냉각 인프라 등 장기 자산 투자
영업현금흐름 (OCF)	580 (+29% YoY)	• 클라우드 고객 청구와 대금 회수 증가로 현금창출력 개선
현금 기준 CapEx	393.7	• 실제 유형자산 취득을 위한 현금지출 규모
잉여현금흐름 (FCF)	186.3	• 영업현금흐름 580억 달러에서 현금 기준 CapEx 393.7억 달러를 차감
현금 및 단기투자자산	645 (QoQ -123)	• 전분기 말 768억 달러 대비 123억 달러 감소

향후 CapEx 및 현금흐름 전망

구분	전망치	주요 내용
FY 2Q27 CapEx	500억 달러 초과	• 외부 데이터센터 인도 지연분과 추가 GPU·CPU 도입으로 500억 달러 상회 전망
2026년 연간 CapEx	약 1,700억 달러	• 기존 가이드선스 유지, 대규모 AI 인프라 투자 지속
FY27 연간 FCF	흑자 전망	• 투자 확대에도 연간 잉여현금흐름 흑자 유지 예상

자료: Microsoft, 대신증권 Research Center

프런티어 생태계, 모델보다 비용 대비 성과 경쟁

사티아 나델라 CEO는 FY26 연례 주주서한과 이번 어닝콜에서 향후 AI 시대의 핵심 경쟁력이 개별 AI 모델이 아니라 모델, AI 에이전트, 기업 데이터, 반도체와 데이터센터를 결합한 프런티어 생태계(Frontier Ecosystem)에 있다고 강조했다. 여러 기업의 AI 모델이 빠르게 발전하면서 기업들은 하나의 모델만 사용할 필요가 없어지고 있다. 따라서 마이크로소프트는 가장 똑똑한 모델 하나를 보유하는 것보다 고객이 여러 모델을 자유롭게 사용하고 실제 업무를 가장 낮은 비용으로 수행할 수 있는 시스템을 제공하는 데 집중하고 있다.

특히 나델라는 AI 경제성의 핵심 지표로 비용 대비 성과(Cost-to-Outcome)를 강조했다. 기존 챗봇은 사용자가 질문할 때마다 한 번씩 응답하지만, AI 에이전트는 검색, 데이터 조회, 프로그램 실행, 다른 에이전트 호출 등을 반복하면서 수십 분에서 수시간 동안 업무를 수행한다. 이 과정에서 토큰 사용량은 크게 증가한다. 그러나 기업 고객이 원하는 것은 더 많은 토큰이 아니라 실제 업무 결과이기 때문에, 같은 업무를 더 적은 토큰으로 수행하거나 동일한 전력에서 더 많은 토큰을 생성하는 기술이 중요해진다.

이를 위해 마이크로소프트는 여러 데이터센터에서 발생하는 컴퓨팅 수요를 하나의 공통 자원처럼 관리하는 방향으로 인프라를 설계하고 있다. 애저, 파운드리, 코파일럿 등 여러 서비스가 사용하는 컴퓨팅을 수요와 수익성에 따라 유연하게 배분하는 구조다. 예를 들어 코파일럿의 모델 최적화로 기존 GPU 사용량이 줄어들면 남은 용량을 파운드리 고객에게 제공해 추가 매출을 얻을 수 있다. 이번 분기에 확인된 파운드리 매출 확대와 코파일럿 처리량 증가는 이러한 전략이 실제 성과로 연결되기 시작했다는 의미가 있다.

코파일럿·에이전트, 기업 업무 자동화 중심으로 전환

마이크로소프트는 최근 코파일럿(Copilot) 전략의 중심을 범용 소비자 AI에서 기업의 실제 업무를 수행하는 AI 에이전트로 이동하고 있다. 소비자용 코파일럿 자체를 포기한다기보다 소비자용 AI 비서 경쟁의 우선순위를 낮추고, 마이크로소프트 365(Microsoft 365), 팀즈(Teams), 아웃룩(Outlook), 깃허브(GitHub), 다이나믹스(Dynamics), 애저와 연결된 업무용 AI에 집중하는 전략이다.

새로운 코파일럿은 채팅, 장기 업무를 수행하는 코워크(Cowork), 자연어 기반 소프트웨어 개발 기능인 코드(Code), 사용자가 자리를 비워도 계속 업무를 수행하는 오토파일럿(Autopilot)을 하나의 애플리케이션으로 통합했다. 기존 워드·엑셀·파워포인트를 사용자가 직접 오가는 방식에서 코파일럿에 업무를 맡기면 필요한 프로그램과 전문 에이전트를 자동으로 활용하는 업무용 운영체제를 지향한다.

이번 어닝콜에서는 이러한 변화가 실제 사용자 활동으로 이어지고 있다는 점이 확인됐다. 코파일럿 일간 사용자는 전분기 대비 약 50%, 전년 대비 5배 이상 증가했으며, 자동으로 수행한 작업 수는 7월 이후 약 2배 증가했다. 사용자 만족도도 전분기 대비 약 20% 개선됐다. 전체 코파일럿 요청의 90%는 자동 모드를 통해 처리되는데, 이는 사용자가 직접 모델을 선택하지 않아도 시스템이 작업의 난이도와 비용, 응답속도에 따라 적합한 AI 모델을 자동으로 배정하는 방식이다.

수익모델 역시 사용자당 구독료에서 구독료와 사용량 기반 과금을 함께 적용하는 방향으로 확대되고 있다. 기본 코파일럿은 기업 IT 부서가 직원 전체를 대상으로 구독료를 지불하지만, 복잡한 업무를 장시간 수행하는 에이전트는 사용한 만큼 비용을 지불하는 구조다. 이번 분기 사용량 기반 크레딧을 실제로 소비한 코파일럿 고객 수는 전분기 대비 75% 이상 증가했다. 다만 나델라는 기업들이 단순히 토큰 사용량을 늘리는 것을 원하지 않으며, 구독료로 제공되는 기본 기능을 충분히 활용하면서 추가적인 업무 가치가 발생할 때 사용량 과금을 적용하는 것이 중요하다고 설명했다.

기업 데이터·멀티모델, 실제 업무 기록이 새로운 학습 자산으로

마이크로소프트는 개별 AI 모델보다 기업이 축적하는 데이터와 업무 맥락을 더욱 중요한 장기 자산으로 보고 있다. 현재 마이크로소프트 파운드리(Microsoft Foundry)는 OpenAI, 엔트로픽, 미스트랄(Mistral) 등 외부 모델과 자체 MAI 제품군을 포함해 1만1,000개 이상의 모델을 제공한다. 이번 분기 파운드리 고객은 12만 곳을 넘어섰고 매출은 전년 대비 3배 이상 증가했다. 여러 모델을 활용해 에이전트를 구축하는 고객 수는 전년 대비 5배 증가했으며, 파운드리에서 제공하는 오픈소스 모델 관련 매출은 10배 늘어났다. 이는 기업들이 하나의 모델에 모든 업무를 맡기기보다 업무별 성능과 비용에 따라 여러 모델을 선택하기 시작했음을 보여준다.

이를 위해 마이크로소프트는 AI 모델과 AI가 실제 업무를 수행하도록 만드는 하네스(Harness)를 분리하고 있다. 하네스는 AI 모델이 사용할 수 있는 기업 데이터, 장기 기억, 도구, 행동 규칙, 실행 환경을 연결하고 작업 진행을 관리하는 시스템이다. 예를 들어 OpenAI 모델을 사용해 구축한 기업용 에이전트라도 모델을 엔트로픽이나 MAI로 바꿀 때 기존 기업 데이터와 업무 흐름을 처음부터 다시 구축할 필요가 없도록 설계하는 것이다. 최근에는 깃허브에서 개발한 멀티모델 하네스를 코파일럿 전체에 확대 적용하고, 이를 파운드리 고객에게도 제공하기 시작했다. 이 하네스는 20개 이상의 프런티어 모델을 지원하며, 실행 엔진을 러스트(Rust) 기반으로 전환해 작업 시작과 첫 토큰 생성에 필요한 시간을 줄였다.

여기서 더 중요한 변화는 실제 업무 수행 기록이 향후 AI 모델을 추가 학습시키는 데이터가 된다는 점이다. 나델라는 이번 어닝콜에서 유니레버(Unilever)가 프런티어 모델을 활용해 주문 접수와 거래 조건 관련 분쟁 처리를 자동화한 뒤, 그 과정에서 축적된 실행 기록을 사용해 자신들의 MAI 모델을 학습시킨 사례를 소개했다. 실행 기록에는 AI가 어떤 정보를 검색했고, 어떤 도구를 호출했으며, 어떤 순서로 업무를 처리했는지가 남는다. 기업은 이 기록을 활용해 범용 프런티어 모델이 수행하던 특정 업무를 더 작고 저렴한 자체 모델로 이전할 수 있다.

즉 처음에는 가장 강한 프런티어 모델을 사용해 업무를 자동화하지만, 작업이 반복되면서 충분한 데이터가 축적되면 해당 기업 업무에 특화된 모델을 학습시키는 구조다. 이후에도 새로운 업무나 어려운 문제에는 다시 프런티어 모델을 사용하고, 반복적인 작업은 저렴한 모델로 이전할 수 있다. 이는 AI 사용 자체가 다음 AI를 위한 학습 데이터를 만드는 순환 구조라는 점에서 중요하다.

Microsoft IQ와 Agent 365, 기업 AI의 지식·통제 계층 구축

이러한 기업 고유의 지식과 업무 맥락을 연결하는 것이 마이크로소프트 아이큐(Microsoft IQ)다. 워크 아이큐(Work IQ)는 직원의 이메일, 문서, 회의, 채팅, 프로젝트 정보 등을 연결해 누가 어떤 일을 하고 어떤 관계 속에서 협업하는지 이해하도록 한다. 패브릭 아이큐(Fabric IQ)는 ERP, CRM, 재고, 생산, 회계 등 기업의 실제 운영 데이터를 연결하며, 파운드리 아이큐(Foundry IQ)는 기업 규정과 공식 문서, 지식 데이터베이스 등 에이전트가 참고해야 할 정보를 제공한다. 웹 아이큐(Web IQ)는 외부 웹에서 새로운 정보를 가져오는 역할을 한다. 이들을 결합하면 AI가 단순히 문서를 검색하는 수준을 넘어 기업의 조직과 업무 관계, 현재 사업 상태를 종합적으로 이해할 수 있게 된다.

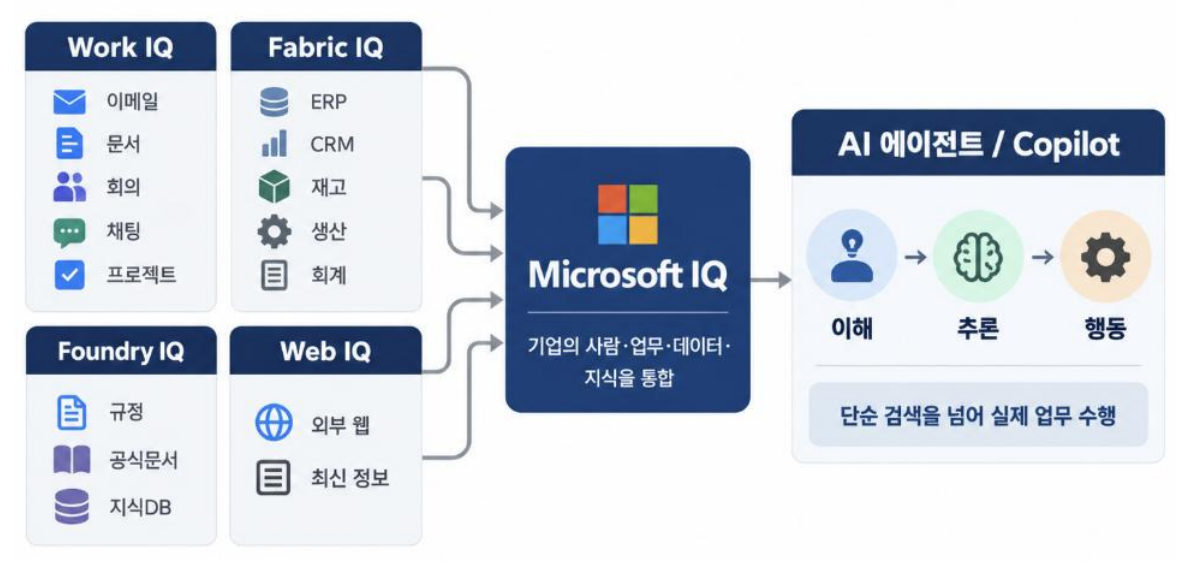
특히 패브릭 아이큐는 기업의 설비, 제품, 고객, 주문, 재고 등을 객체로 정의하고 이들 간의 관계와 규칙을 연결한다는 점에서 팔란티어의 온톨로지(Ontology)와 유사한 방향성을 가진다. 예를 들어 특정 제품의 납품이 지연되면 해당 제품에 필요한 부품, 공급업체, 생산시설, 재고 현황과 고객 계약을 함께 분석해 필요한 조치를 제안하는 방식이다. 여기에 워크 아이큐가 보유한 이메일과 회의 정보를 더하면 시스템에 기록된 공식 데이터뿐 아니라 담당자들이 실제로 어떤 문제를 논의하고 있는지도 함께 이해할 수 있다. FY26 기준 워크 아이큐가 연결하는 데이터는 17EB를 넘어섰으며, 이번 분기 마이크로소프트 아이큐를 통해 기업 데이터와 업무 맥락을 활용하는 고객도 약 6만 곳으로 전분기 대비 25% 증가했다.

여기에 에이전트 365(Agent 365)는 기업이 직원의 계정과 권한을 관리하듯 AI 에이전트의 신원, 보안, 접근권한, 비용과 성과를 관리하는 통제 계층을 제공한다. 등록된 에이전트 수는 이번 분기 5,000만 개를 넘어 전분기 대비 30% 이상 증가했다. 마이크로소프트는 AI 모델의 자능과 시스템 접근권한을 분리하고, 필요한 작업에 대해서만 권한을 부여하는 방식을 강조한다. 이를 위해 퍼셉션(Perception)은 에이전트의 위험한 행동을 감시하고 차단하며, 여러 감시 에이전트가 다른 에이전트의 행동을 계속 검토하도록 설계됐다. 기업 고객이 퍼뷰(Purview)를 통해 감사한 에이전트 활동도 누적 76억 건을 넘어 전분기 대비 5배 증가했다.

마이크로소프트는 이를 실제 고객 현장에 적용하기 위해 약 25억 달러를 투자하고 6,000명의 산업·AI 엔지니어를 고객사에 투입하는 프런티어 컴퍼니(Frontier Company)도 운영하고 있다. 기업의 업무를 직접 이해하고 데이터와 AI 에이전트를 연결하면서 사용 결과를 반복적으로 개선하는 역할이다. 특히 마이크로소프트는 여러 고객이 기업 데이터의 소유권과 학습 결과를 외부 AI 모델 제공업체에 종속시키지 않으면서 자신만의 학습 시스템을 만들 수 있어야 한다고 강조한다.

그림 9. 기업 데이터·문맥·관계를 지배하는 Microsoft IQ

기업 데이터와 업무 맥락을 연결해 AI 에이전트가 실제 일을 이해하고 수행하도록 하는 통합 계층



MAI-Base-2, 프런티어급 자체 모델과 저비용 특화 모델 확대

마이크로소프트는 10월 15일 자체 파운데이션 모델 MAI-Base-2의 사전학습을 완료했다고 발표했다. MAI-Base-2는 총 2.8조 개의 파라미터를 보유하지만 입력에 따라 필요한 전문가 모델만 선택하는 전문가 혼합 구조(Mixture of Experts)를 적용해 실제 연산에는 약 800억 개의 파라미터만 활성화한다. 2026년 8~9월 엔비디아 GB200 GPU 1만6,000개를 이용해 38일 동안 총 50조 개의 토큰을 학습했다. 이미지와 텍스트를 처음부터 함께 학습하는 멀티모달 구조를 적용했으며, 전체 학습량에서 이미지·텍스트 데이터가 차지하는 비중은 18%였다.

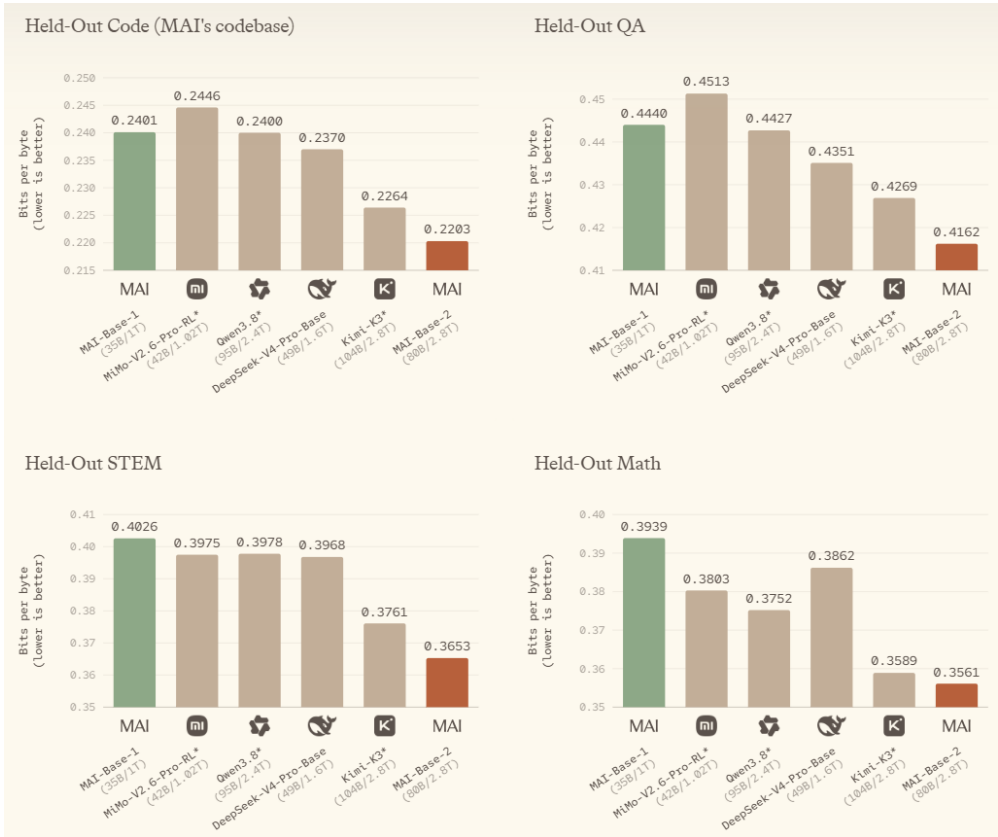
이번 모델은 대규모 모델을 얼마나 적은 연산으로 학습할 수 있는지에서도 개선을 보였다. 마이크로소프트 자체 평가에서 MAI-Base-2는 테스트한 모든 오픈웨이트 모델보다 높은 사전학습 성능을 기록했으며, 딥시크(DeepSeek)의 V4-Pro-Base와 동일한 성능에 도달하는 데 필요한 학습 연산량이 4.2~12.3배 적었다. MAI-Base-1과 비교해도 학습 효율이 6.3~12.7배 개선됐다. 이를 위해 학습 데이터의 품질을 높이고, 중복되거나 가치가 낮은 데이터를 줄였으며, 모델 구조와 학습 설정을 함께 최적화했다. 특히 다른 AI 모델의 출력물을 증류하는 방식에 의존하기보다 자체적으로 수집·정제한 데이터 파이프라인을 구축하고 학습에 활용했다.

무스타파 술레이만 마이크로소프트 AI CEO는 사내 타운홀에서 MAI-Base-2가 GPT-5.6 Sol의 기반 모델로 알려진 OpenAI의 Spud보다 사전학습 평가에서 높은 성능을 기록했으며, 학습에 사용한 컴퓨팅은 약 25%에 불과했다고 주장했다. 그리고 현재 MAI-Base-2도 실제 지시 수행과 추론 능력을 높이기 위한 사후학습을 진행하고 있으며, 이미 MAI-Base-3 또한 개발 준비를 시작했음을 밝혔다.

이번 어닝콜에서는 자체 모델 개발의 목표를 더욱 분명히 설명했다. 나델라는 최고 수준의 외부 프런티어 모델이 출시된 이후 3~6개월 안에 유사한 성능을 훨씬 적은 학습 비용으로 제공하는 것을 목표로 제시했다. 즉 프런티어 모델 기술 자체의 경쟁에는 참여하되, 반드시 최고 성능 모델 하나를 보유하는 것보다 해당 성능을 얼마나 낮은 비용으로 실제 서비스에 적용할 수 있는냐를 중요하게 보고 있다. 이는 모든 작업에 가장 크고 비싼 프런티어 모델이 필요하지 않다는 마이크로소프트의 철학을 보여준다. 간단한 분류나 행동 선택은 작은 모델이 수행하고, 복잡한 추론만 프런티어 모델에 맡기면 전체 토큰 비용을 크게 낮출 수 있다는 전략이다.

결국 MAI 개발의 핵심 목적은 오픈AI나 엔트로픽 같은 외부 모델을 완전히 대체하는 것보다 자체적으로도 프런티어급 성능을 확보하고, 반복적으로 사용하는 기업용 AI 업무를 저렴한 특화 모델로 이전할 수 있는 선택권을 갖는 데 있다. 외부 프런티어 모델은 계속 활용하되 MAI 모델과 자체 반도체를 함께 최적화하면서 코파일럿과 기업용 에이전트의 원가와 제품 개발 속도를 직접 통제하려는 전략이다.

그림 10. MAI-Base-2 모델 성능 벤치마크



자료: Microsoft, 대산증권 Research Center

Maia 자체 칩과 AI 슈퍼팩토리, 메모리·전력 효율 개선

마이크로소프트는 최근 차세대 자체 AI 가속기 Maia 300을 공개하고 2027년부터 애저 내 Maia 시스템을 수 GW 규모로 확대할 계획이다. 기존 Maia가 일부 서비스에서 성능과 경제성을 검증하는 단계였다면, 앞으로는 코파일럿과 기업용 에이전트의 대규모 추론을 담당하는 주력 시스템으로 역할이 확대된다. 다만 엔비디아와 AMD의 GPU를 전면 대체하는 것이 아니라 대규모 학습과 외부 고객 워크로드에는 최신 GPU를 계속 사용하면서 반복적으로 많은 토큰을 생성하는 자체 서비스의 추론에서는 Maia 비중을 높이는 방향이다. 이번 어닝콜에서도 나델라는 엔비디아의 첫 Vera Rubin GPU를 반입했다고 밝히며 외부 GPU와 자체 칩을 함께 확대하는 전략을 재확인했다.

Maia 300은 TSMC의 2나노급 N2P 공정과 칩렛 구조를 활용하고 HBM4 8개를 탑재해 칩당 메모리 대역폭이 최대 18TB/s에 달한다. 한 랙에는 Maia 300 96개가 탑재되며, 총 27.6TB의 HBM과 1.7PB/s의 메모리 대역폭을 제공한다. 마이크로소프트는 Maia 200 대비 최대 7배 높은 랙당 토큰 처리량과 주요 GPU 시스템 대비 같은 전력에서 약 40% 더 많은 토큰 처리량을 목표로 하고 있다.

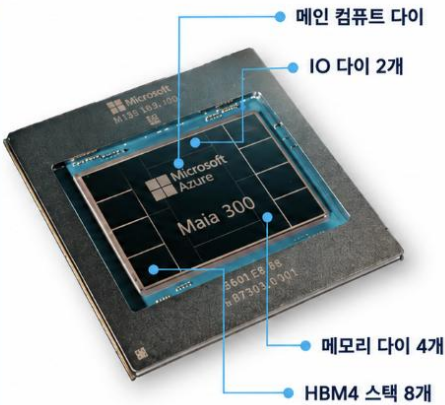
특히 Maia 300은 단순한 연산 성능보다 데이터를 얼마나 적게 이동시키면서 추론을 수행할 수 있는지에 초점을 맞췄다. AI 에이전트가 긴 업무를 수행하면 이전 대화와 중간 작업 결과를 저장하는 KV 캐시가 계속 커지고, 이를 HBM에서 읽고 쓰는 과정에서 많은 메모리 대역폭과 전력이 소모된다. Maia 300은 저장밀 연산인 MXFP4를 활용해 모델 가중치와 KV 캐시가 차지하는 공간을 줄이고, 칩 내부의 고속 메모리인 SRAM을 적극적으로 활용해 HBM과 연산장치 사이의 데이터 이동을 줄인다. 네트워크도 이터넷 기반의 고속 연결 구조로 설계해 여러 칩과 랙이 하나의 시스템처럼 작동하도록 했다. 결국 같은 메모리와 전력으로 더 긴 문맥을 처리하고 더 많은 에이전트를 동시에 실행하려는 구조다.

데이터센터 자체도 하나의 대규모 학습 시스템으로 연결하고 있다. FY26 신규 데이터센터 88개를 추가한 데 이어 이번 분기에도 5개 대륙에 24개 데이터센터를 추가했다. 특히 미국 애틀랜타와 위스콘신의 페어워터(Fairwater) 데이터센터를 고속 광통신망으로 연결해 서로 다른 지역의 연산 자원을 하나의 AI 슈퍼팩토리처럼 활용하는 구조를 구축했다. 대규모 모델의 사전학습뿐 아니라 강화학습과 평가처럼 여러 단계의 작업을 반복하는 과정에서 GPU의 유휴시간과 데이터 이동 병목을 줄이는 것이 목적이다.

실제 이번 분기에는 주요 지역에서 엔비디아 GB300의 반입부터 서비스 가동까지 걸리는 시간을 전분기 대비 약 20% 단축했고, 코파일럿 워크로드의 처리량도 전분기 대비 50% 이상 높였다. 이는 새로운 GPU를 추가하는 것뿐 아니라 기존 장비를 더 빠르게 가동하고 소프트웨어를 최적화하는 것만으로도 상당한 추가 처리량을 확보할 수 있다는 의미다. 결국 MAI 자체 모델, 코파일럿과 기업용 에이전트, Maia 자체 칩, 애저 데이터센터를 함께 설계해 동일한 전력과 투자비로 더 많은 AI 업무를 수행하는 것이 마이크로소프트의 장기적인 수직 통합 전략이다.

그림 11. Maia 200 & 300 칩 스펙

주요 사양	Maia 200	Maia 300
공정	모놀리식 N3P	멀티 칩렛 N2P 컴퓨트
FP4 연산 성능	10.1	13.8
FP8 연산 성능	5.1	6.9
전력	750 W	950 W
HBM	6 x HBM3E 스택	8 x HBM4 스택
HBM 대역폭	7.1 TB/s	18 TB/s
HBM 용량	216 GB	288 GB
백엔드 네트워크	2단계 이더넷 스케일업 + AI Transport	2단계 이더넷 스케일업 + AI Transport
스케일업 대역폭(양방향)	2.8 TB/s	3.2 TB/s
랙당 탑재 수	48개 마이아스/랙	96개 마이아스/랙
클러스터 규모	~6천 개	~1.2만 개 가속기



자료: Microsoft, 대신증권 Research Center

[Compliance Notice]

금융투자업규정 4-20조 1항5호사목에 따라 작성일 현재 사전고지와 관련한 사항이 없으며, 당사의 금융투자분석사는 자료작성일 현재 본 자료에 관련하여 재산적 이해관계가 없습니다. 당사는 동 자료에 언급된 종목과 계열회사의 관계가 없으며, 당사의 금융투자분석사는 본 자료의 작성과 관련하여 외부 부당한 압력이나 간섭을 받지 않고 본인의 의견을 정확하게 반영하였습니다.

(담당자: 리서치센터)

본 자료는 투자자들의 투자판단에 참고가 되는 정보제공을 목적으로 배포되는 자료입니다. 본 자료에 수록된 내용은 당사 리서치센터의 추정치로서 오차가 발생할 수 있으며, 정확성이나 완벽성은 보장하지 않습니다. 본 자료를 이용하시는 분은 동 자료와 관련한 투자의 최종 결정은 자신의 판단으로 하시기 바랍니다.