

A 30 éves Recognita története

A Neumann János Számítógép-tudományi Társaság Informatikatörténeti Fóruma (NJSZT iTF) és az Óbudai Egyetem (ÓE) a Recognita 30 évéről tartott rendezvényt.

A rendezvény időpontja: **2019. november 08. (péntek) 14:00 – 18:00**

Helyszín: **Óbudai Egyetem** (Budapest III. Bécsi út 96/B) F01 terem

A program:

I. Fejlődésünk története – Milyen utat jártunk be és hogyan, hova jutottunk, honnan indulhat tovább az új generáció?

Tick József: Megnyitó ▶

Reszler Ákos: Bevezető ▶ 

Várszegi György: A Recognita fejlődési stációi ▶ 

Novák Attila: A stabilitás fenntartása a kihívások közepette, avagy miből lesz a cserebogár ▶ 

Berliner Péter: Generációk jönnek-mennek, de a Recognita örök – a Recognita értékek friss szemmel ▶ 

II. Szakmai szekció – Mit főzünk a fazékban?

Szalay Imre: Köszöntés az NJSZT nevében, emléklap átadása Kovács Emőkének ▶

Marosi István: A karakterfelismerés buktatói ▶ 

Tállai Benedek: Hozzászólás ▶

Urbán Zoltán: Mi az a PDF és miért szeretjük ▶ 

Reszler Ákos: Összefoglalás, zársz ▶

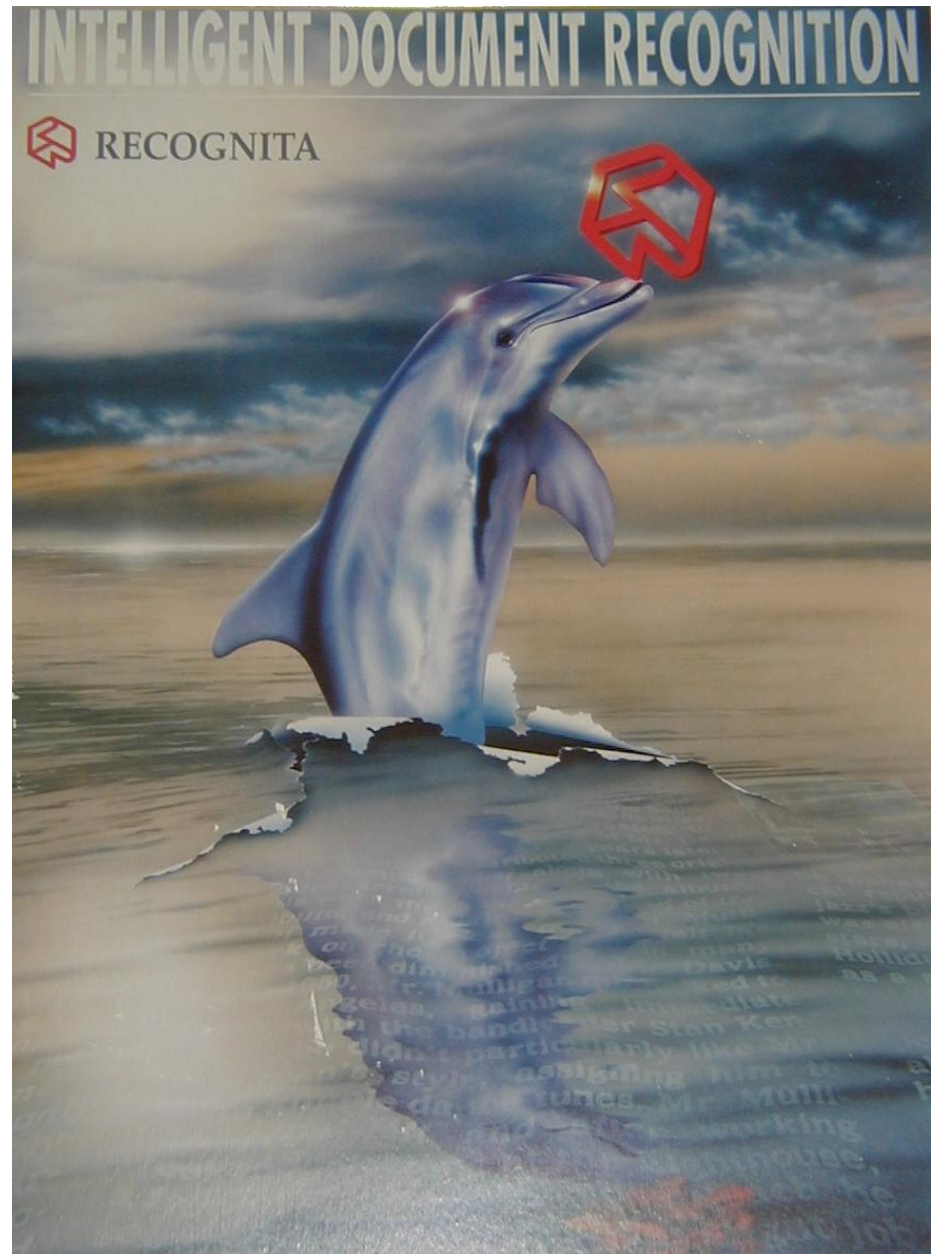


A 30 éves Recognita múlt, jelen és jövő

Dr. Reszler Ákos
vezérigazgató
Kofax-Recognita Zrt

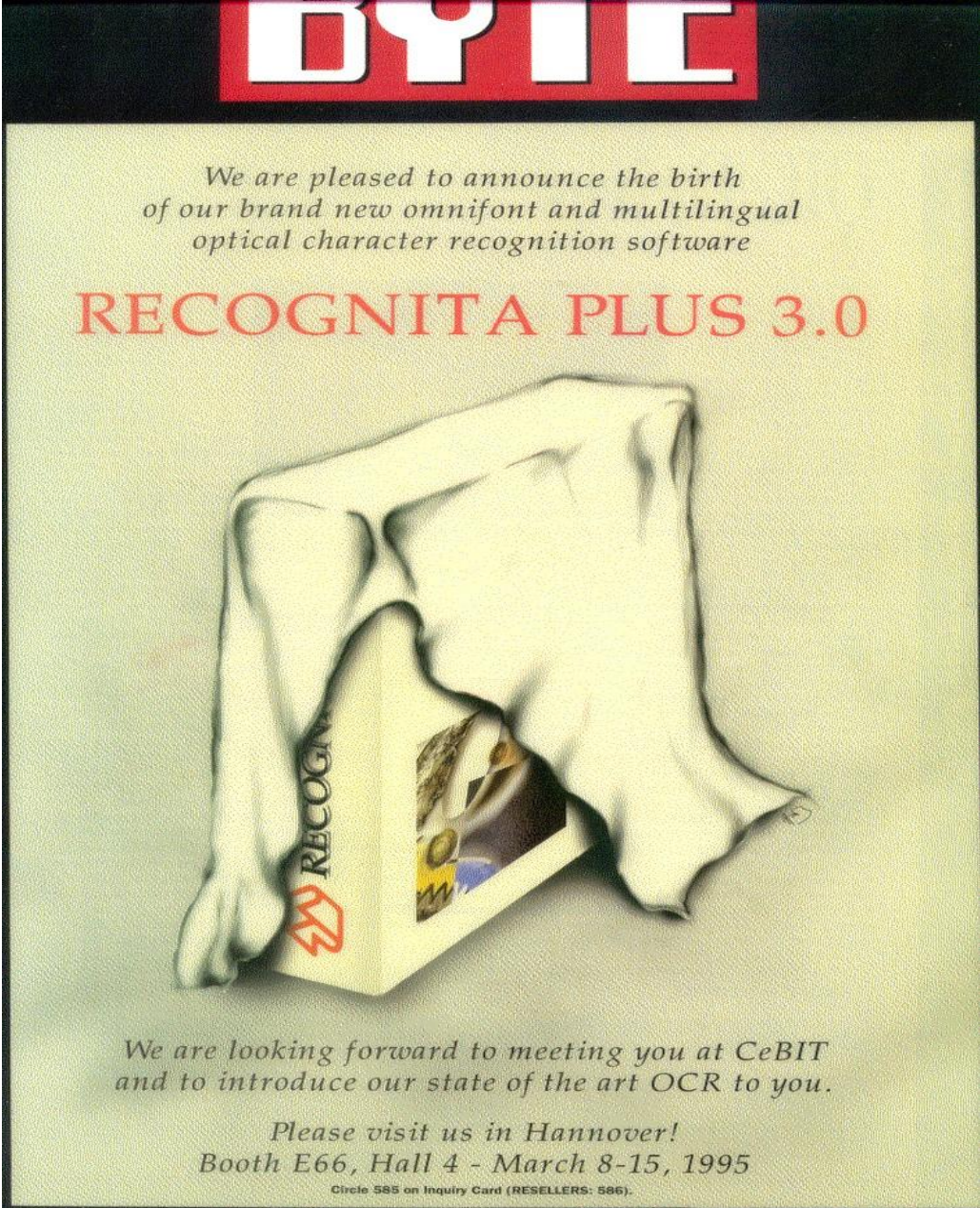
Recognita plakát

ezt mindenki
szerette



BYTE-hirdetés 1995-ből

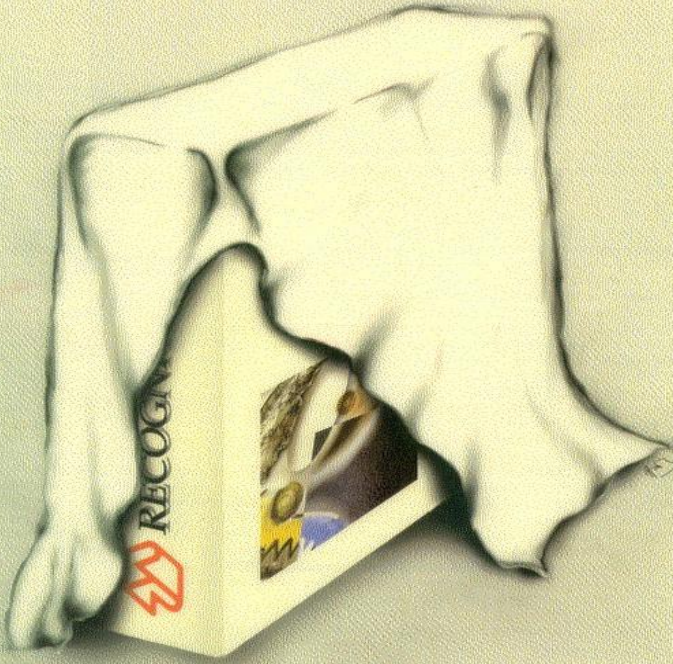
“a nem igazán lett kész” verzió



BYTE

*We are pleased to announce the birth
of our brand new omnifont and multilingual
optical character recognition software*

RECOGNITA PLUS 3.0

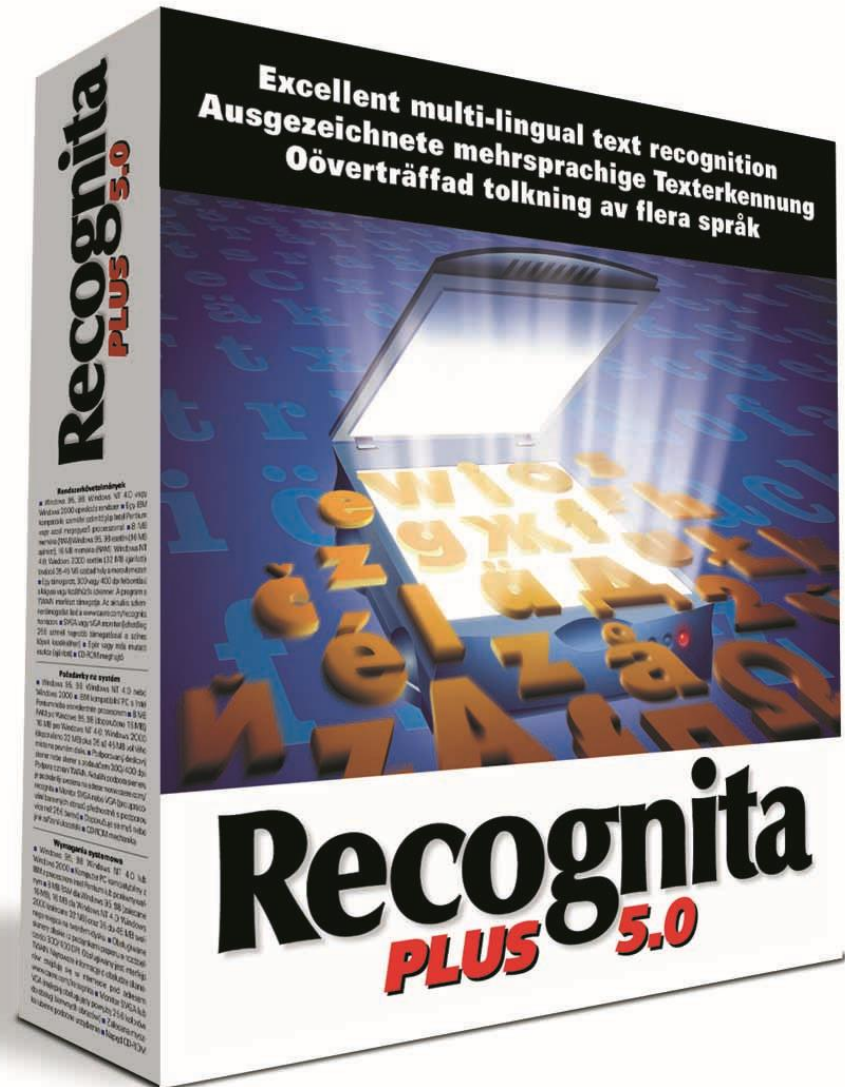


*We are looking forward to meeting you at CeBIT
and to introduce our state of the art OCR to you.*

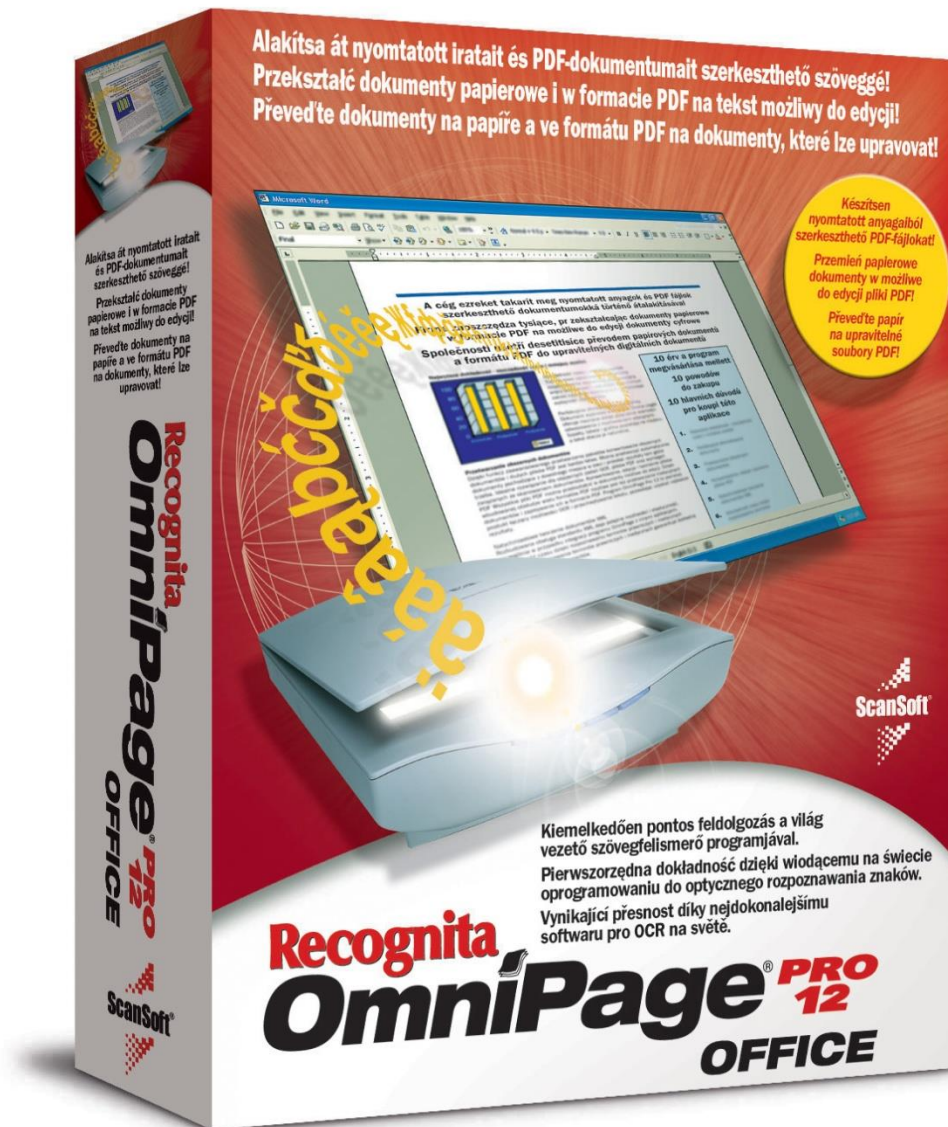
*Please visit us in Hannover!
Booth E66, Hall 4 - March 8-15, 1995*

Circle 585 on Inquiry Card (RESELLERS: 586).

Recognita Plus 5.0 – az utolsó R.P.



Recognita - OmniPage



Kofax-Recognita a Váci úton



The background of the entire page is a light gray gradient with several realistic water droplets of various sizes scattered across it. The droplets have highlights and shadows, giving them a three-dimensional appearance.

A RECOGNITA FEJLŐDÉSI STÁCIÓI

COPYRIGHT

NOT RESERVED

1989-2019

XXX

A KÖRNYEZET

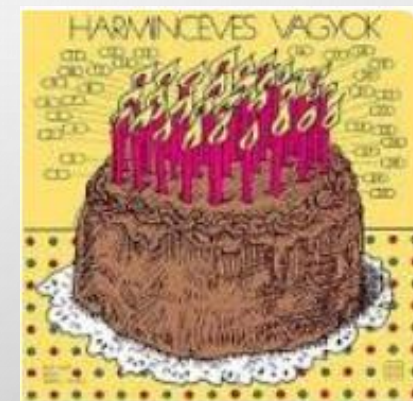
- KEDVEZŐ INFORMATIKAI KÖRNYEZET - SZERZŐI JOGOK ELISMERÉSE
 - 1969-ES III. TÖRVÉNY: KORSZERŰ, NYUGATI NORMÁKNAK MEGFELELŐ,
 - 1971 MAGYARORSZÁG ALÁÍRTA AZ EGYSÉGES SZERZŐI JOGI EGYEZMÉNYT
- IBM PROJECT CHESS: NYÍLT ARCHITEKTÚRA, NEM KIZÁRÓLAG IBM KOMPONENSEKBŐL
 - 1979 INTEL 8088 MIKROPROCESSZOR
 - 1981 IBM PC ÉS IBM KOMPATIBILIS PC-K (*IBM PC TECHNICAL REFERENCE MANUAL*)
IBM BIOS, MS DOS/MS WINDOWS
 - EGYSÉGESÍTETT HW SW KÖRNYEZET, AMIRE ÉRDEMES VOLT FEJLESZTENI
 - IGÉNY ÉS LEHETŐSÉG AZ EGYÉNI DESKTOP PRODUCTIVITY-RE
- MEGJELENNEK A JÓ KÉPMINŐSÉGET PRODUKÁLÓ ASZTALI SZKENNEREK
 - 1986 WORLD FIRST DESKTOP HALFTONE SCANNER - MICROTEK

A BÖLCSŐ

- SZÁMÍTÁSTECHNIKAI KOORDINÁCIÓS INTÉZET (SZKI)
 - 1969-BEN JÖTT LÉTRE NÁRAY ZSOLT FŐKONSTRUKTŐR VEZETÉSÉVEL A KGST EGYSÉGES SZÁMÍTÓGÉP RENDSZER (AZ ESZR)
 - Ő ALAPÍTOTTA MEG 1969-BEN AZ SZKI-T
FELADATA A SZÁMÍTÁSTECHNIKAI KAPACITÁSOK (KUTATÓ-FEJLESZTŐ INTÉZETEK, GYÁRAK) ÖSSZEFOGÁSA, KÉPVISELETE
 - AZ INTÉZET SZÁMOS MUNKATÁRSA SZERZETT TAPASZTALATOT NYUGAT-EURÓPA NEVES SZOFTVERFEJLESZTŐ INTÉZMÉNYÉBEN
 - MIKET FEJLESZTETT
 - R-10, R-12, R-15 SZÁMÍTÓGÉPEK; M-05X, M-08X MIKROSZÁMÍTÓGÉPEK, PROPER SZÁMÍTÓGÉP
 - ALKALMAZÁSI RENDSZEREK MIKROSZÁMÍTÓGÉPRE
 - MESTERSÉGES INTELLIGENCIA ESZKÖZÖK ÉS ALKALMAZÁSOK
 - KÉPFELDOLGOZÓ RENDSZEREK
 - SZOFTVERTECHNOLÓGIAI RENDSZER A SZOFTVERFEJLESZTÉS MINDEN CIKLUSÁNAK SZÁMÍTÓGÉPES TÁMOGATÁSÁRA
 - DOKUMENTUMFELDOLGOZÓ RENDSZEREK AZ ADATOK RÖGZÍTÉSÉTŐL, A FELDOLGOZÁSON KERESZTÜL A NYOMDÁIG
 - TERMELÉSIRÁNYÍTÓ RENDSZEREK A KÖNYVELÉSIG ÉS BÉRSZÁMFEJTÉSIG
 - MÉRNÖKI TERVEZÉS STATIKAI ÉS GÉPÉSZETI SZÁMÍTÁSAINAK SZÁMÍTÓGÉPES TÁMOGATÁSA
 - BIOELEKTROMOS JELEK ÉRTÉKELÉSÉT VÉGZŐ MIKROSZÁMÍTÓGÉP A GYÓGYSZERGYÁRTÁSBAN
 - **KARAKTERFELISMERŐ RENDSZER KAMERA KÉPEKEN**

NYITÁNY

- KELL EGY CSAPAT
 - KOVÁCS EMŐKE A PRAGMATIKUS KUTATÓ, MAROSI ISTVÁN A BRILIÁNS ALGORITMUS FEJLESZTŐ ÉS PROGRAMOZÓ,
TÁLLAI BENEDEK AZ ÜZLETEMBER (1991-BEN MINDHÁRMAN SZÉCHENYI DÍJASOK LETTEK)
- EARLY ADOPTERS
 - NÉHÁNY VISZONTELADÓ ELSŐSORBAN NÉMET NYELVTERÜLETEN
 - 3,000 MÁRKA FELETTI ÁR -> NÉHÁNY SZÁZ ELADOTT TERMÉK ÉVENTE
- 89 A VÁLTOZÁSOK ÉVE
 - AZ ÍGÉRETES MŰHELYEK ÖNÁLLÓSODÁSA SZKI TÖBBSÉGI TULAJDONNAL: IQSOFT, RECOGNITA,...
 - SZKI-RECOGNITA TULAJDONOSOK: SZKI, EGY ANGOL MAGÁNBEFEKTETŐ ÉS DOLGOZÓK
- 1990 GYORS VÁLLALAT-GRÜNDOLÁS KELET NÉMETORSZÁGBAN PONT AZ EGYESÍTÉS ELŐTT
- 1991 LEÁNYVÁLLALAT KALIFORNIÁBAN



PÉNZÜGYI KIHÍVÁS

- A TERJESZKEDÉS KIÉHEZTETI A CÉGET – TELJES VAGYONVESZTÉS
 - A MAGYAR BANKOK VÁLLALAT-ÉRTÉKELÉSÉNEK ALAPJA AZ ASZTALOK ÉS SZÁMÍTÓGÉPEK SZÁMA
 - EZ HATÁROZZA MEG A VÁLLALAT ÉRTÉKET
 - A SZELLEMI TULAJDON NEM SZÁMÍT
 - A TULAJDONOSOK NEM TUDNAK TŐKÉT EMELNI
 - NAPI FIZETÉSI GONDOK, HADIKOMMUNIZMUS
 - A VEZETŐSÉG SZEMÉLYES KOCKÁZATVÁLLALÁSA SZÜKSÉGES A HITELHEZ JUTÁSHOZ
- 1993-1994 MAVA BEFEKTETÉS
 - SZILÁRD PÉNZÜGYI ALAPOK
 - PROGRESSZÍV VEZETŐ DR RESZLER ÁKOS
 - AKI NEM MEGTAKARÍTANI AKARJA A KÖLTSÉGET, HANEM MEGTERMELNI

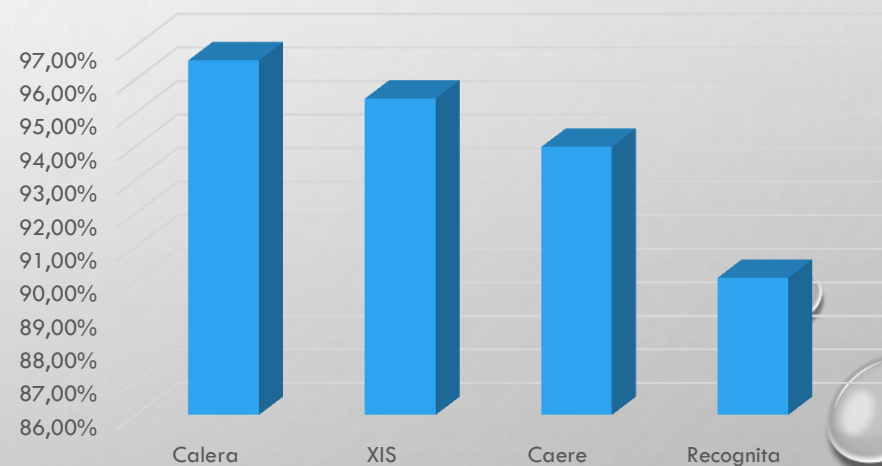
SZERVEZETTSÉGI KIHÍVÁS

- KREATÍV KÁOSZ
 - HIHETETLEN MOTIVÁCIÓ ÉS CSAPATSZELLEM
 - FOLYAMATOS LENDÜLETES FEJLESZTÉS
 - FLOPPYN BEGYŰJTÖTT KÓDRÉSZLETEK
 - TESZTELŐ CSAPAT NÉLKÜL
 - SZABÁLYOZOTT FOLYAMATOK NÉLKÜL
 - SZALADGÁLÁSOK A REPTÉRRE, HOGY VISSZAHOZZUK JAVÍTÁSRA A TERMÉKET
- A MEGNÖVEKEDETT CSAPAT MÁΣ SZERVEZŐDÉST KÍVÁN MEG
 - ÚJ MODERN ARCHITEKTÚRA
 - 1996 ISO 9001
 - KOMOLY FESZŰLTΣÉGEK ÉΣ NÉHÁNY KULCSFEJLESZTŐ ELVESZTÉSE

PONTOSSÁGI KIHÍVÁS

- ALULMARADUNK A SZKENNERGYÁRTÓK SZERZŐDÉSEIÉRT FOLYTATOTT VERSENGÉSBEN
 - AZ AMERIKAI LEÁNYVÁLLALAT KÉTSÉGEKET FOGALMAZ MEG A TECHNOLOGIA MINŐSÉGÉVEL KAPCSOLATBAN
- MEGMÉRÉTTETÉS
 - BENEVEZÜNK A UNLV (UNIVERSITY OF NEVADA LAS VEGAS) FELMÉRÉSBE
 - BIZTONSÁG-POLITIKAILAG FONTOS TECHNOLOGIA
 - DOE TÁMOGATOTT ÉVES FELMÉRÉS
 - TESZT METODOLÓGIA, TESZT KÉSZLETEK, KIÉRTÉKELŐ RENDSZER
 - AZ ELSŐ POFON A LEGNAGYOBB

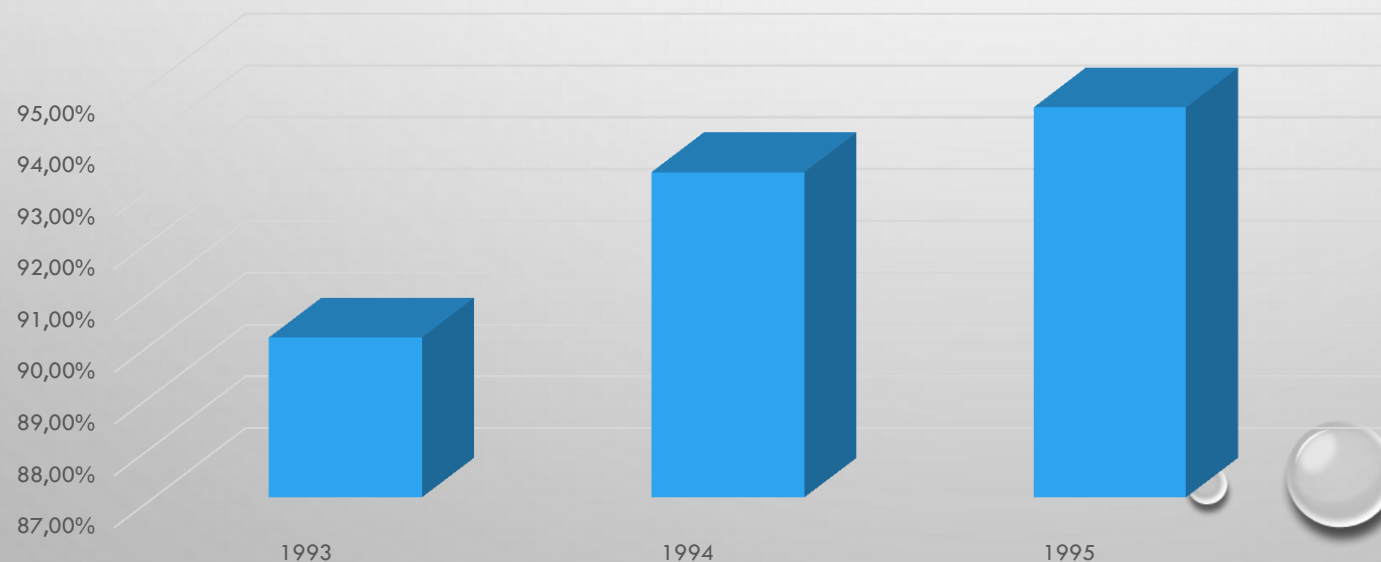
Word accuracy on magazin samples in 1993



PONTOSSÁGI KIHÍVÁS

- 3 ÉV ALATT 1996-RA A TESZTADATOK FELHASZNÁLÁSA, A FEJLESZTÉS ÉS A MÓDSZERTAN ADAPTÁLÁSA DRÁMAI NÖVEKEDÉST EREDMÉNYEZETT
- MINDEZ EGY 100 NYELVET ENGEDÉLYEZŐ TECHNOLOGIÁVAL

Recognita word accuracy on magazin samples



MINŐSÉGI KIHÍVÁS

- CAERE AKVIZÍCIÓ 1996-BAN (OCR PIACVEZETŐ)
 - HOZZÁFÉRÉS A TECHNOLOGIÁHOZ
 - FELTÉTELEZETT EURÓPAI ÉS OEM PIAC MEGSZERZÉSE
 - EGY FELTÖREKVŐ VERSENYTÁRS KIIKTATÁSA
- KIVÍVTUK A MEGBECSÜLÉST, BEVONTAK A PIACVEZETŐ TERMÉK FEJLESZTÉSÉBE
- EZ A TERMÉK (OMNIPAGE) SZÁZSZOR NAGYOBB PÉLDÁNYSZÁMBAN KELT EL
 - AZ ESETLEGES HIBÁK NEM/NEHEZEN KEZELHETŐEK A KIADÁS UTÁN
- HOGY RADIKÁLISAN CSÖKKENTSÜK A KOCKÁZATOT
 - MEGERŐSÍTETTÜK FOLYAMATAINKAT (TERVEZÉS, DOKUMENTÁLÁS, PREZENTÁLÁS)
 - ÁTVETTÜK A QA KONCEPCIÓJUKAT. MEGERŐSÍTETTÜK A QA SZERVEZETET ÉS FOLYAMATOKAT
 - ÉS EZEKTŐL KICSIT LE IS LASSULT AZ INNOVÁCIÓ

TÖBBTERMÉKES KIHÍVÁS

- A SCANSOFT AKVIZÍCIÓ 2000-BEN
 - A CAERE BEFEKTETŐK ELÉGEDETLENSÉGE A CAERE MENEDZSMENTTEL
 - A SCANSOFT JOBB SZEREPLÉSE A MICROSOFT OCR TENDEREN (200 DPI GRAY MINTÁN ÉS AZ ÁZSIAI NYELVEK TÁMOGATÁSÁVAL)
- A SCANSOFT-NAK VOLT A DESKTOP OCR MELLETT EGY MÁSIK PIACVEZETŐ MEGOLDÁSA A PAPERPORT – DESKTOP DOCUMENT MANAGEMENT SYSTEM
- HARMADIKKÉNT CSATLAKOZOTT A PDF TERMÉKCSALÁD ÉS A MINDENT ÁTHATÓ PDF TECHNOLÓGIA
- A PÁRHUZAMOSAN FUTÓ TERMÉKKIBOCSÁTÁSOK ELMOZDULÁST IGÉNYELTEK A BANZAI STÍLUSTÓL
- A SCANSOFT (KORÁBBAN XIS) TOVÁBBVITTE A XEROX “TIME TO MARKET” METODOLÓGIÁJÁT, AMI TUDTA KEZELNI A PÁRHUZAMOS, HASONLÓ SÚLYÚ TERMÉKKIBOCSÁTÁSOKAT

TECHNOLÓGAI KIHÍVÁS

- KLIENS-SZERVER

- AZ ECOPY AKVIZÍCIÓ 2009-BEN VÉGET VETETT A KIZÁRÓLAGOS DESKTOP ORIENTÁCIÓNAK
- AZ MFP (MULTI FUNCTIONAL PERIPHERY) VEZÉRLÉSE KLIENS-SZERVER MEGOLDÁS
 - A KLIENS OLDALON EMBEDDED JAVA TECHNOLÓGIÁK ÉS WEB SERVICE-EK
 - A SZERVER OLDALON WEB SZERVER, SQL ADATBÁZISOK, HIGH AVAILABILITY, FAIL OVER TECHNOLÓGIÁK

- FELHŐ KIHÍVÁS

- OCR SZERVER, LOKÁLIS SZERVER NÉLKÜLI MFP SZKENNELŐ MEGOLDÁS

JELENLÉGI KIHÍVÁSOK

- CONTAINER TECHNOLÓGIÁK TÁMOGATÁSA
- AZ AGILIS METODOLÓGIA MEGTÁMOGATÁSA KITERJESZTETT TESZTAUTOMATIZÁLÁSSAL
- A TECHNOLÓGIABESZÁLLÍTÓK MINŐSÉGBIZTOSÍTÁSÁNAK ERŐSÍTÉSE
- MOTIVÁLT, HATÉKONY ÉS FELELŐS MÁTRIX SZERVEZET
- A SOK MILLIÓS ÖRÖKSÉG
- GENERÁCIÓVÁLTÁS
- MESTERSÉGES INTELLIGENCIA

Recognita

caere

ScanSoft

NUANCE

KOFAX

A stabilitás fenntartása kihívások közepette

avagy, miből lesz a cserebogár...

eCopy

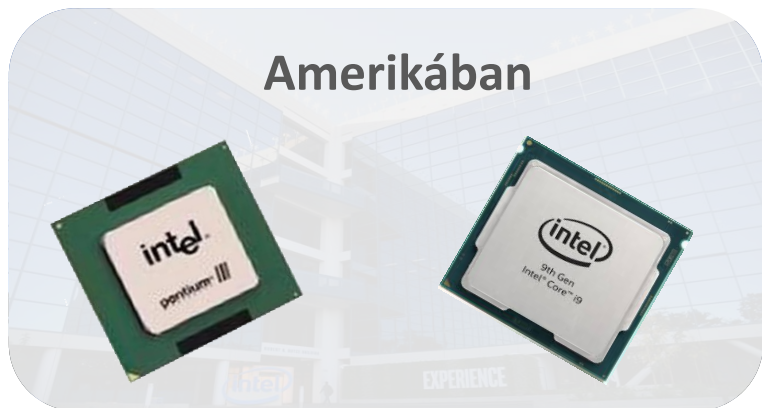
copitrak

Notable
solutions

X-Solutions®

equitrac

Safe Com

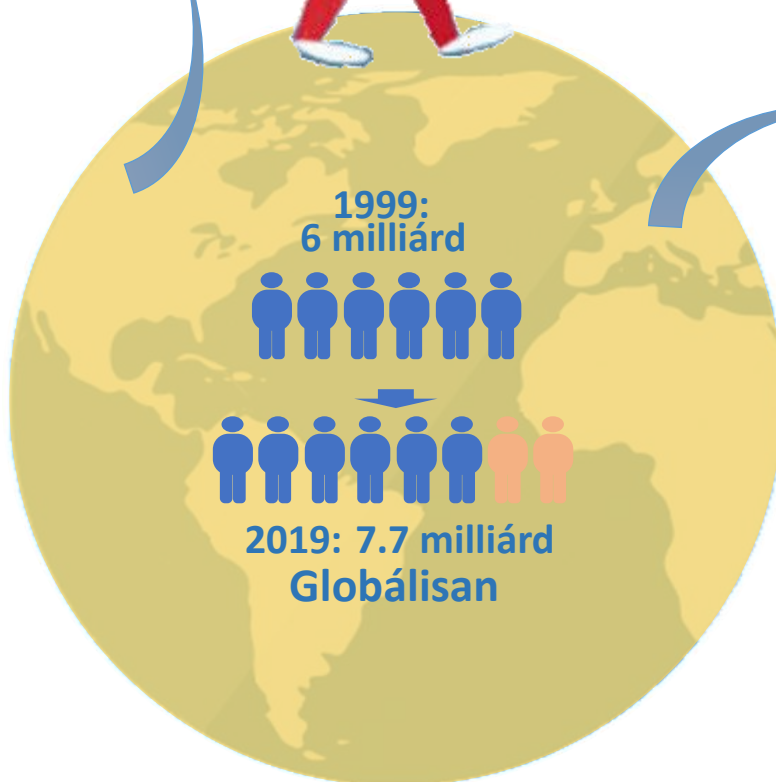


Amerikában

- Pentium III vs. Core i9
 - 3,100 x nagyobb teljesítmény
 - 700 x több tranzisztor

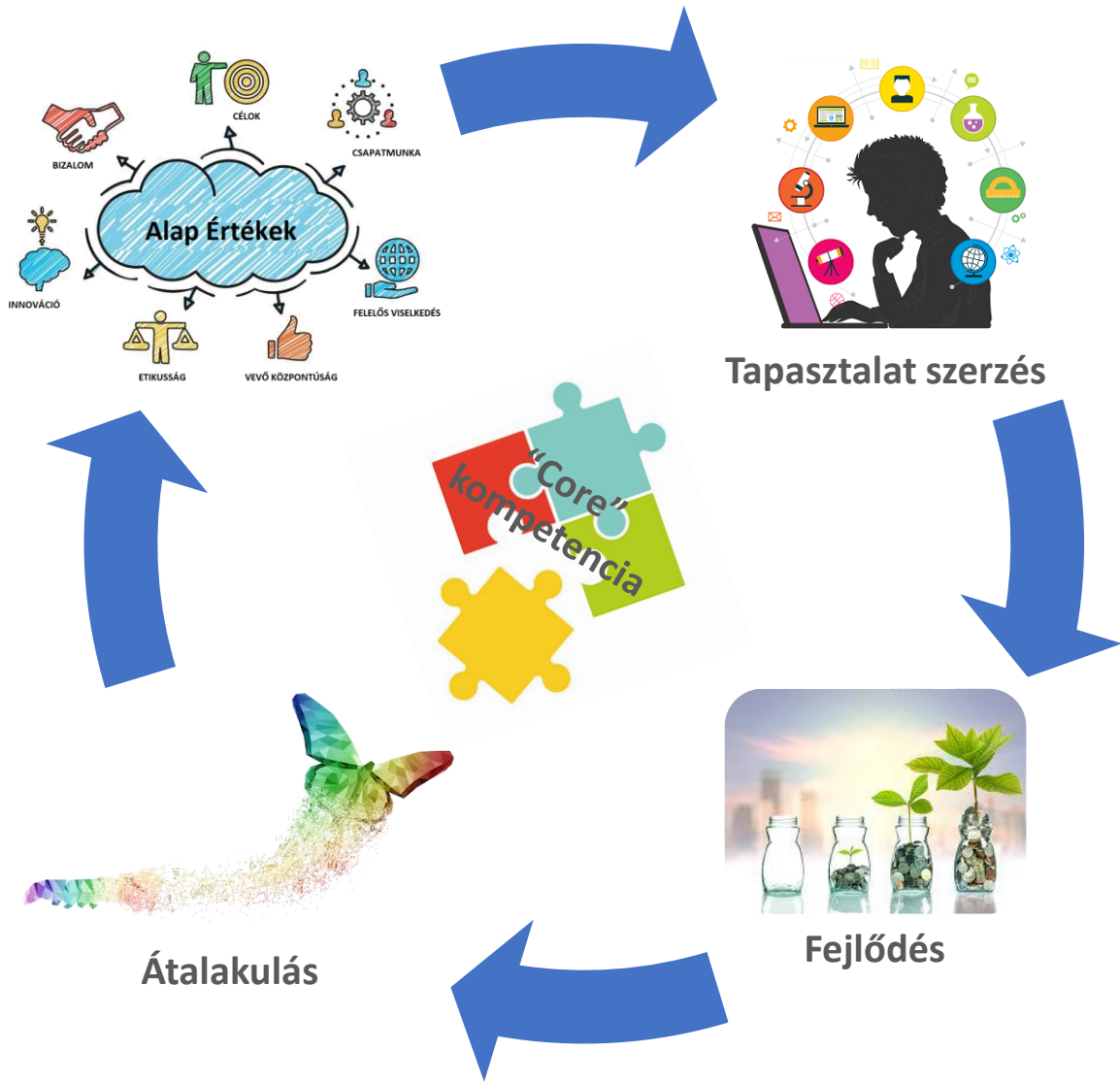


- 4, 500 munkanap
- 35, 500 munkaóra
- 128 millió másodperc
- 64, 000 kilométer



Közép-Európa szívében









Kultúra


A kreatív káoszból a jól
szervezett kiszámíthatóságig




*„Minden idők legnagyobb felfedezése,
hogy az ember meg tudja változtatni a
jövőjét pusztán azáltal, hogy
megváltoztatja a hozzáállását.”*

Oprah Winfrey

A történelem a közélet és Európa inspirálták




Sok az eskaláció.
Ez az új verzió rossz
minőségét mutatja.



Azért sok az eskaláció, mert
mindenki a release mielőbbi kiadásán
dolgozott, kicsit összegyűlt.
*A múltkor kevés volt, akkor azt
mondta, nincsenek
eladások...*

Optimálisan miként történhetett volna?



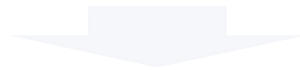
**Képzeld,
2 héttel előbb tudtunk
releaselni és az eskalációs
várakozási idő csak
5 nappal nőtt**

**Ez az pro-aktív
Kreativitás az üzleti siker
alapja. Jó döntés, látványos
eredmény.**

A sikeres átalakulás kulcsai a robosztus folyamatok és a bizalom építés volt

Folyamatok definiálása

- Belső standardok lefektetése
- Dokumentált szerepkörök és felelősségek
- ISO 9001



Kiszámíthatóság

Bizalom építés

- Pro-aktivitás vs. re-aktív szokások
- Kommunikáció: úgy értsék ahogy szeretnénk
- A sikerek láthatóvá tétele, hivatkozása (precedence)



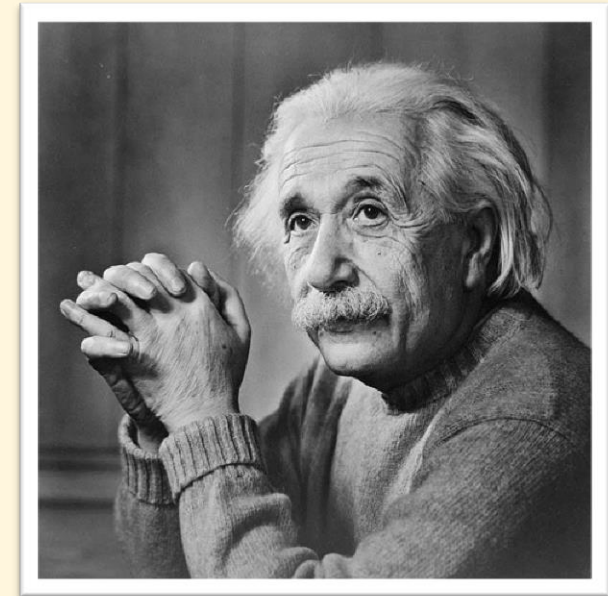
Felhatalmazás

2003-ra a ScanSoft felszámolta az észak-amerikai fejlesztő bázisát és az összes akkori termék fejlesztését Budapestre konszolidálta.



Képességek

Az alkalmazkodás és
dominancia egyensúlya



*„Mindent olyan egyszerűen kell csinálni,
amennyire csak lehetséges, de semmivel
sem egyszerűbben.”*

Albert Einstein

今すぐスキャン スキャン

フォルダ 機能ペイン

スキャン設定 サムネイル All-in-One Search プロパティ

表示の変更 並べ替え 戻る 表示

更新 アイテムを等間隔に整列 更新

開く

デスクトップオプション オプション

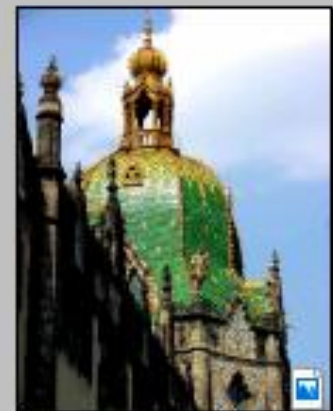
フォルダ

- マイ PaperPort ドキュメント
 - FAX
 - ウェブ ページ
 - サンプル
 - 写真
 - プレゼンテーション
 - 記事
 - 銀行取引明細書
 - 写真
 - 受領書
 - 税金
 - 投資
 - 不動産
 - 名刺
- マイウェブ ページ

ドキュメント¥マイ PaperPort ドキュメント¥サンプル¥写真



Autumn Scene



Budapest Landmark



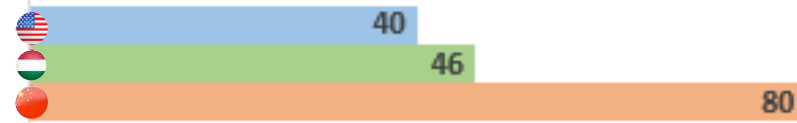
Majestic Tree

Mi át-”küzdöttük” magunkat a bizalomépítés és hatékonyság buktatóin

A mai modellek segítségével sok minden kap értelmet és magyarázatot amit megtapasztaltunk

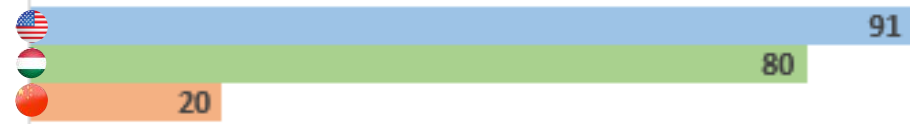
Power Distance (hatalmi távolság):

Mennyire elfogadott a társadalomban az erő egyenlőtlen eloszlása



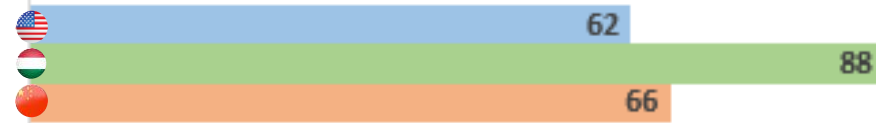
Individualism (individualizmus):

Mennyire helyezi az egyén az érdekeit a csoport érdekei elé



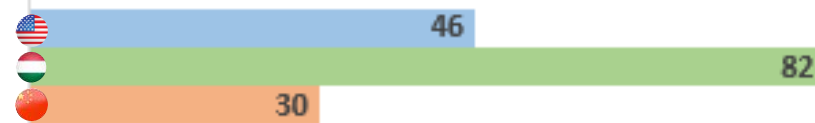
Masculinity (maszkulinitás):

Mennyire társul a sikeresség érzetéhez a másik legyőzése (vs. belső béke, boldog élet, stb.)



Uncertainty Avoidance (bizonytalanság kerülés):

Mennyire igyekeznek kiiktatni / viselik nehezen az egyének a bizonytalanságot



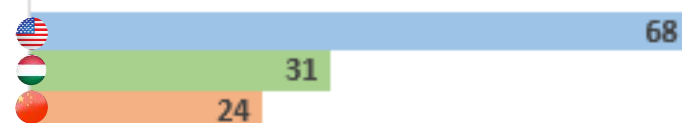
Long Term Orientation (hosszútávú orientáció):

Mennyire pragmatikus, hosszútávú értékek mentén gondolkozik a társadalom



Indulgence (boldogság):

Mennyire optimista, a boldogságot mindig pozitívumként értékelő társadalom



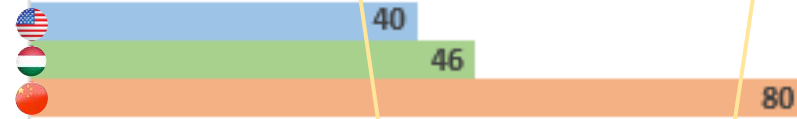
Magyarország kulturális eltérései Amerikától és Kínától Geert Hofstede modelje alapján

Negatív sztereotípiák? (magyarok az amerikaiak szemével):

Borúlító és negatív. Fafejű, nem együttműködő.

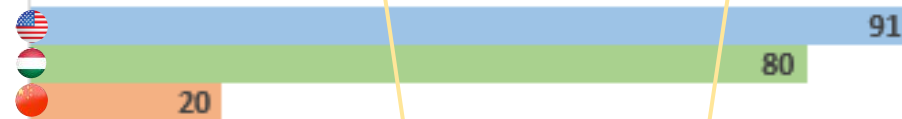
Power Distance (hatalmi távolság):

Mennyire elfogadott a társadalomban az erő egyenlőtlen eloszlása



Individualism (individualizmus):

Mennyire helyezi az egyén az érdekeit a csoport érdekei elé



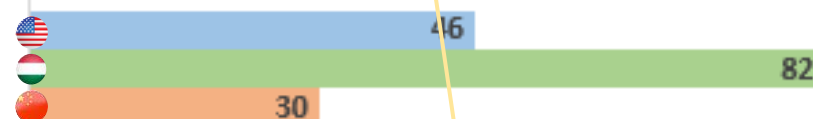
Masculinity (maszkulinitás):

Mennyire társul a sikeresség érzetéhez a másik legyőzése (vs. belső béke, boldog élet, stb.)



Uncertainty Avoidance (bizonytalanság kerülés):

Mennyire igyekeznek kiiktatni / viselik nehezen az egyének a bizonytalanságot



Long Term Orientation (hosszútávú orientáció):

Mennyire pragmatikus, hosszútávú értékek mentén gondolkozik a társadalom



Indulgence (boldogság):

Mennyire optimista, a boldogságot mindig pozitívumként értékelő társadalom

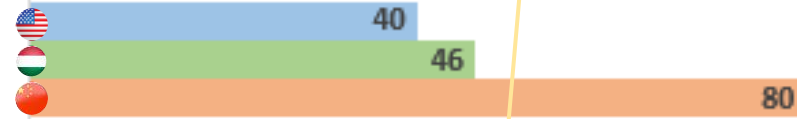


Bizalom hiány a kínai együttműködés során:

Elrejtik a problémákat, a vezetők nem képviselik a vállalat érdekeit.

Power Distance (hatalmi távolság):

Mennyire elfogadott a társadalomban az erő egyenlőtlen eloszlása



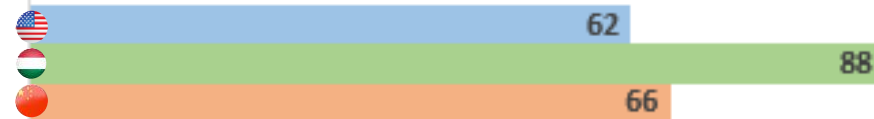
Individualism (individualizmus):

Mennyire helyezi az egyén az érdekeit a csoport érdekei elé



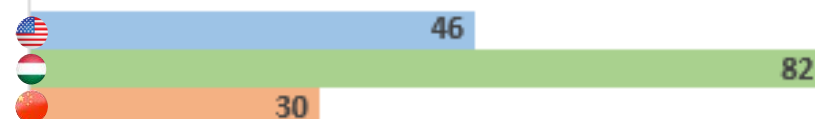
Masculinity (maszkulinitás):

Mennyire társul a sikeresség érzetéhez a másik legyőzése (vs. belső béke, boldog élet, stb.)



Uncertainty Avoidance (bizonytalanság kerülés):

Mennyire igyekeznek kiiktatni / viselik nehezen az egyének a bizonytalanságot



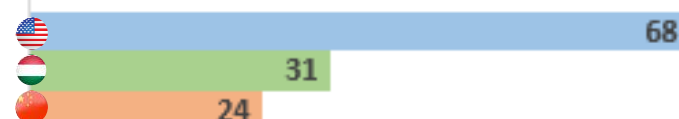
Long Term Orientation (hosszútávú orientáció):

Mennyire pragmatikus, hosszútávú értékek mentén gondolkozik a társadalom



Indulgence (boldogság):

Mennyire optimista, a boldogságot mindig pozitívumként értékelő társadalom





Tehetség

Megszerezni és megtartani



„Egy kiváló mérnök három közepessel ér fel.”

Elon Musk

Az új felvételek...

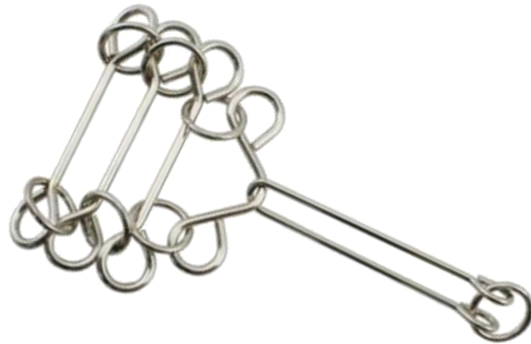
... sosem könnyűek



Viszont amikor

- 35 embert kell felvenni 6 féle képesítéssel,
- új szakterületre, nincs „mibe” integrálni
- mindezt fél éven belül,
- hogy azután 70 tapasztalt mérnök munkáját vegye át...

...az inkább ilyen:



mint ilyen:





Interjú

- Szakma & HR, nincs további kör

Teszt

- Gyakorlatias, kalibrált

Előszűrés

- 250 jelöltből 200 megkeresés



Átlátszóság



Fejlődés



Első benyomás



Kultúra
és
légkör



„Recognition”



„Hideg van odakint”



Lojalitás

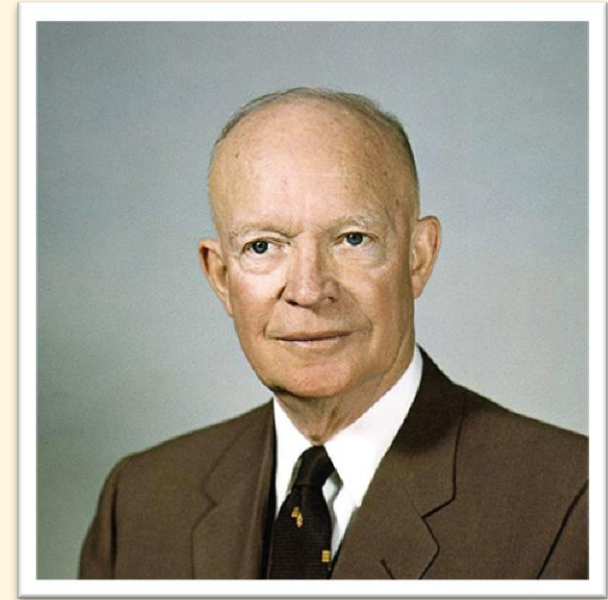


Anyagi juttatások



Management

Jó terv és végrehajtás,
tapasztalat és összetartás



„A motiváció az, amikor ráveszed az embereket, hogy tegyék, amit te akarsz, és gondolják azt, hogy ők akarják.”

Dwight David Eisenhower

Együttműködés



Összetartás



Átlátszóság



Ha esik, ha fúj: kedd 10 AM vezetői értekezlet

Köszönöm a figyelmet!



Az előadóról:



Novák Attila vagyok. A Kofax-Recognita ZRT-nél fejlesztési igazgatóként dolgozom. Informatika és üzleti területekről szerzett MSc és MBA végzettséggel egyaránt rendelkezem. 1999-ben csatlakoztam a Caere-Recognita vállalathoz amelyet megvásárolt a ScanSoft és később felvette a Nuance nevet.

Munkásságom folyamán részt vettem több akvizíció végrehajtásában és az azt követő integrációban. OEM, felhő és szerver alapú nagyvállalati megoldások fejlesztésének egyaránt voltam felelős vezetője.

Mélyen hiszek abban, hogy a fenntartható siker és versenyelőny alapjai az erős és egészséges vállalati kultúra, valamint a nagy gonddal felépített csapatok. Szabadidőmben 7 éves kislányom boldog édesapja vagyok.

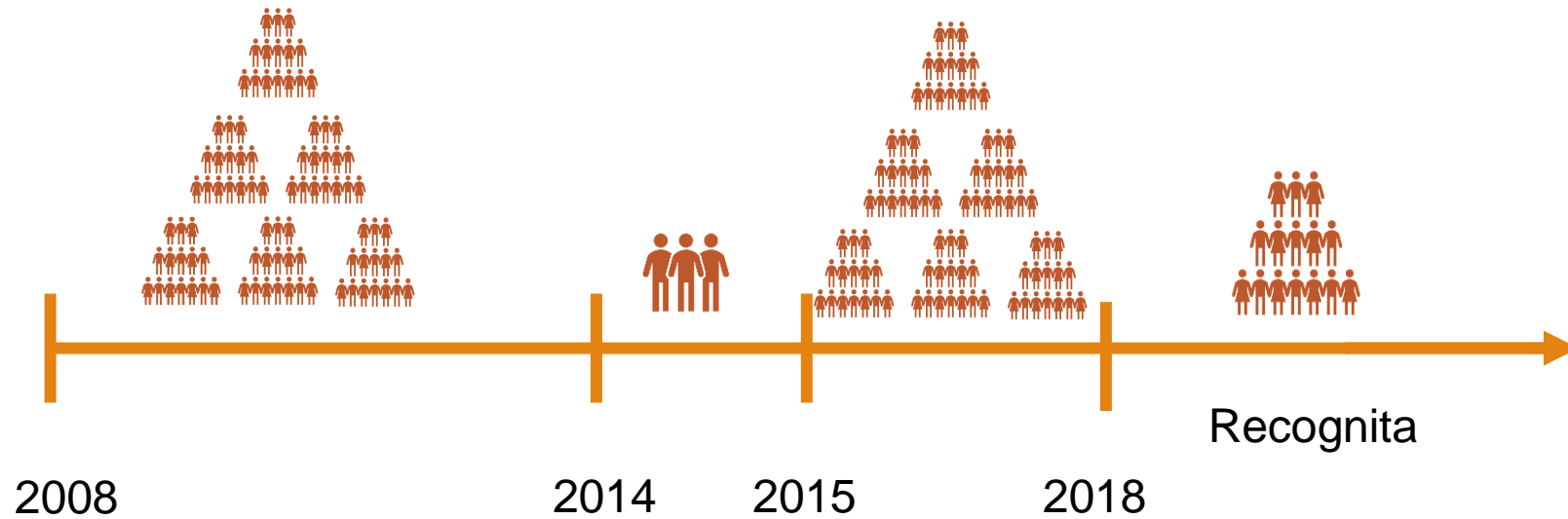
<https://www.linkedin.com/in/novakattila/>

Generációk jönnek-mennek...

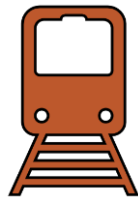
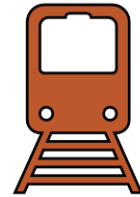
...de a Recognita örök

BERLINER PÉTER - A RECOGNITA ÉRTÉKEK FRISS SZEMMEL

Előzmények



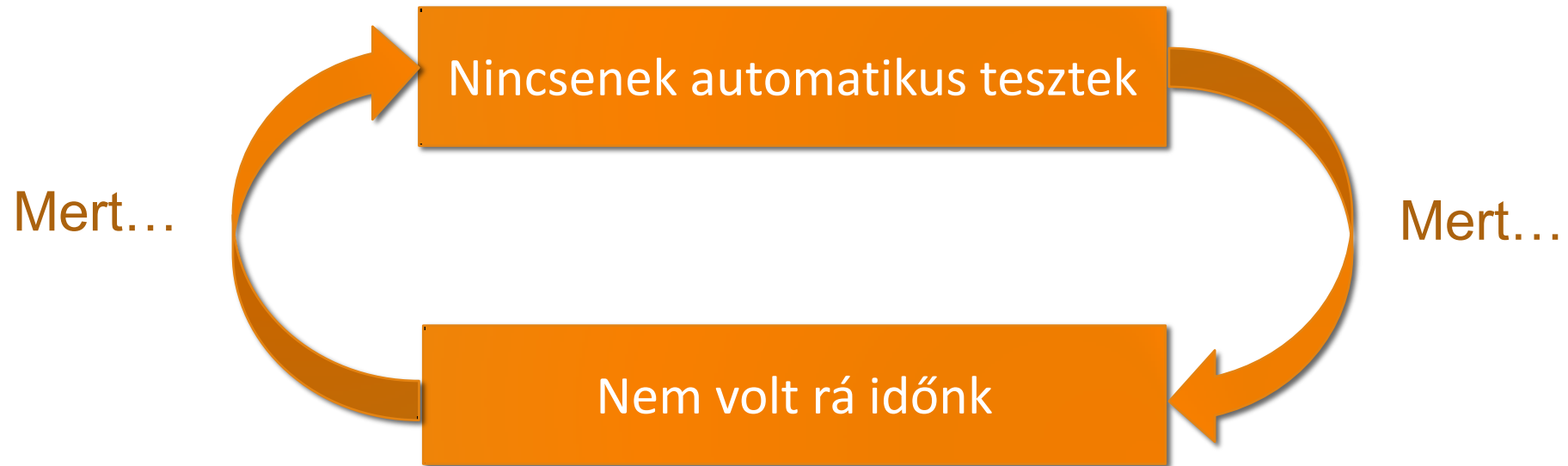
A váltás



Recognita

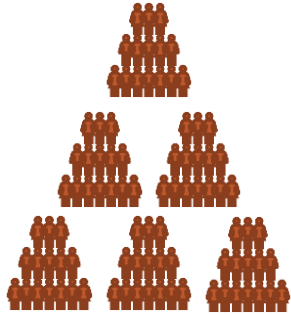


Kis kitérő az automatizált tesztek világába...



Tapasztalatok

Nagy cég - kis cég - arany középút?



2000+ alkalmazott itthon

Szigorú folyamatok

Évekig ugyanaz

2-3 fejlesztő

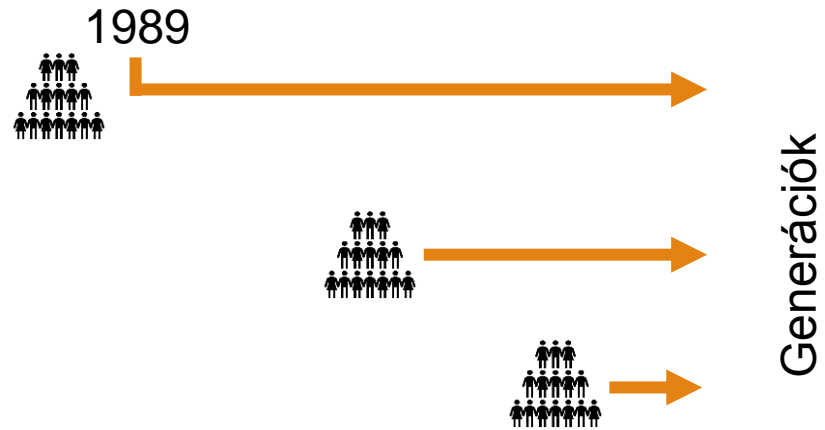
Ad-hoc fejlesztés

Naponta új irány



„Szürke” hétköznapiok...

Stabilitás és Rugalmasság



Összefoglalva...

Köszönöm a figyelmet!



A karakterfelismerés buktatói

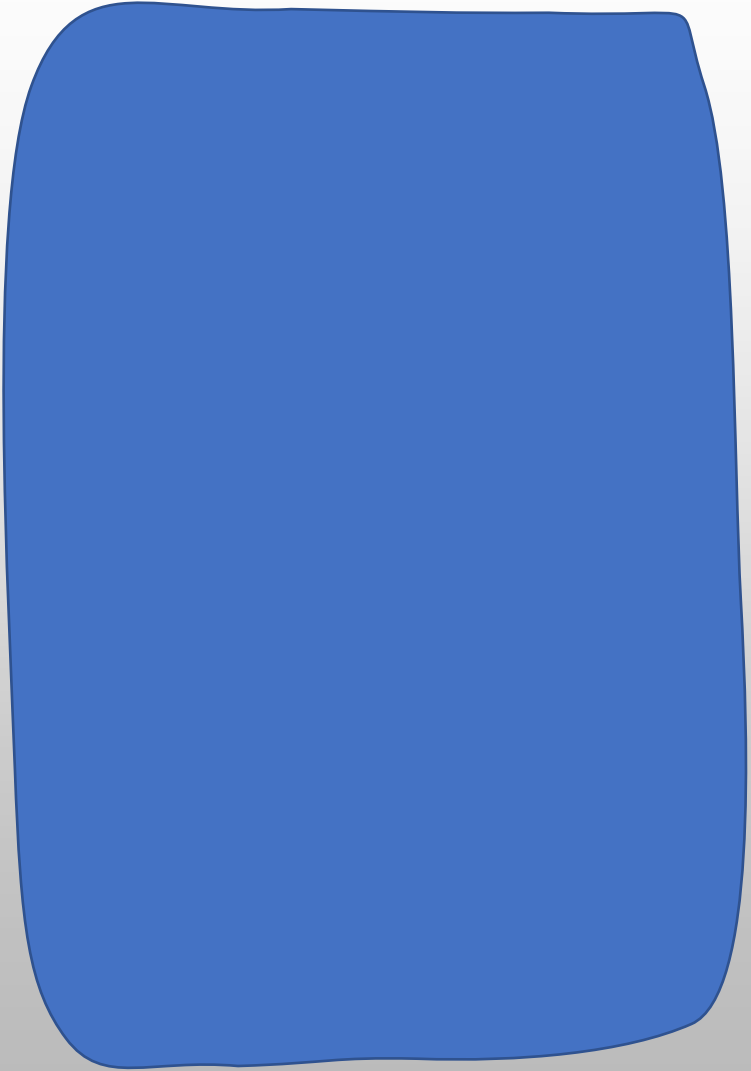
Szakmai szekció –
Mit főzünk a fazékban?

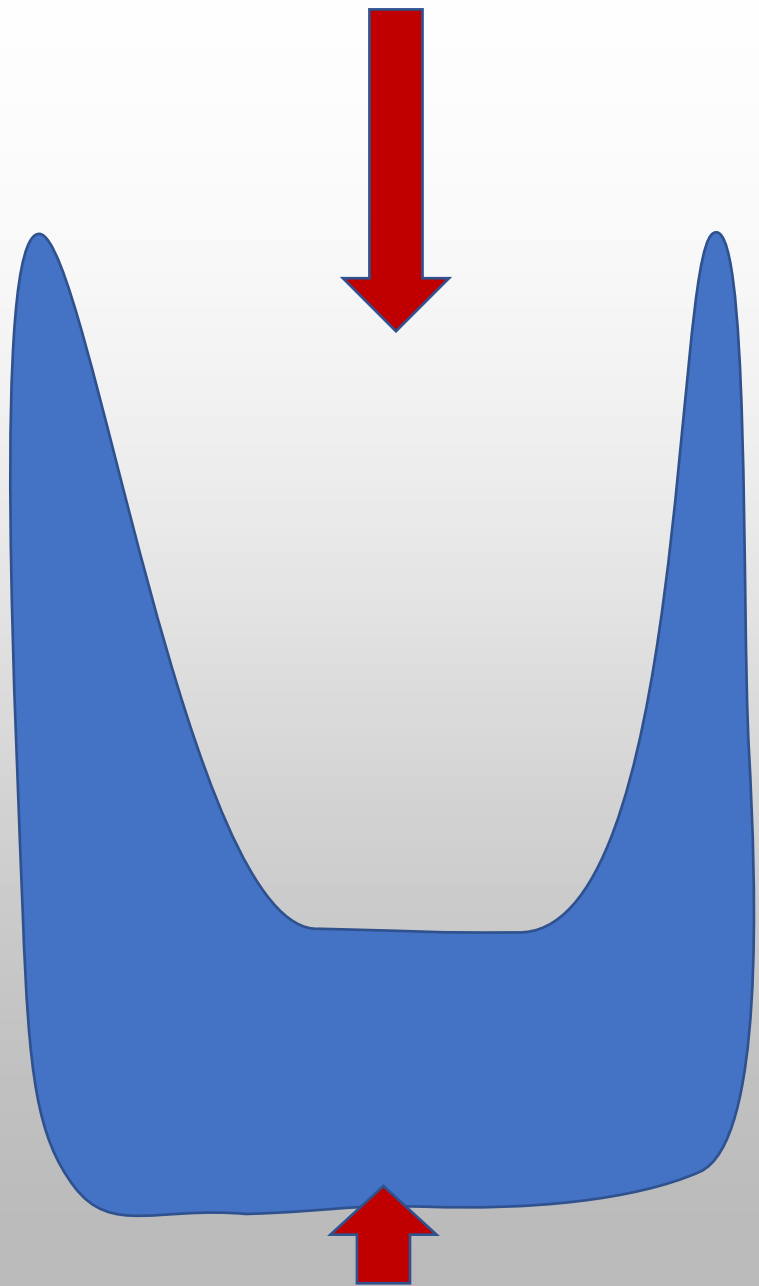
dr. Marosi István

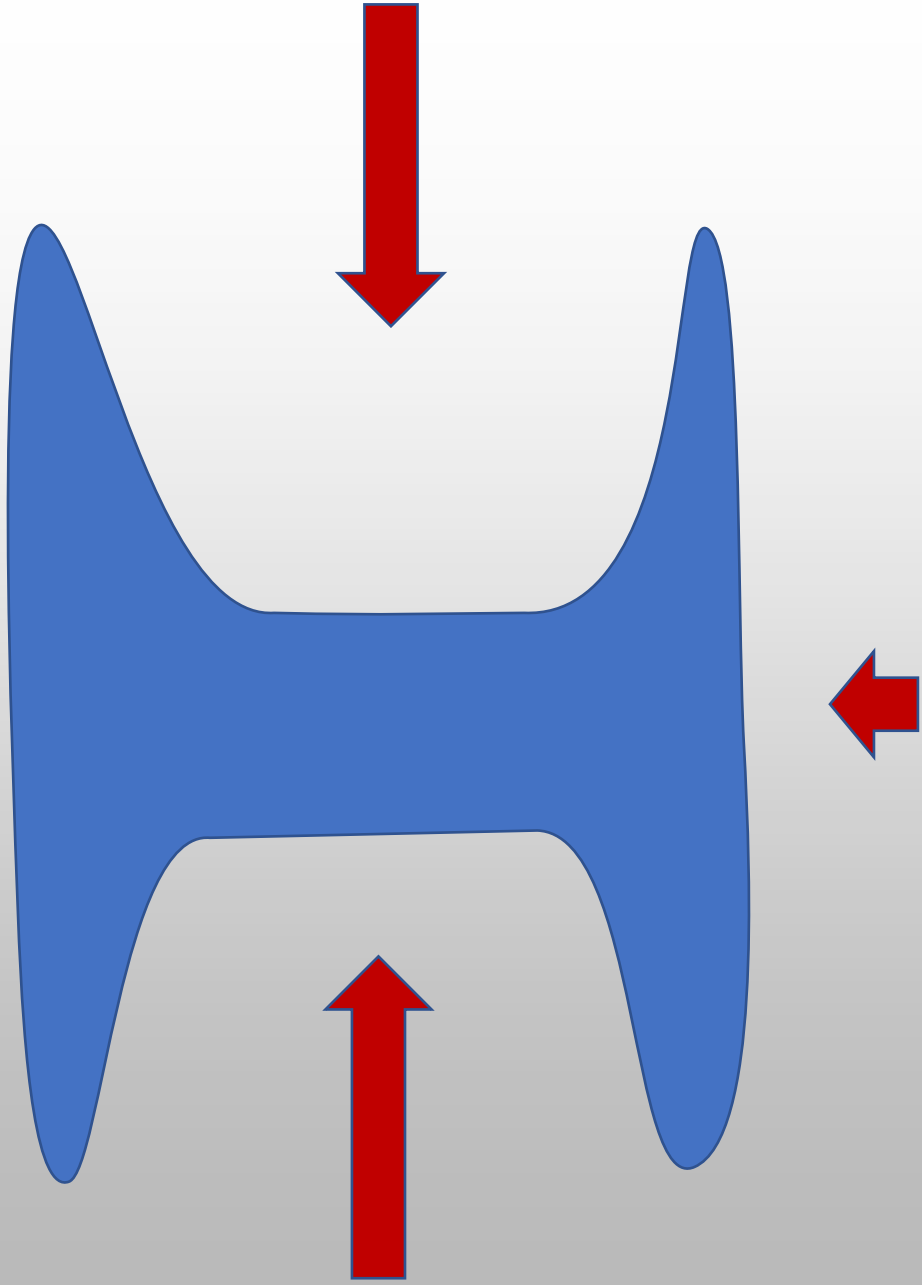
A karakterfelismerés buktatói

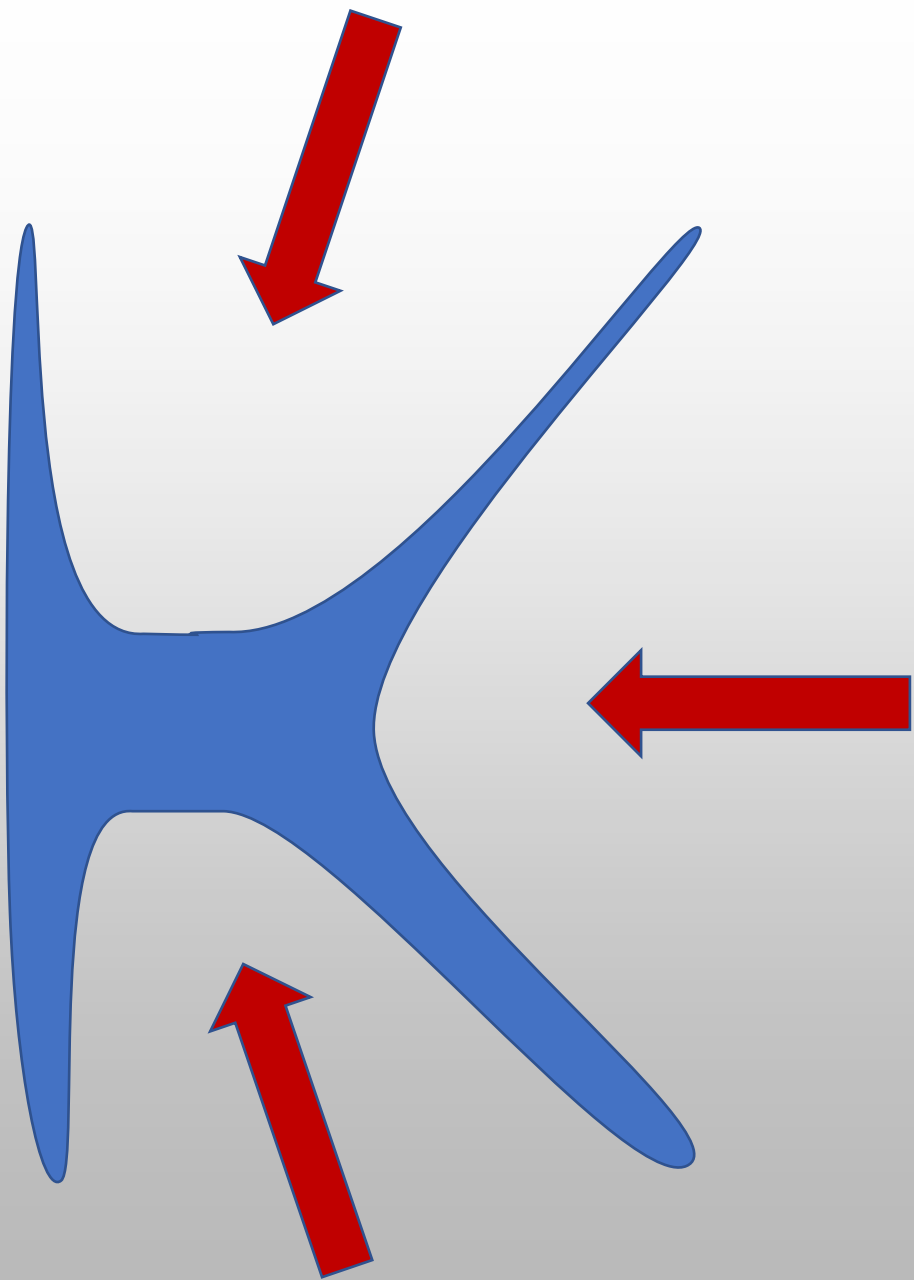
- A kezdetek – 35 éve
 - 1984, Kovács Emőke, SzKI
 - M08X számítógép
 - Kamera felbontás: 384×288 pixel
 - Megvilágítástól erősen függ a kép
 - Nincs éles kép
 - Matrix-matching nem járható út
 - Ötlet: gyurma





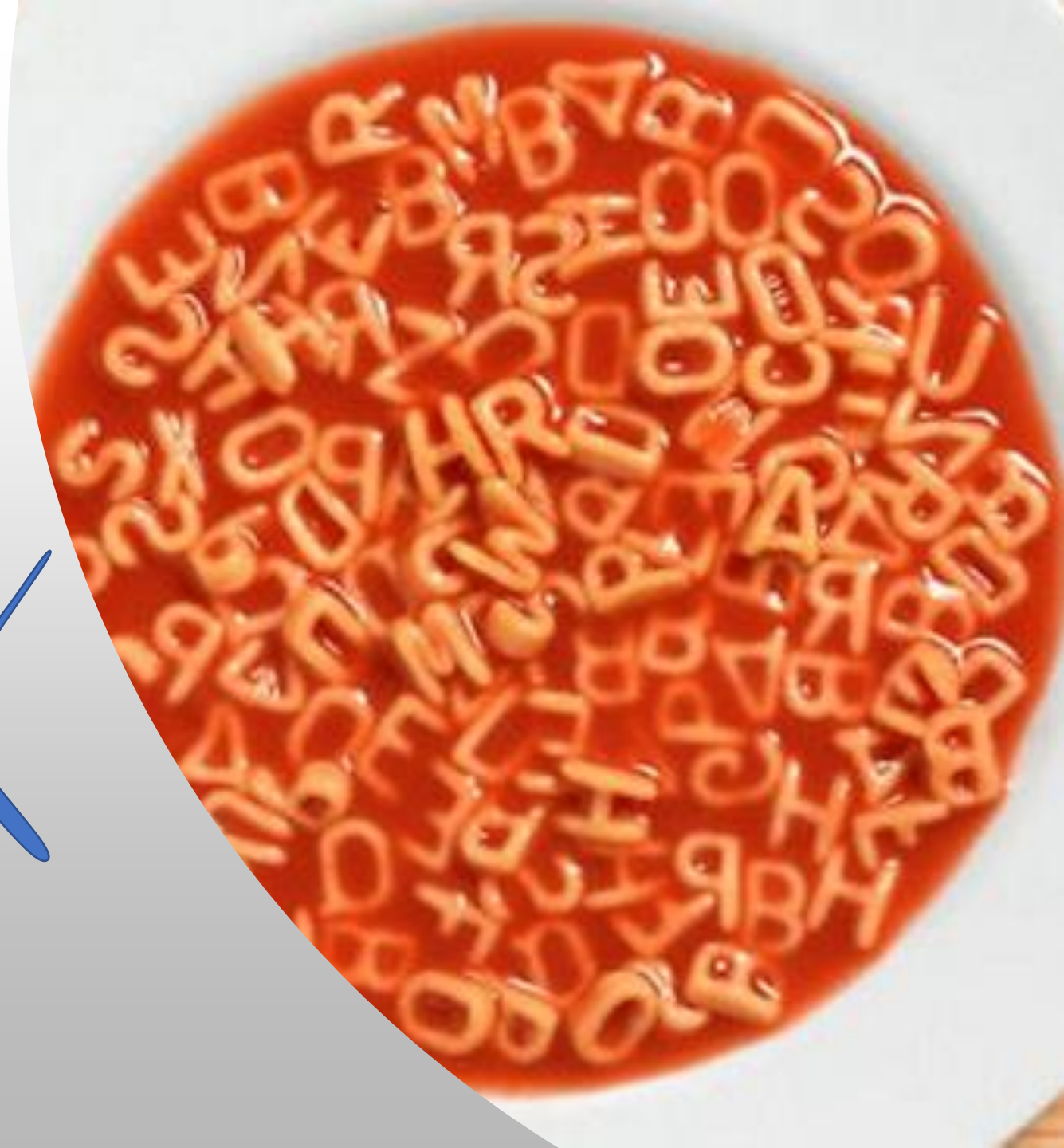






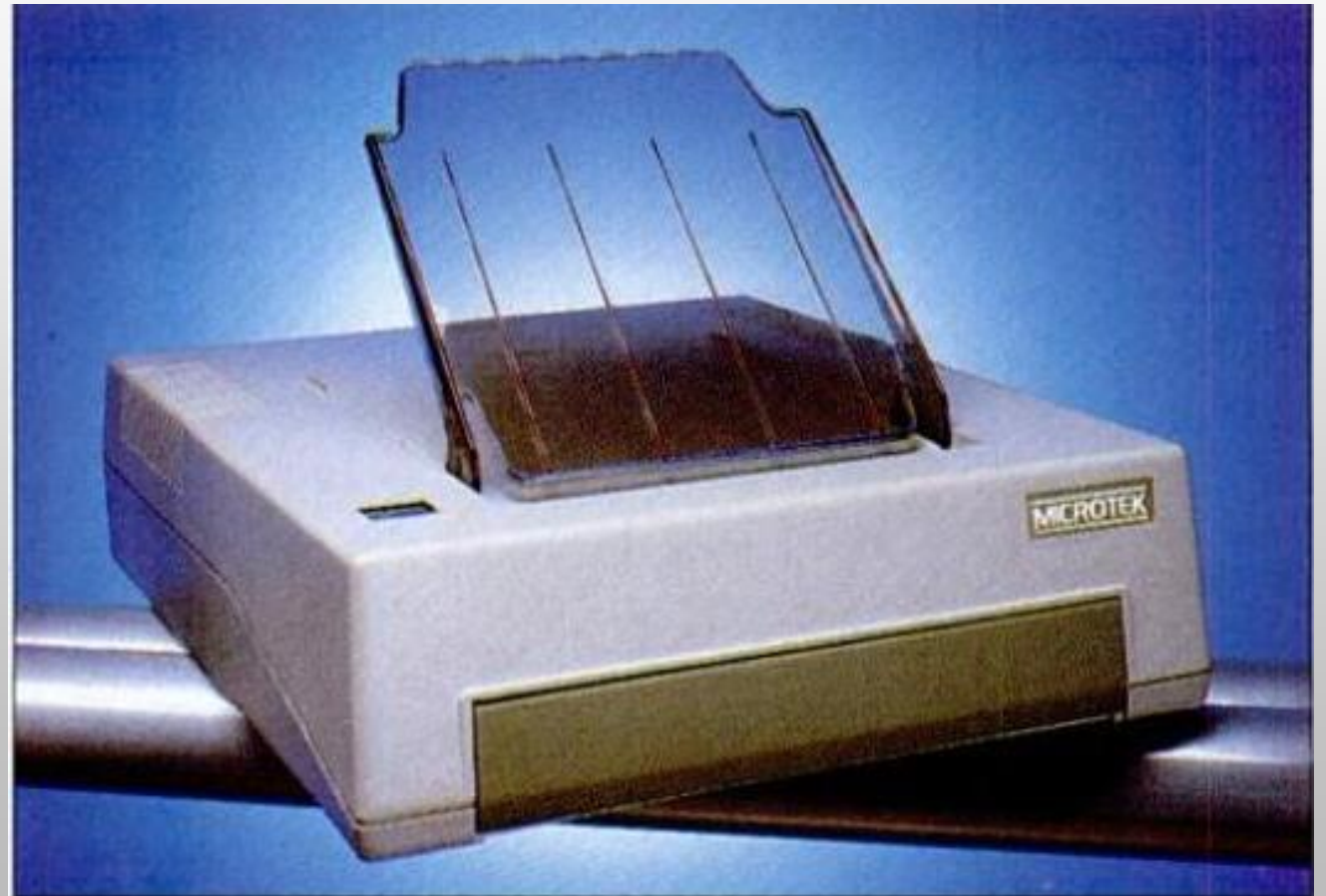
A karakterfelismerés buktatói

- Feature-ök:
 - A konkáv öblök helye és iránya
- Kézzel rajzolt karakter-kártyák
- Első program:
 - Kontúrkövetés alapú
 - Bedrótozott logika
 - If blokkok sorozata



A karakterfelismerés buktatói

- 1985:
 - Kovács Emőke, Marosi István
- Új hardware
 - Microtek scanner
 - MS-300A: 300 dpi BW
 - IBM PC klón (Proper-16)
 - Multitech grafikus kártya
 - 1024×768 monochrome
 - (CGA BW: 640×200)
- Grafikus – ablakos interfész
 - BitBlt műveletek
 - Egér nincs – szimulátor
 - Scanner és BMP file kezelés
 - ASM-ben írtuk
 - 1985-ben természetesnek tűnt



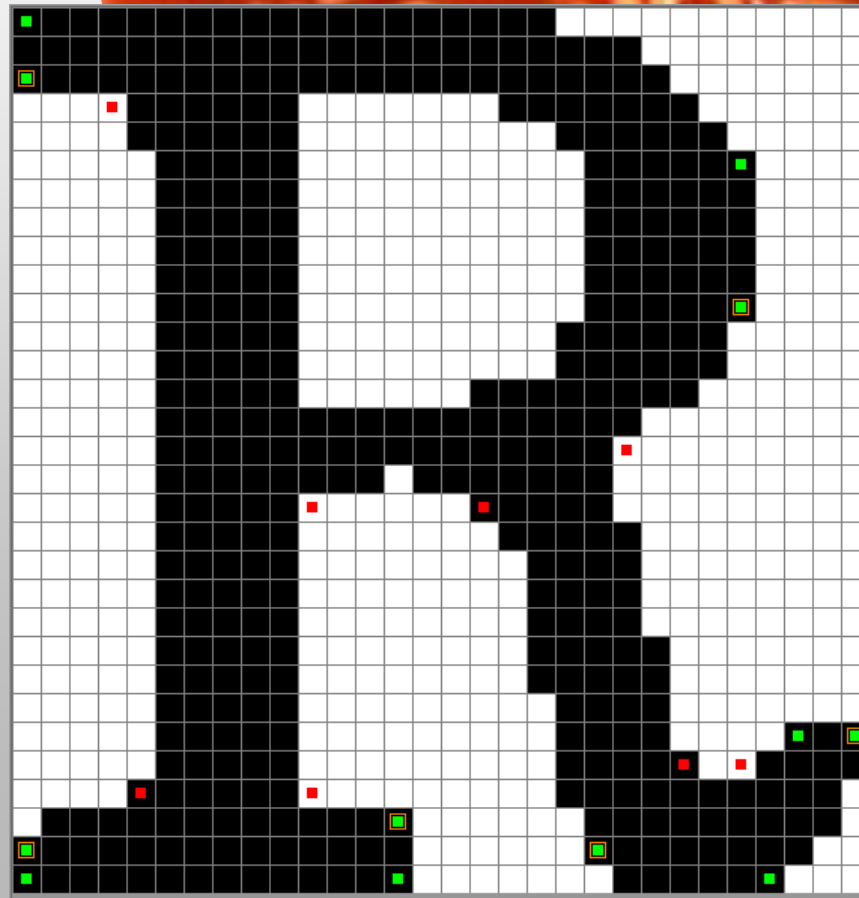
A karakterfelismerés buktatói

- 16 bites IBM PC
 - Csak 64k-t lehet könnyen címezni
 - Egy A4-es BW oldal 1.04 MB
 - Csak egy része fér be!
 - 128 képsor, gyűrűpuffer
- 1986: Első „nagy” bug:
 - Program > 32k
- 1996: 32 bites kód
 - `movlp es::di,cx::dx`
→ `mov edi,edx`
- 2012: 64 bites kód
 - Ugyanabból a forrásból 3 architektúra
→ `mov rdi,rdx`



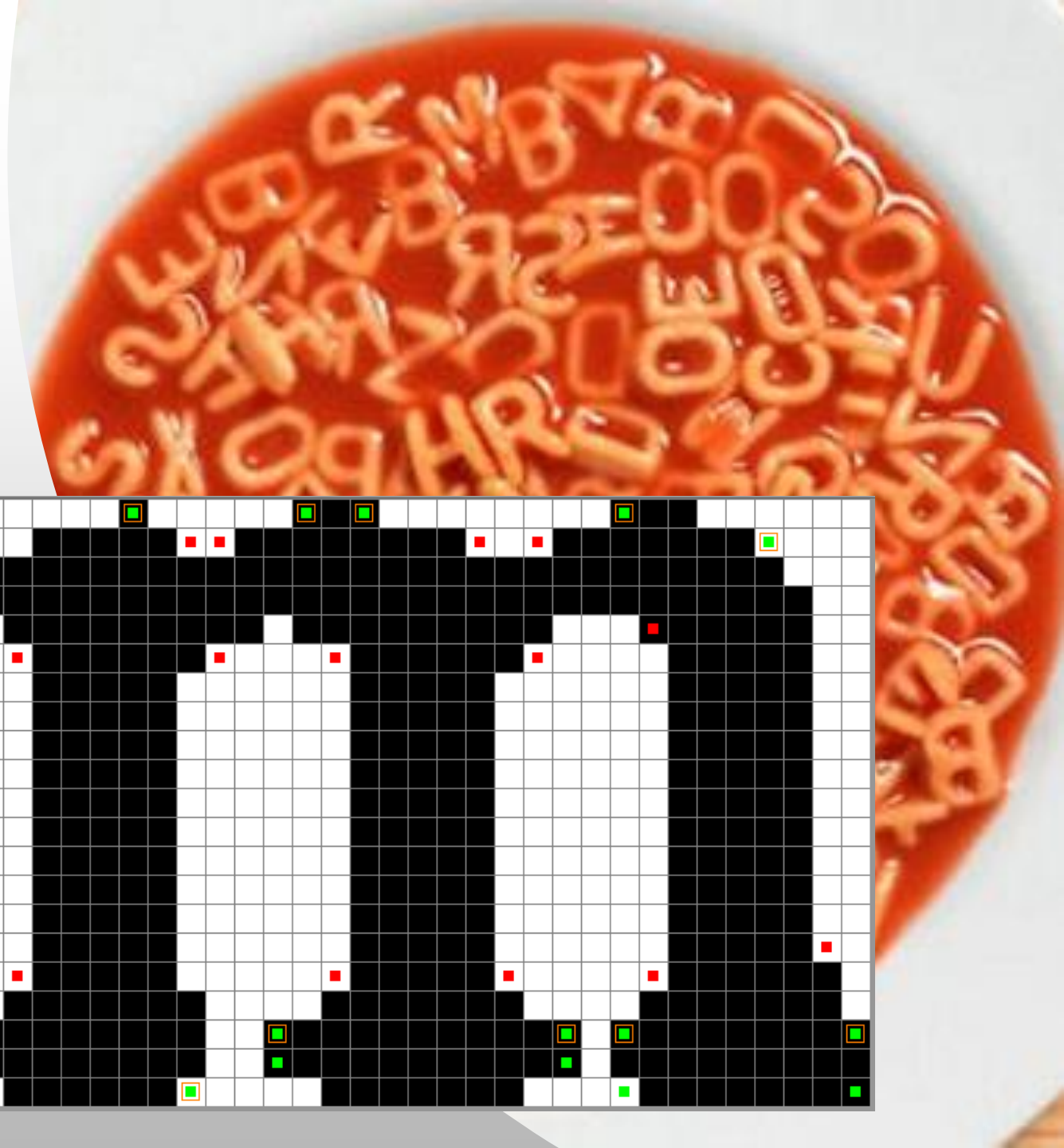
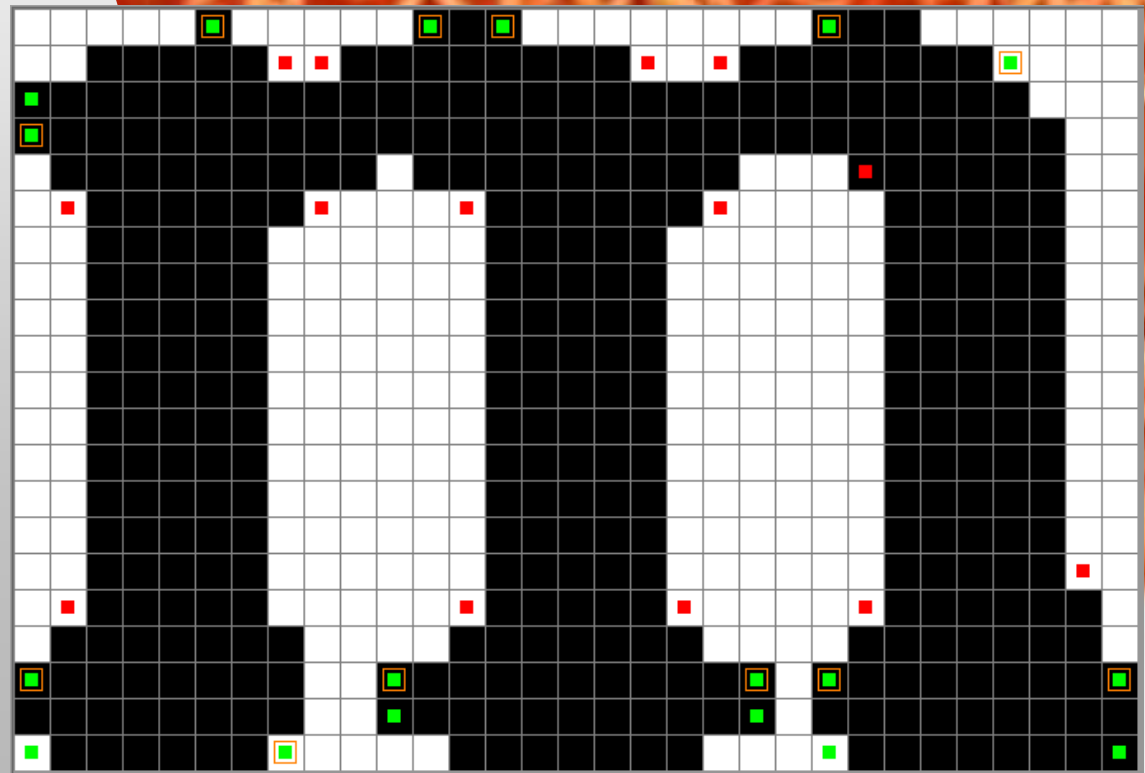
A karakterfelismerés buktatói

- Feature-ök (1985):
 - Konvex (zöld) és konkáv (piros) töréspontok helye a kontúron
 - A pontok száma az elsődleges feature!
- Betanított betűalakok
 - Nincs bedrótozva semmi
 - K-Nearest Neighbors osztályozó
 - Egyszerű tanítás
 - PAT file: címkézett pattern gyűjtemény
 - Tanuló program összevonja a hasonlókat
 - Viszonylag kevés alak is elegendő



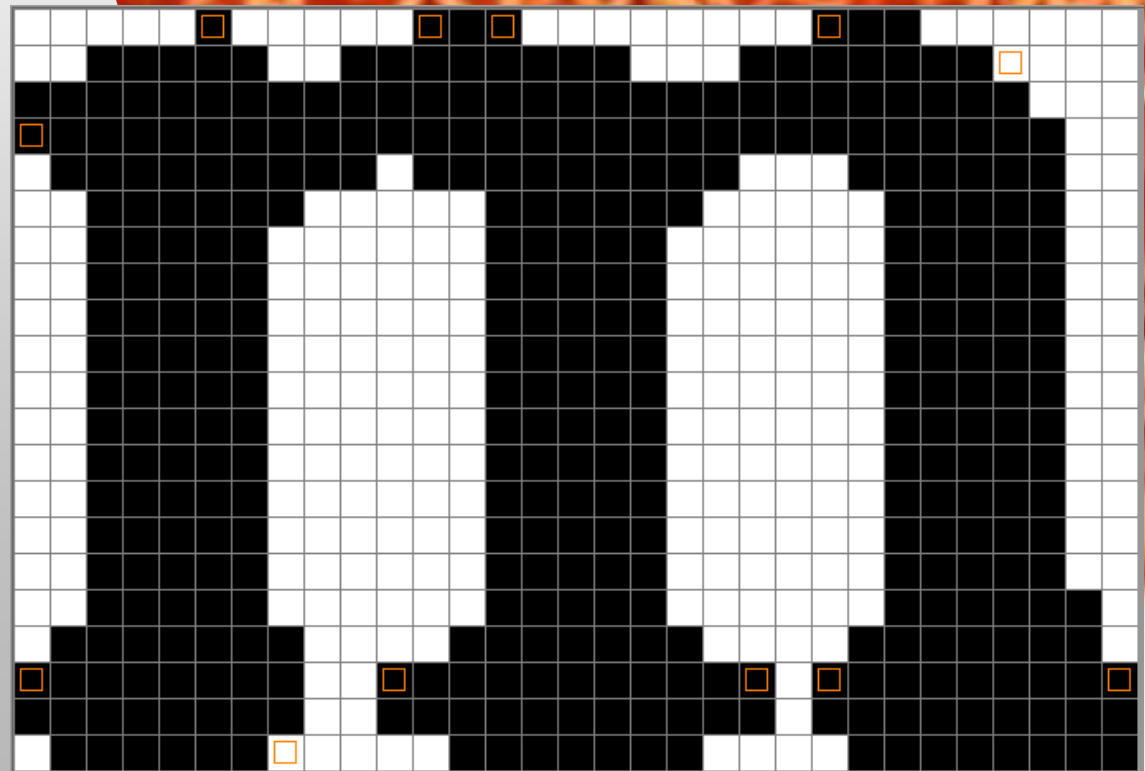
A karakterfelismerés buktatói

- Mennyi a „viszonylag kevés” alak?
 - Néha meglepően sok:
 - Talpak hajlamosak eltűnni
 - Mindig kimarad valamilyen alak
 - Érzékeny pontok, 2ⁿ alak generálás
- Feature-ök 2.0:
 - Töréspontok helyett ívek



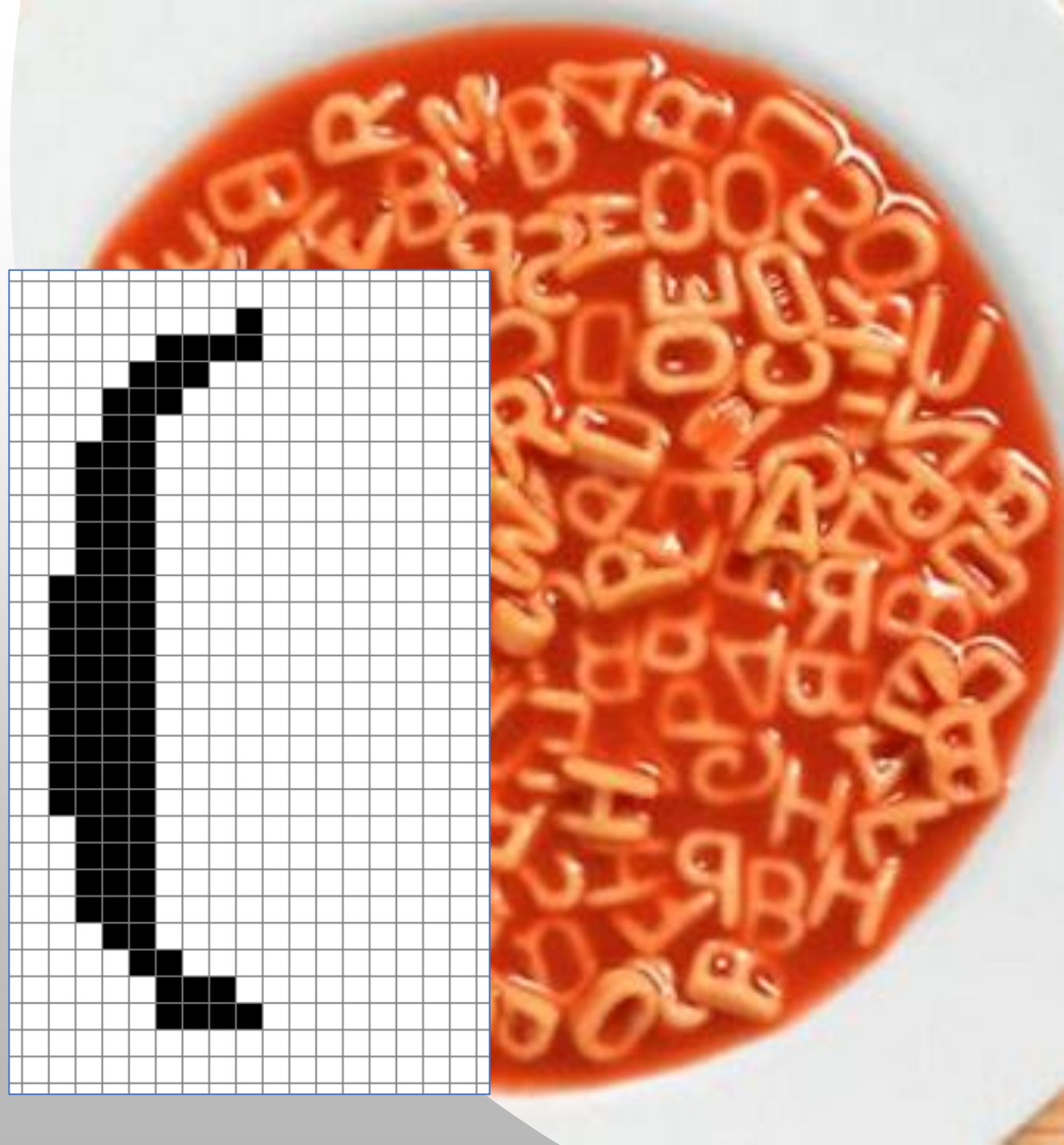
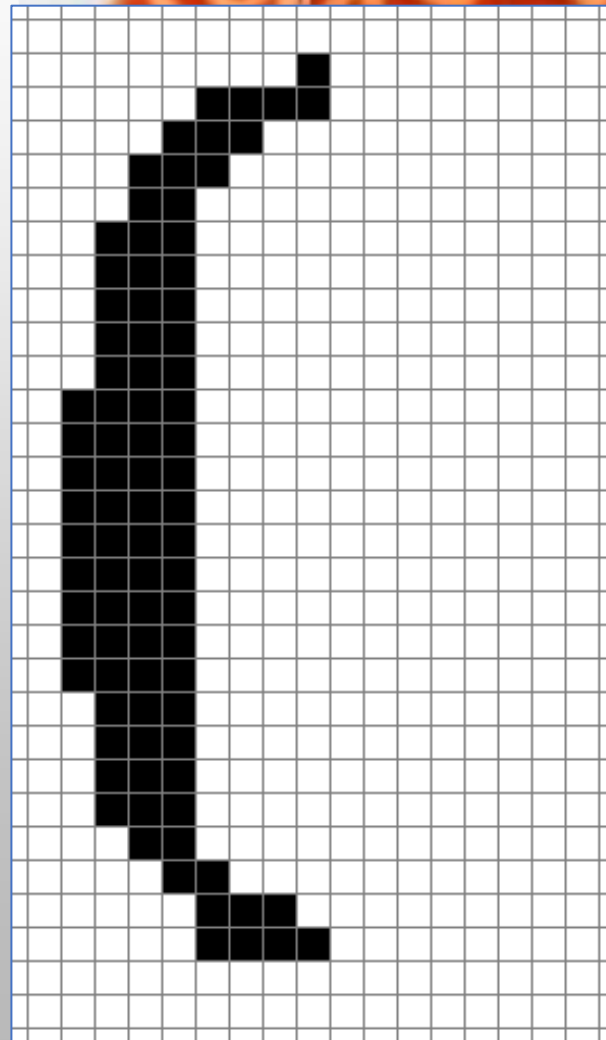
A karakterfelismerés buktatói

- Mennyi a „viszonylag kevés” alak?
 - Néha meglepően sok:
 - Talpak hajlamosak eltűnni
 - Mindig kimarad valamilyen alak
 - Érzékeny pontok, 2ⁿ alak generálás
- Feature-ök 2.0 (1990):
 - Töréspontok helyett ívek
 - Sokkal kevesebb változatosság
 - Viszont gyengébb alak-leírás
 - ISIT rutinok szükségesek



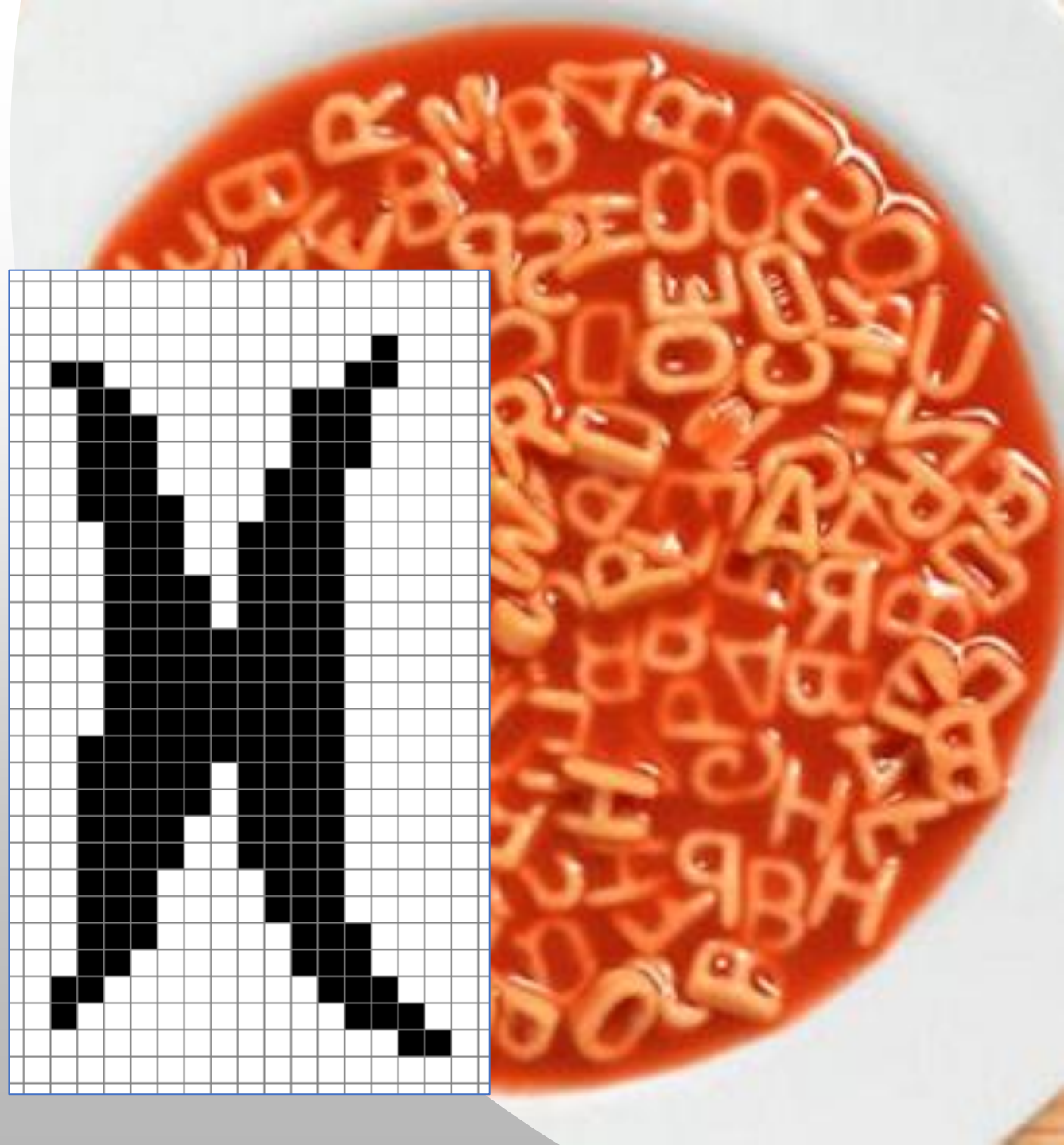
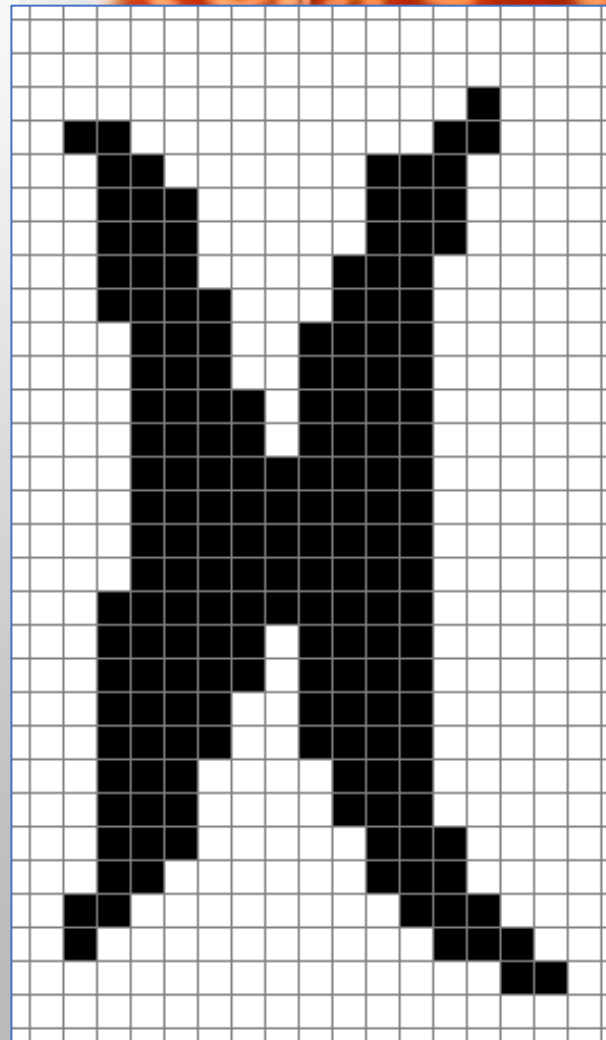
A karakterfelismerés buktatói

- Nagyon egyszerű alakok
 - Más feature is kell hozzájuk



A karakterfelismerés buktatói

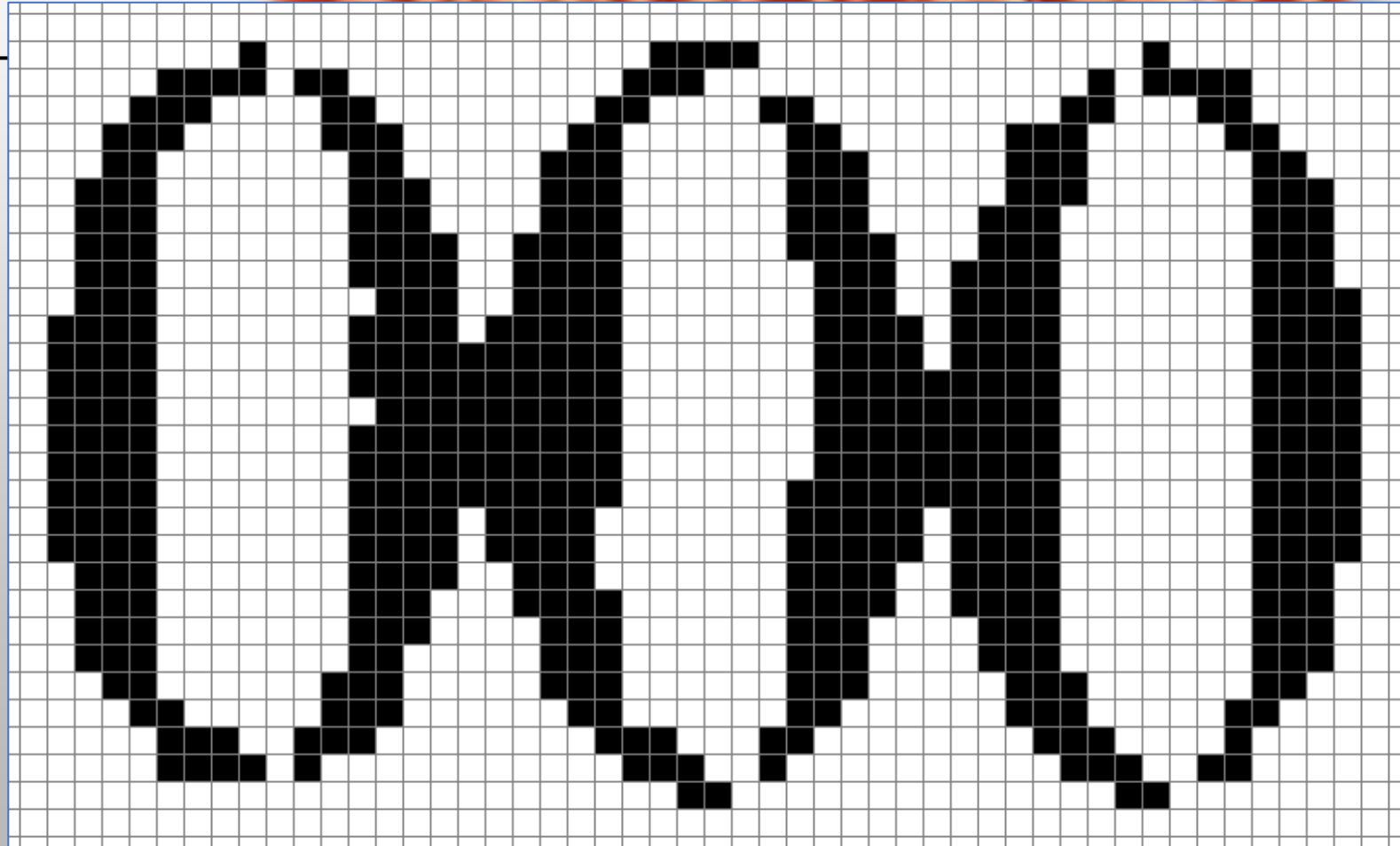
- Nagyon egyszerű alakok
 - Más feature is kell hozzájuk
- A program így látja, pixel-közelből



A karakterfelismerés buktatói



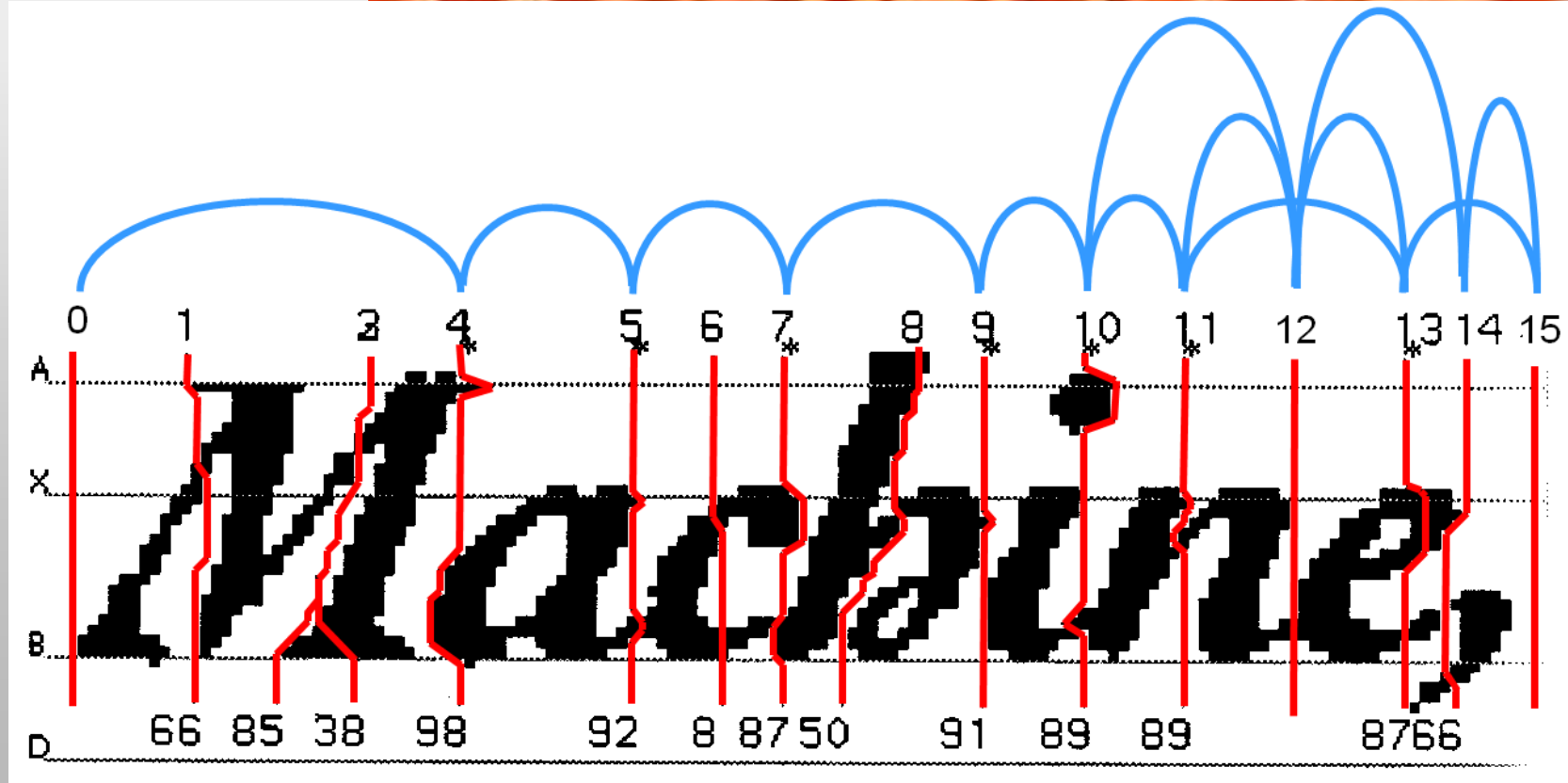
- Karakter szegmentálás
 - A felismerési problémák nagy része szegmentálás alapú!



A karakterfelismerés buktatói

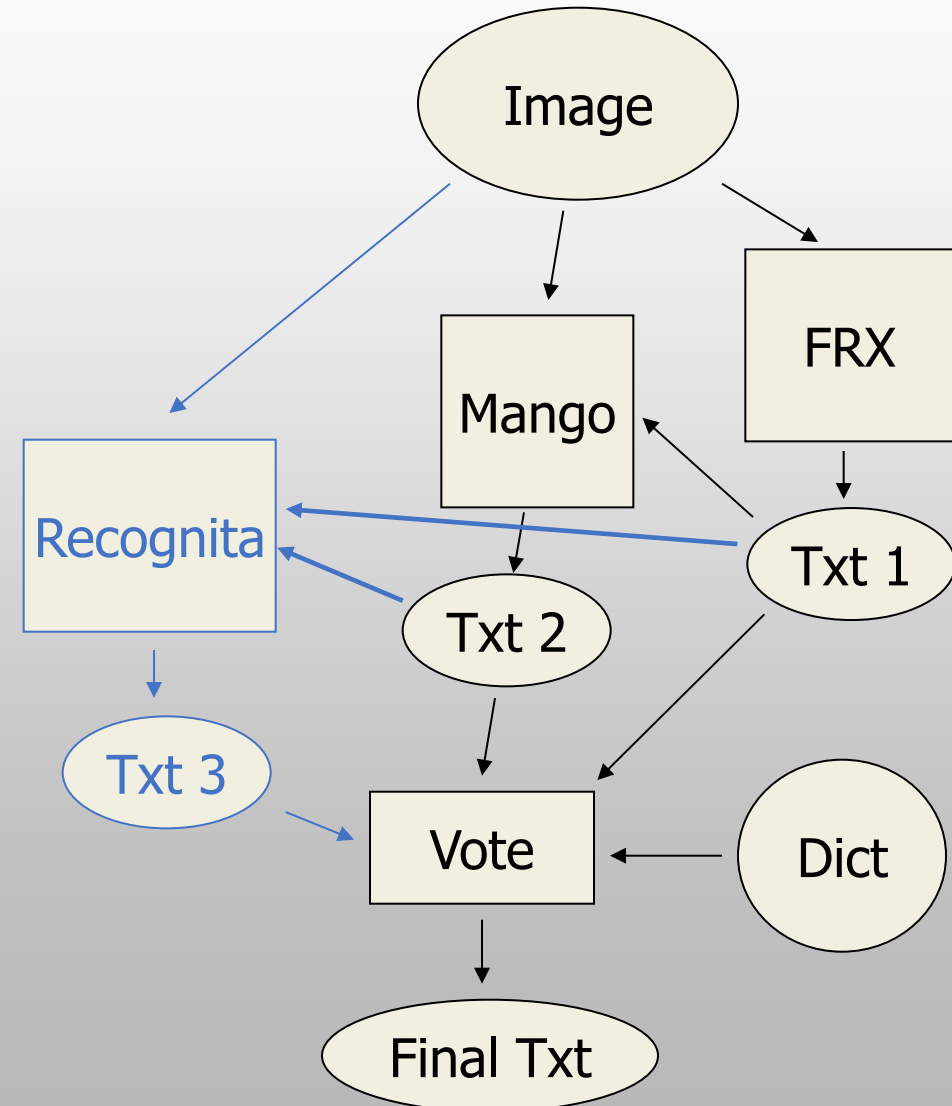


- Karakter szegmentálás
 - A felismerési problémák nagy része szegmentálás alapú!
 - Iteratív processz, sok lehetséges szegmentálást kell kipróbálni, felismerni



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine:
 - 4 nyugati: Recognita, M-Text, Fireworx, Mango
 - Mindegyikben:
 - Elő-tanított osztályozás
 - Szegmentálás
 - Adaptív osztályozás
 - Adaptív szegmentálás
 - Nyelvi analízis
 - Recognita és Mango:
 - Belső vote-olás az algoritmusok mélyén
 - 1 ázsiai: kínai partnertől



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás:
 - Scanner kezelés
 - Képfájl-ok
 - PDF fájl-ok

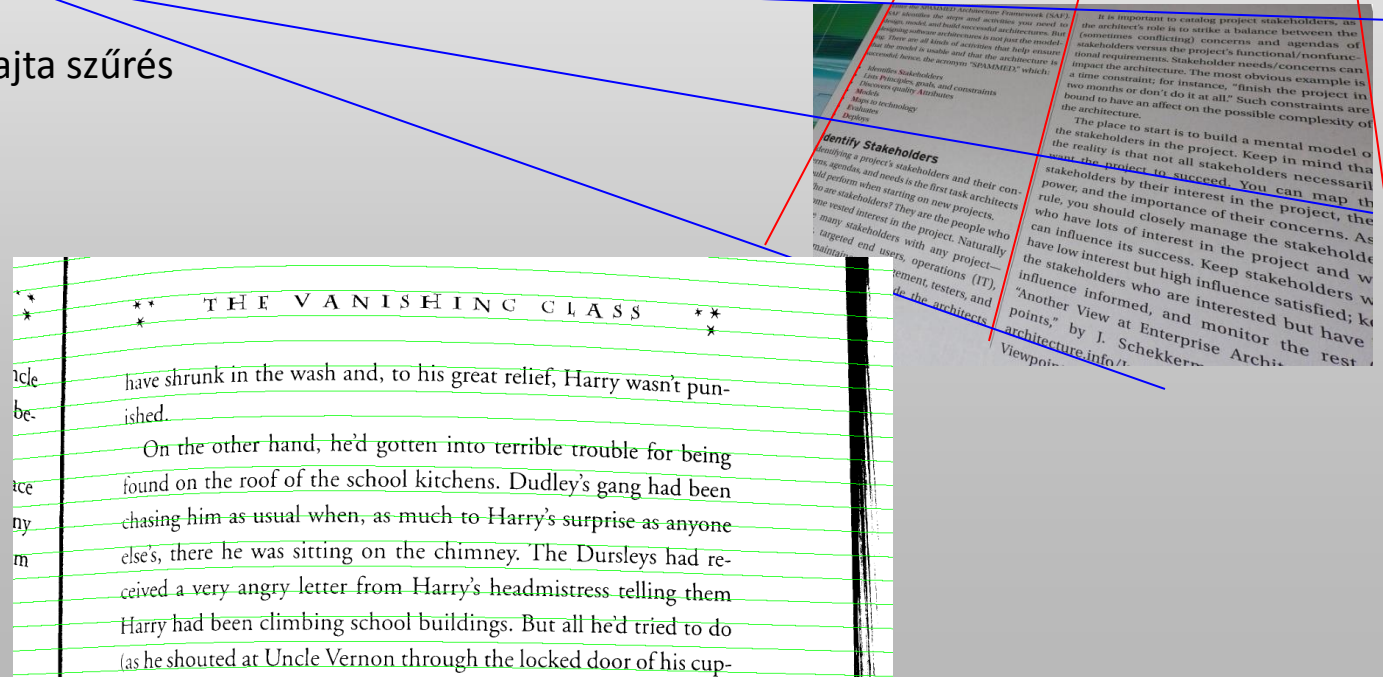


Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás:
 - 2D/3D de-skew, de-speckle, mindenfajta szűrés és bővítkezés
 - Sorkiegyenesítés
 - Binarizálás

Color Separation

De-speckle, de-skew



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás
 - Mindenféle felismerés:
 - Barcode
 - OMR
 - Pontmátrix
 - Kézzel írt számok
 - Nyelv felismerés



YES NO

! " # \$ % & ' () * + , - . / 0
A B C D E F G H I J K L M N O P
Q R S T U V W X Y Z a b c d e f g h i j k l m n
o p q r s t u v w x y z 0 1 2 3 4 5 6 7 8 9

郊区局

تدریبات

Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás
 - Mindenféle felismerés
 - Zónázás:
 - Hol vannak a szövegek, táblázatok, grafikák?

zone 0, Vertical

zone 2, Vertical

zone 6, Graphic

zone 7, Flow

zone 12, Vertical

zone 3, Vertical

zone 8, Vertical

zone 4, Flow

zone 13, Flow

zone 5, Graphic

zone 9, Flow

zone 14, Table

zone 15, Flow

zone 16, Flow

WEDO上海速報 2003.12

今日のおススメ CD

大陸MUSIC

月刊排行榜 2003年10月号

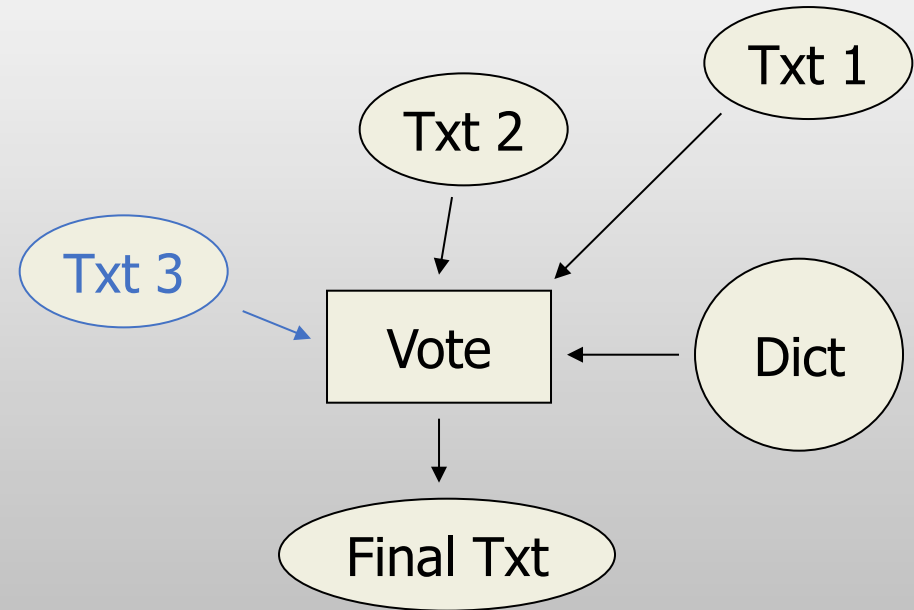
Rank	Song	Artist
1	你的爱给了谁	零点乐队
2	是不是梦	艾敬
3	绽放	李健
4	加速度	瞿颖
5	九月的天	沈丹
6	你记得吗	胡彦斌
7	对岸	金海心
8	天使的诅咒	魏雪漫
9	流浪狗	李泉 feat. 范晓萱
10	风雨彩虹, 铿锵玫瑰	田震

中国音楽航線 -Music Airline- 中国国際広播電台日本語放送

日本向け放送 毎週(土)&(日) 18:55-24:20の間に計6回放送
 AM 1044KHz、短波(SW) 7.190MHz ネットラジオ: <http://online.cri.com.cn/radio4u/japanese.html>
<http://cfioradio.cri.com.cn/japan/airline/index.htm> (ブラボー中国) <http://www.seiryu.com.cn/bclass/>

Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás
 - Mindenféle felismerés
 - Zónázás
 - Vote-olás:
 - Külső vote-olás:
Az engine-ek outputja alapján egy jobb eredmény kialakítása



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás
 - Mindenféle felismerés
 - Zónázás
 - Vote-olás
 - Form-ok kezelése:
 - Form elemek felismerése
 - Kitöltendő mező
 - Check mark
 - „Fésű”
 - Template illesztés

Note: This is a fictitious sample file used to demonstrate the form capabilities of OmniPage and PDF Converter Professional.

PTX RESEARCH

PTX is an affirmative action/equal employment opportunity employer. Discrimination because of race, color, religion, sex, handicap, sexual orientation or national origin is prohibited. In order to be considered for an internship, you must submit a completed application form along with a cover letter and your resume.

Name(s) of Internship(s) Applied For:

1. Internship1 _____ 3. _____
2. Internship2 _____ 4. _____

Contact Information

Name: László Fero _____
Address: Fay Andrea u 20/A _____
City: Póráz _____ State: _____ Zip Code: 2013 _____
Telephone Numbers: (Home) +3620120466 _____ (Mobile) +3620346679 _____
Email Address: laszlo.fero@ptx.com _____

Are you legally eligible to work in the U.S.? Yes No
Are you requesting college credit hours for your internship? Yes No
If Yes, College Name: _____
If you do not receive an internship at PTX, would you be interested in a fee-based program?
 Yes No

Education

TYPE OF SCHOOL	NAME AND LOCATION	DEGREE/DATE	MAJOR
High School	_____	_____	_____
College	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

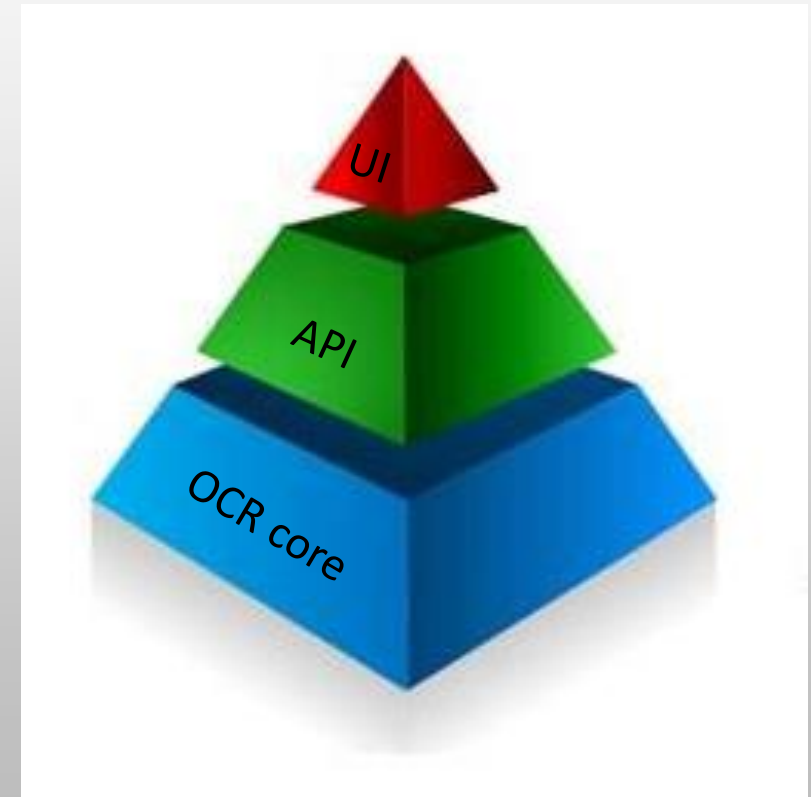
Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás
 - Mindenféle felismerés
 - Zónázás
 - Vote-olás
 - Form-ok kezelése
 - Output dokumentum előállítás:
 - PDF processzálás
 - Több lapos formattált dokumentumok (pl. docx, xlsx, html)



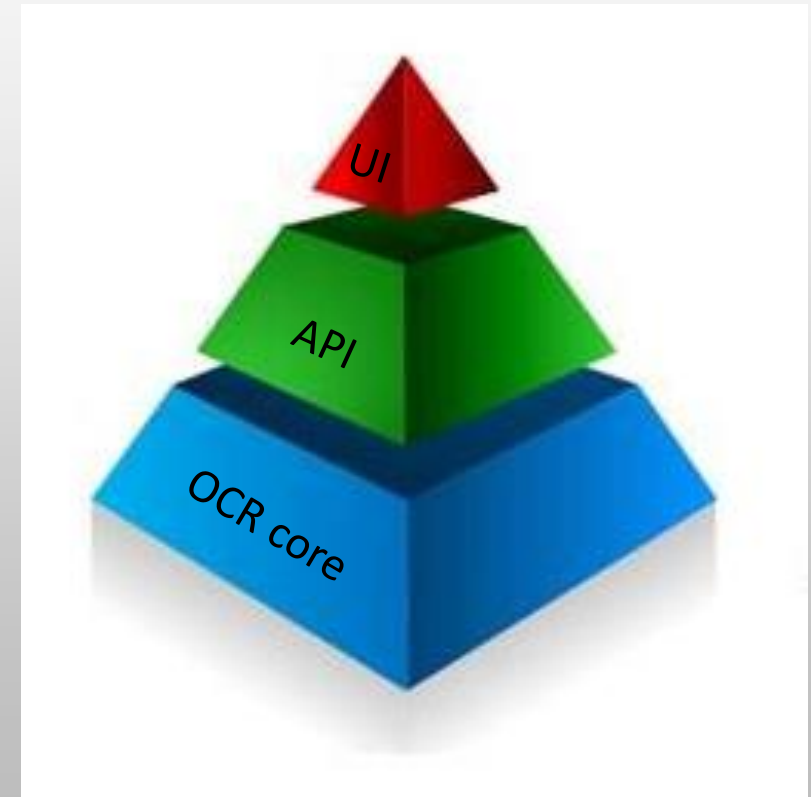
Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás
 - Mindenféle felismerés
 - Zónázás
 - Vote-olás
 - Form-ok kezelése
 - Output dokumentum előállítás
 - Interfészek:
 - Windows UI
 - C, C#, Java API



Az OCR nem csak karakterfelismerés!

- Az osztályozás csak kis része az OCR-nek
- OmniPage: 10.4 millió sor
 - 5 OCR engine
 - Kép beolvasás
 - Kép előfeldolgozás
 - Mindenféle felismerés
 - Zónázás
 - Vote-olás
 - Form-ok kezelése
 - Output dokumentum előállítás
 - Interfészek
 - És még sok más...



Köszönöm a figyelmet



A RECOGNITA ÉS A PDF FÁJLFORMÁTUM

avagy

MI AZ A PDF ÉS MIÉRT SZERETJÜK?

A NEUMANN JÁNOS SZÁMÍTÓGÉP-TUDOMÁNYI TÁRSASÁG
INFORMATIKATÖRTÉNETI FÓRUMA (NJSZT ITF) ÉS
AZ ÓBUDAI EGYETEM

„NAGY SZÁMÍTÁSTECHNIKAI MŰHELYEK” SZOROZATA

A 30 ÉVES RECOGNITA TÖRTÉNETE

2019. 11. 8.

Urbán Zoltán
zoltan.urban@kofax.com

MI A PDF?

- A Portable Document Format (PDF) az Adobe Systems által kifejlesztett nyílt szabványú, dokumentumok tárolására alkalmas fájlformátum, amely alkalmas szöveget, ábrát és képeket tartalmazó dokumentum leírására eszközfüggetlen és felbontásfüggetlen formában.
- A legelterjedtebb technológia a végleges dokumentumok tárolására
 - A megjelenítés hardver-, operációs rendszer és felbontásfüggetlenül, képernyőn és nyomtatásban egyenrangúan eredethű
 - „Csomagformátum”: minden típusú kapcsolódó adat becsatolható a fájlba
 - Ideális formátum papíralapú dokumentumok digitalizálására
 - Az eredeti oldalak szkennelt képe látható
 - A mögöttes szöveg alapján kereshető, katalogizálható

A PDF FÁJL - ALAPISMERETEK

- A PDF fájlok tartalmazhatnak
 - Lapokat, rajtuk
 - Szövegeket
 - Vektoros ábrákat, vonalakat
 - Rasztergrafikákat
 - Annotációkat (kiemelés, szövegbuborék stb.)
 - A szövegekhez tartozó betűkészleteket és karakterkódolási információkat
 - Többé – kevésbé ☺
 - Címkéket (tags)
 - Egyes PDF alszabványokban kötelező
 - Dokumentuminformációkat (metadata)
 - Biztonsági elemeket (titkosítás, digitális aláírás)

A PDF FÁJL - ALAPISMERETEK



rasztergrafika

BECHTEL
REPORT 2006

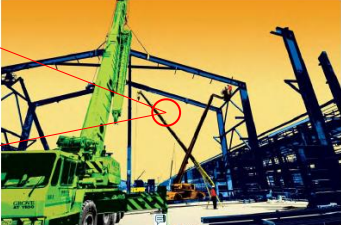
Projects around the world underscored our leadership in the aluminum industry.

Mining & Metals

FERROUS, NONFERROUS, PRECIOUS, AND LIGHT METALS, AS WELL AS INDUSTRIAL MATER

Bechtel's leadership in the aluminum industry manifests once again in 2005. In Bahrain, we completed an expansion of the Alcoa smelter that made it one of the world's largest aluminum production sites, with a capacity of more than 830,000 tonnes per year. The project included an innovative training program for hundreds of local workers, and it set a world record for fastest startup at an aluminum smelter. Based partly on our performance in Bahrain, we won another major project in nearby Oman, where we are providing engineering, procurement and construction management for the new Sohar aluminum smelter.

The work site at the Alcoa smelter in Bahrain.



Safety-First
Number of non-injury incidents per 100 workers per year

Year	Actual	Target
2001	0.14	0.10
2002	0.14	0.10
2003	0.14	0.10
2004	0.14	0.10
2005	0.14	0.10

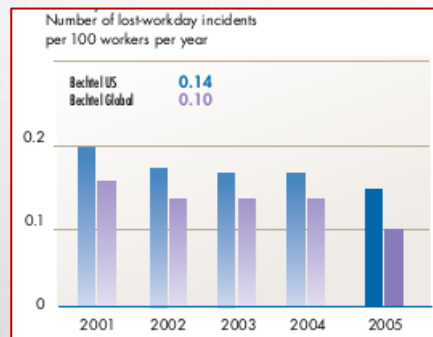
2.2311 The security for the loans shall be single family, owner occupied residences of good design and construction, in good condition, and comparable in value to other homes in the neighborhood.

2.2312 Borrower must have a good credit rating and have adequate income to support the loan.

2.2313 Loans shall be for \$10,000 or more and shall be fully insured by the FIA to the maximum extent permitted under the law.

1 of 1 80%

A PDF FÁJL - ALAPISMERETEK



vektorgrafika


BECHTEL
11-2007 2006

Projects around the world underscored our leadership in the aluminum industry.

Mining & Metals

FERROUS, NONFERROUS, PRECIOUS, AND LIGHT METALS, AS WELL AS INDUSTRIAL MATER

Bechtel's leadership in the aluminum industry was manifest once again in 2005. In Bahrain, we completed an expansion of the Albu smelter that made it one of the world's largest aluminum production sites, with a capacity of more than 830,000 tonnes per year. The project included an innovative training program for hundreds of local workers, and it set a world record for fastest startup at an aluminum smelter. Based partly on our performance in Bahrain, we won another major project in nearby Oman, where we are providing engineering, procurement, and construction management for the new Solar aluminum smelter.



2.2311 The security for the loans shall be single family, owner occupied residences of good design and construction, in good condition, and comparable in value to other homes in the neighborhood.

2.2312 Borrower must have a good credit rating and have adequate income to support the loan.

2.2313 Loans shall be for \$10,000 or more and shall be fully insured by the FHA to the maximum extent permitted under the law.

Stability First
Number of lost-workday incidents per 100 workers per year

Year	Bechtel US	Bechtel Global
2001	0.18	0.15
2002	0.16	0.13
2003	0.15	0.13
2004	0.15	0.13
2005	0.14	0.10

A PDF FÁJL - ALAPISMERETEK

Bechtel's leadership in the aluminum industry was manifest once again in 2005. In Bahrain, we completed an expansion of the Alba smelter that made it one of the world's largest aluminum production sites, with a capacity of more than 830,000 tonnes per year. The project included an innovative training program for hundreds of local workers, and it set a world record for fastest startup at an aluminum smelter. Based partly on our

szöveg

BECHTEL
11.2007-2008

Projects around the world underscored our leadership in the aluminum industry.

Mining & Metals

FERROUS, NONFERROUS, PRECIOUS, AND LIGHT METALS, AS WELL AS INDUSTRIAL MATER

Bechtel's leadership in the aluminum industry was manifest once again in 2005. In Bahrain, we completed an expansion of the Alba smelter that made it one of the world's largest aluminum production sites, with a capacity of more than 830,000 tonnes per year. The project included an innovative training program for hundreds of local workers, and it set a world record for fastest startup at an aluminum smelter. Based partly on our performance in Bahrain, we won another major project in nearby Oman, where we are providing engineering, procurement, and construction management for the new Solar aluminum smelter.

© 2007 Bechtel Corporation. All rights reserved. Bechtel is a registered trademark of Bechtel Corporation.


2.2311 The security for the loans shall be single family, owner occupied residences of good design and construction, in good condition, and comparable in value to other homes in the neighborhood.

2.2312 Borrower must have a good credit rating and have adequate income to support the loan.

2.2313 Loans shall be for \$10,000 or more and shall be fully insured by the FHA to the maximum extent permitted under the law.

Safety First
Number of lost-workday incidents per 100 workers per year

Year	Number of lost-workday incidents per 100 workers per year
2001	0.14
2002	0.10
2003	0.10
2004	0.10
2005	0.10



A PDF FÁJL - ALAPISMERETEK

- 2.2311 The security for the loans shall be single family, owner occupied residences of good design and construction, in good condition, and comparable in value to other homes in the neighborhood.
- 2.2312 Borrower must have a good credit rating and have adequate income to support the loan.
- 2.2313 Loans shall be for \$10,000 or more and shall be fully insured by the FHA to the maximum extent permitted under the law.

Szkennelt (raszteres) szöveg


BECHTEL
11/2007 2008

Projects around the world underscored our leadership in the aluminum industry.

Mining & Metals

FERROUS, NONFERROUS, PRECIOUS, AND LIGHT METALS, AS WELL AS INDUSTRIAL MATER

Bechtel's leadership in the aluminum industry was manifest once again in 2005. In Bahrain, we completed an expansion of the Albu smelter that made it one of the world's largest aluminum production sites, with a capacity of more than 830,000 tonnes per year. The project included an innovative training program for hundreds of local workers, and it set a world record for fastest startup at an aluminum smelter. Based partly on our performance in Bahrain, we won another major project in nearby Oman, where we are providing engineering, procurement, and construction management for the new Salas aluminum smelter.



2.2311 The security for the loans shall be single family, owner occupied residences of good design and construction, in good condition, and comparable in value to other homes in the neighborhood.

2.2312 Borrower must have a good credit rating and have adequate income to support the loan.

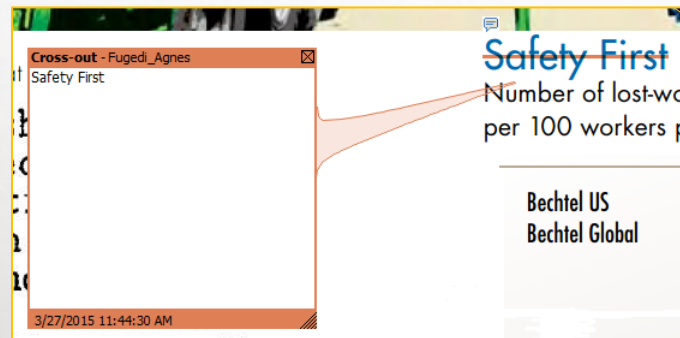
2.2313 Loans shall be for \$10,000 or more and shall be fully insured by the FHA to the maximum extent permitted under the law.

Safety First
Number of lost-workday incidents per 100 workers per year

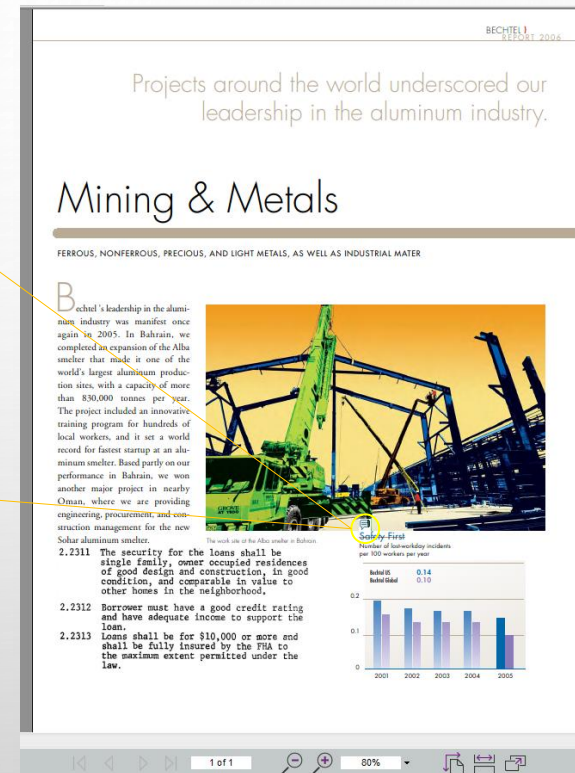
Year	Number of lost-workday incidents per 100 workers per year
2001	0.18
2002	0.15
2003	0.12
2004	0.10
2005	0.08

1 of 1 80%

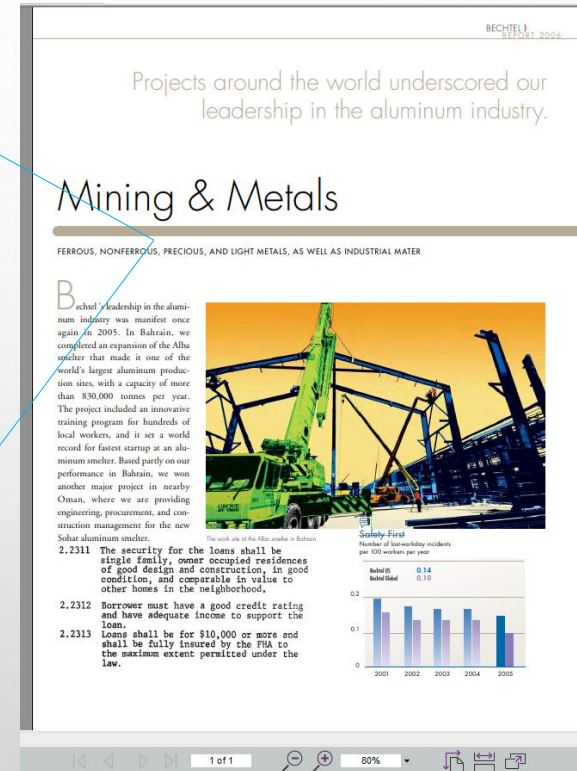
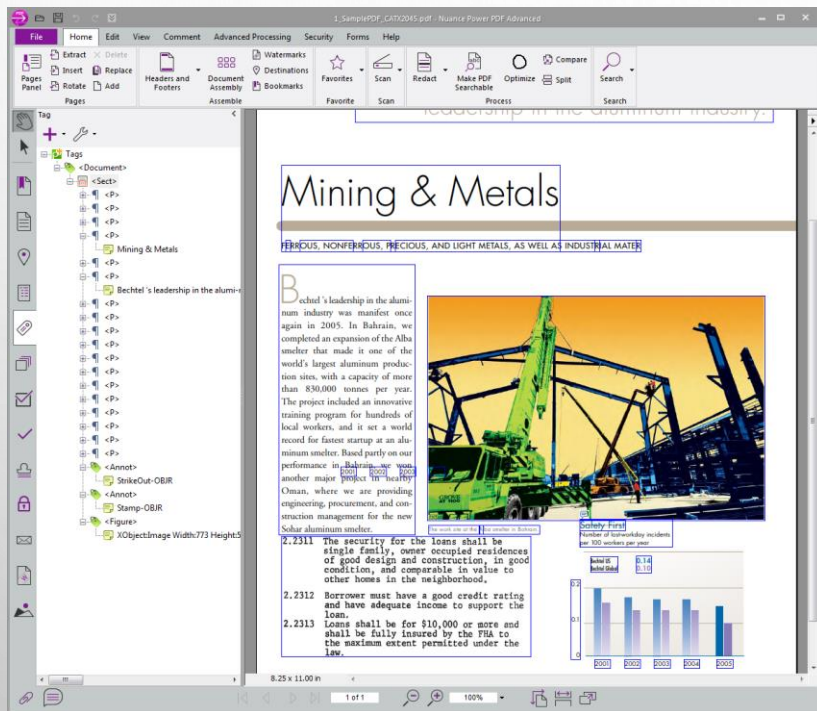
A PDF FÁJL - ALAPISMERETEK



Annotáció - áthúzás



A PDF FÁJL - ALAPISMERETEK



Címkék

HOGY KERÜL A (PDF) CSIZMA AZ (OCR) ASZTALRA?

- PDF fájlok kezelésénél felmerülő feladatok
 - PDF készítése
 - Leggyakrabban „nyomtatás” útján
 - Papíralapú dokumentumok digitalizálása
 - PDF feldolgozása
 - Lapok törlése, beszúrása, átrendezése
 - Szöveg kinyerése
 - Gépi feldolgozás számára (pl. számlák)
 - Bemenet gyengénlátók felolvasó szoftvereinek
 - Redaktálás
 - Konvertálás (MS Word DOCX, ePub stb.)
 - „végleges” ↔ szerkeszthető, tördelhető formátum
 - Címkézés (PDF/A-1a, PDF/UA)



PAPÍRALAPÚ DOKUMENTUMOK DIGITALIZÁLÁSA

IMAGE-ON-TEXT PDF FÁJLOK

- Az egyes oldalakon a szkennelt képek látszanak
- A képek „mögött” nem látható módon az OCR által felismert szöveg pozícióhelyesen
- A hagyományos keresés a PDF-kezelő alkalmazásokban a szokott módon működik, mintha nem is a szkennelt képet látnánk
- Másolás / beillesztés is alkalmazható a szöveges részek kiexportálására
- További lehetőség: a szkennelt képek optimalizálása (MRC technológia)
 - Sokkal kisebb méret komoly minőségromlás nélkül
 - Alapja a szöveges és grafikus területek különválasztása (képfeldolgozás!)
 - A szöveges részek nagyobb felbontásban de monokromatikusan tárolva (jobb tömöríthetőség)
 - A grafikák kisebb felbontásban, veszteségesen tömörítve

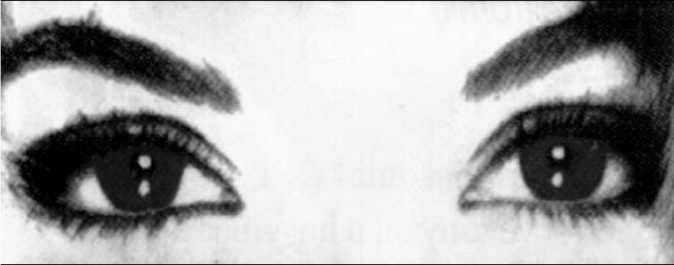
SZÖVEGKINYERÉS, KONVERTÁLÁS, CÍMKÉZÉS

- Nem minden szöveg, ami annak látszik (pl. rasztergrafika feljebb)
- A PDF fájlok megjelenítéséhez nem szükséges a szövegkarakterek azonosíthatósága
- A PDF fájlok megjelenítéséhez nem szükséges a szövegrészek sorrendjének az ismerete
- A PDF fájlok nem kell, hogy tartalmazzanak logikai struktúrát
 - bekezdés, fejléc, lábléc, táblázatok, hasábok stb.
 - konvertálásnál, címkézéskor ezeket kell megtalálni és visszaadni a kimeneten
- A fentiek kezeléséhez szükség van képfeldolgozási és felismerési algoritmusokra!

GYANUS.PDF

File | C:/Users/UZ/Desktop/GYANUS.PDF

LÉZER KEZELÉS
UTÁN NEM KELL SZEMÜVEG



**Gyors, pontos,
fájdalommentes
lézersugárzás után
végleges eredmény**

Szt. István krt. 24.
Tel.: 340-4390

*Untitled - Notepad

File Edit Format View Help

```
#$%&'( ))*!"*  
(+,*-.$,$-"$!  
0+.(*.* **&)+!  
2,"3 5"0$! 7"3 9:3  
;*(3< >: ?@:>A?|
```

100% Windows (CRLF) UTF-8

SZÖVEGKINYERÉS, KONVERTÁLÁS, CÍMKÉZÉS

- Nem minden szöveg, ami annak látszik (pl. rasztergrafika feljebb)
- A PDF fájlok megjelenítéséhez nem szükséges a szövegkarakterek azonosíthatósága
- A PDF fájlok megjelenítéséhez nem szükséges a szövegrészek sorrendjének az ismerete
- A PDF fájlok nem kell, hogy tartalmazzanak logikai struktúrát
 - bekezdés, fejléc, lábléc, táblázatok, hasábok stb.
 - konvertálásnál, címkézéskor ezeket kell megtalálni és visszaadni a kimeneten
- A fentiek kezeléséhez szükség van képfeldolgozási és felismerési algoritmusokra!

KÖSZÖNÖM A FIGYELMET!