

Pojednostavljenje ulaznih procesa u grafičkim poduzećima

Hanževački, Vedran

Undergraduate thesis / Završni rad

2016

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Graphic Arts / Sveučilište u Zagrebu, Grafički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:216:225059>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-11-15**



Repository / Repozitorij:

[Faculty of Graphic Arts Repository](#)



SVEUČILIŠTE U ZAGREBU
GRAFIČKI FAKULTET

ZAVRŠNI RAD

Vedran Hanževački



Sveučilište u Zagrebu
Grafički fakultet

Smjer: Tehničko - tehnološki

ZAVRŠNI RAD

POJEDNOSTAVLJENJE ULAZNIH PROCESA U GRAFIČKIM PODUZEĆIMA

Mentor:

Izv. prof. dr. sc. Dubravko Banić

Student:

Vedran Hanževački

Zagreb, 2016

Sažetak

Ovim radom pojašnjavaju se procesi koji utječu na automatizaciju, odnosno pojednostavljenje ulaznih procesa za neko grafičko poduzeće.

Objašnjavanjem i razradom pojmova kao što su indeksiranje, strojno učenje i slično, dolazi se do eksperimentalnog dijela koje prikazuje primjenu jedne razine automatizacije korištenjem programa za elektroničku poštu.

Koristi se takav primjer jer je elektronička pošta danas glavni ulazni proces kod mnogih poduzeća. Strojnim učenjem i umjetnom inteligencijom se takav jednostavni sustav može dovesti na mnogo višu i napredniju razinu, koja se već danas koristi i na taj način strojevi zamjenjuju čovjeka u mnogim područjima.

Strojnim prevođenjem se uklanjaju jezične barijere te su globalna komunikacija i poslovanje znatno olakšani.

Međusobnom integracijom takvih metoda i sustava, dobiva se napredni sustav koji može obavljati čovjeku nedostižne radnje u iznimno kratkom vremenskom periodu.

Ključne riječi: automatizacija, elektronička pošta, sustav, grafičko poduzeće

Abstract

This paper explains the processes which affect the automation and simplification of the input processes of a graphic company. After explaining and elaborating of concepts such as indexing, machine learning, etc., comes the experimental part which shows one level of automation using the email client.

Reason why this example is used is because many companies use email as their main input process. Such a simple system can achieve a much higher and more advanced level by using machine learning and artificial intelligence.

Similar systems are already used today and are replacing humans in many fields of work. Machine translation removes language barriers and makes the global communication and commerce much easier.

With mutual integration of such methods and systems, very advanced systems can be achieved that can perform actions unachievable to humans in an extremely short period of time.

Keywords: automation, electronic mail, system, graphic company

Sadržaj

1. Uvod	1
2. Grafička industrija	3
3. Komunikacijski sustav	5
3.1. Grafički komunikacijski proces	6
4. Automatizacija	8
4.1. Automatizacija radnog procesa	10
5. Indeksiranje	13
5.1. Ručno indeksiranje	13
5.2. Poluautomatsko indeksiranje	14
5.3. Automatsko indeksiranje	14
5.4. Vrste automatskog indeksiranja	15
5.4.1. Metoda ekstrakcije.....	15
5.4.2. Metoda dodjeljivanja ključnih riječi	15
5.4.3. Lingvističke i statističke metode	16
6. Strojno učenje	17
6.1. Nadgledano učenje	18
6.2. Nenadgledano učenje	19
7. Programi	20
7.1. Gmail	20
7.2. E-mail klijent (Thunderbird)	22
7. Strojno prevođenje	24
7.1. Princip rada	24
7.2. Problemi	27
8. Eksperimentalni dio	28
8.1. Definiranje problema	28
8.2. Konfiguracija programa	29
8.3. Rezultati i rasprava	31
9. Zaključak	32
10. Literatura	34

1. Uvod

Već od davnina se očituju počeci grafičke industrije i njezinih proizvoda. Čovjek je oduvijek težio iskazivanju svojih zamisli, stavova, mišljenja, stoga je u početku počeo pisati po kamenu (špilje), a kasnije i na kamenim pločama i drvetu s namjerom da se njegova misao širi. Zbog te želje za prenošenjem znanja, nastaju i metode za umnožavanje pisanih djela. U Kini se već od 3. st. koriste metode umnožavanja pomoću drvenih slova – drvorez.

Možemo reći da se grafička industrija razvija već tisućama godina, dok je najveći i najbrži razvoj doživjela u zadnjih nekoliko stoljeća.

Danas, kada postoji mnogo različitih tehnika tiska, kada su naklade velike i konstantno potrebne, a vrijeme dragocjeno, prijeko je potrebno usmjeriti razvoj tiskarstva ka uštedi vremena, optimizaciji, integraciji, automatizaciji i standardizaciji. Sve zajedno će omogućiti velike uštede vremena i novca.

Već i sada je automatizacija na prilično visokoj razini i konstantno se unaprjeđuje. Tiskarski strojevi gotovo da i ne trebaju ljudsku ruku (osim održavanja ili uklanjanja mogućih kvarova), informacije nesmetano putuju unutar grafičkog poduzeća – priprema, tisak, dorada te između ostalih komponenata sustava.

Svaka informacija ima svoji početak i kraj. Ideja potencijalnog klijenta početak je informacije. Ta ideja razradom prerasta u koncept i digitalnu verziju te moguću maketu. Kako bi se ideja ostvarila potrebno je grafičko poduzeće koje će ju realizirati i umnožiti.

Informacija o ideji danas najčešće dolazi u poduzeće elektroničkom porukom koju određena osoba mora pregledati i preusmjeriti drugoj osobi koja ima znanja iz traženog područja te je zadužena za taj posao. Za to je potrebno vrijeme koje se sastoji od čitanja svake pošte te preusmjeravanja istih drugim zaposlenicima unutar poduzeća.

Pošto je vrijeme danas vrlo dragocjeno, za takve poslove može se koristiti sustav koji će automatski pretražiti sadržaj svake pristigle poruke po definiranim ključnim riječima i po njima zaključiti o kakvom je poslu riječ te poštu proslijediti nadležnoj osobi.

Sustav može biti u većoj ili manjoj ovisnosti o čovjeku te će ovisno o tome funkcionirati na različite načine te s različitim karakteristikama.

2. Grafička industrija

Grafička industrija je grana industrije koja se bavi izradom grafičkih proizvoda kao što su novine i časopisi, knjige i brošure, agitacijski i reklamni materijal (plakati, katalogi), obrasci, trgovačke knjige i poslovne tiskanice, kalendari, akcidencije (posjetnice, čestitke), posebni grafički proizvodi (novčanice, čekovi), grafička galanterija (bilježnice, mape, albumi). Uz sami tisak, u tu granu ulaze i knjigoveštvo te prerada papira zajedno s proizvodnjom ambalaže od kartona i ljepenke (kartonaža).

Glavni elementi proizvodnoga procesa su grafička priprema, tisak i dorada [1].

Grafička tehnologija je djelatnost koja se bavi postupcima oblikovanja, reprodukcije i umnožavanja teksta i ilustracija, tj. grafičkom pripremom, tiskom i grafičkom doradom. Postupcima prethodi izrada snimke, tj. definiranje osnovnog izgleda proizvoda i tehnologije kojom će se on izraditi. Snimka je osnova za procjenu potrebnih materijala, proračun troškova te oblikovanje grafičkoga proizvoda tijekom grafičke pripreme.

Grafička priprema obuhvaća obradu i oblikovanje teksta (izradu tiskarskoga sloga) i ilustracija (reprofotografija), njihovo povezivanje u cjelinu (prijelom) i izradu tiskovne forme. Danas su postupci grafičke pripreme u potpunosti računalno podržani i razvojem stolnog izdavaštva u velikoj mjeri spojeni, a neki se grafički proizvodi pojavljuju samo kao elektronička izdanja, krajnjem korisniku izravno dostupna na zaslonu njegova računala, s mogućnošću daljnje obrade ili ispisa pisačem. No većina se grafičkih proizvoda i dalje izrađuje tradicionalnim tehnikama tiska, pri kojima je u završnom dijelu grafičke pripreme potrebno izraditi tiskovnu formu, tj. matricu s tiskovnim elementima (koji će tijekom tiska nositi tiskarsku boju) i slobodnim površinama [2].

Konačni izgled grafički proizvod poprima tijekom grafičke dorade, postupcima knjigoveštva ili proizvodnje ambalaže [2].

Kao što se vidi iz navedenog, obujam grafičke industrije je veliki, odnosno, postoji mnogo proizvoda koji iz nje izlaze. Tako jedna tiskara ili knjigovežnica može izrađivati mnogo različitih grafičkih proizvoda, a zahtjevi mogu biti česti i raznoliki što u konačnici može oduzimati mnogo vremena za praćenje, čitanje, filtriranje i prosljeđivanje zahtjeva.

Razvojem sustava koji bi u velikoj mjeri taj posao odrađivao sam, uštedilo bi se mnogo vremena i novac koji bi se mogli iskoristiti na bolji način.

To se može ostvariti ako su zadovoljeni određeni uvjeti poput kvalitetnog komunikacijskog sustava, odnosno definiranog komunikacijskog procesa, a u velikoj mjeri sustav bi bio efektivniji ukoliko bi postojala automatizacija radnog procesa.

3. Komunikacijski sustav

Komunikacija je proces razmjene informacija preko dogovorenog sustava znakova, odnosno to je proces slanja informacija sebi ili bilo kojem drugom entitetu, najčešće putem jezika. Riječ komunikacija doslovno znači: podijeliti, učiniti nešto općim ili zajedničkim. Komunikacija je obično opisana prema 3 glavne dimenzije: sadržaju, formi i cilju. Sadržaj komunikacije i forma kreiraju poruke koje se šalju prema cilju. Cilj može biti sam čovjek, druga osoba (u interpersonalnoj komunikaciji) ili drugi entitet poput grupe, organizacije ili društva. [3]

Slično kao i u mnogim srodnim znanostima, postoji mnogo različitih definicija komunikacije i još uvijek nema jedne koju bi prihvatila većina znanstvenika koji se bave ovim znanstvenim poljem. [3]

Harold Lasswell, američki politolog i komunikolog, opisao je njene sastavnice poznatom, tzv. Lasswellovom formulom: 'Tko, kaže što, kojim kanalom, kome i s kojim efektom?'. [3]

Katherine Miller u knjizi *Communication Theories: Perspectives, Processes, and Contexts* daje pregled čak sedamnaest različitih definicija komunikacije kroz povijest: Weaver, npr. kaže da komunikacija 'uključuje sve procedure kojima ljudi utječu jedni na druge', a Hawes komunikaciju naziva 'strukturiranim ponašanjem u prostor-vremenu sa simboličkom odrednicom'. [3]

U svima njima komunikacija se promatra ili kao proces ili kao stvaranje značenja ili kao prijenos određene informacije ili poruke, a vrlo često i kao kombinacija dvaju pogleda, ili čak kao sve troje zajedno. [3]

Kod teorije Harold Lasswell-a očituje se problem koji može biti podosta velik u svakoj industriji, odnosno u bilo kojem poslovanju. Isto je i u grafičkoj industriji.

Način komuniciranja te korištenje izraza iz grafičke struke vrlo je bitan faktor za kvalitetnu komunikaciju između klijenta i grafičkog poduzeća. Ako taj faktor nije zadovoljen, komunikacija neće biti kvalitetno uspostavljena i krajnji rezultat može biti nezadovoljavajuć grafički proizvod.

Korištenjem sustava s definiranim ključnim riječima koje se konstantno nadopunjavaju te mogućnošću prevođenja sa stranog jezika uvelike se umanjuje mogućnost pogreške ili pogrešnog tumačenja poruke.

Često korišten kanal u ovom slučaju je elektronička pošta koja služi kao posrednik u komunikaciji, a koja bi bila osnova spomenutog sustava.

3.1. Grafički komunikacijski proces

Grafička djelatnost usmjerena je komuniciranju, a osnovno sredstvo komuniciranja je jezik, tj. pismo. Grafički komunikacijski proces sastoji se od:

1. pošiljatelja (grafički proizvod),
2. kanala (udaljenost između pošiljatelja i primatelja) i
3. primatelja (čitatelj grafičkog proizvoda)

U trenutku čitanja informacija se prenosi od pošiljatelja (grafičkog proizvoda) kroz kanal do primatelja (čitatelja). Da bi se neka informacija mogla prenijeti kroz kanal ona mora biti kodirana, odnosno misli moraju biti pretvorene u slovne znakove. Prilikom čitanja dolazi do dekodiranja informacije, tj. slovni znakovi pretvaraju se ponovno u misli.

Za uspješno sporazumijevanje potrebno je poznavati kod tj. pismo i jezik.

Cilj komunikacije je što razumljivije prenošenje informacije u komunikacijskom procesu. Zbog toga su moguća poboljšanja elemenata u grafičkom komunikacijskom procesu, i to:

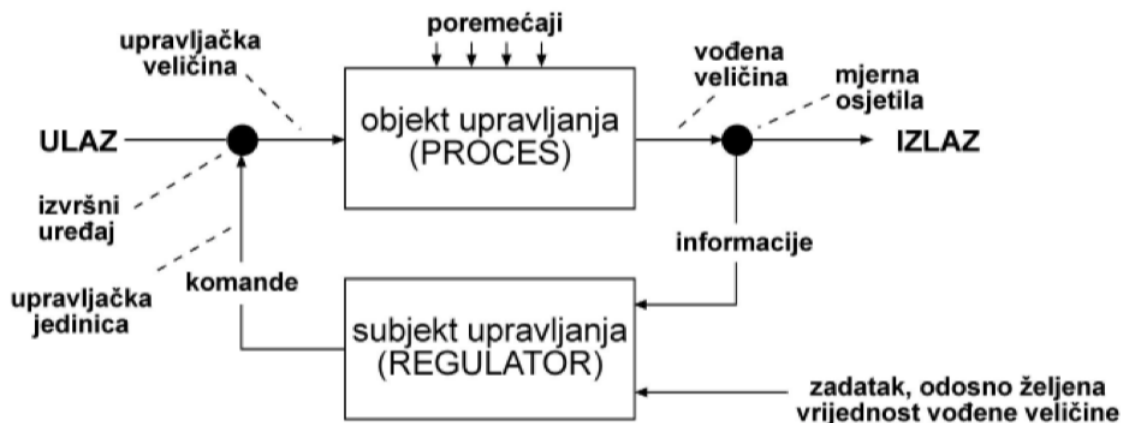
1. pošiljatelja (dobar grafički dizajn, tipografija)
2. kanala (vidni kut, rasvjeta)
3. primatelja (obrazovanje, iskustvo, intelekt)

Komunikacijski proces sastavni je dio automatskog procesa koji može biti upravljan i reguliran.

Upravljeni proces je proces u kojemu uređaj za vođenje automatski usmjerava djelovanje procesa, a da pri tome ne dobiva povratne informacije o svome djelovanju. Zakon vođenja postavlja se na temelju odabranog cilja za održavanje zadanih uvjeta.

Regulirani proces koristi se zbog normalnog funkcioniranja sustava. Regulacija predstavlja način vođenja procesa gdje se odabrana procesna veličina održava na stalnoj vrijednosti ili se mijenja na zadani način, unatoč promjeni uvjeta djelovanja procesa zbog nepredvidivih utjecaja.

Izlaz procesa sa regulatorom je povezan s izvršnim uređajem (Slika 1).

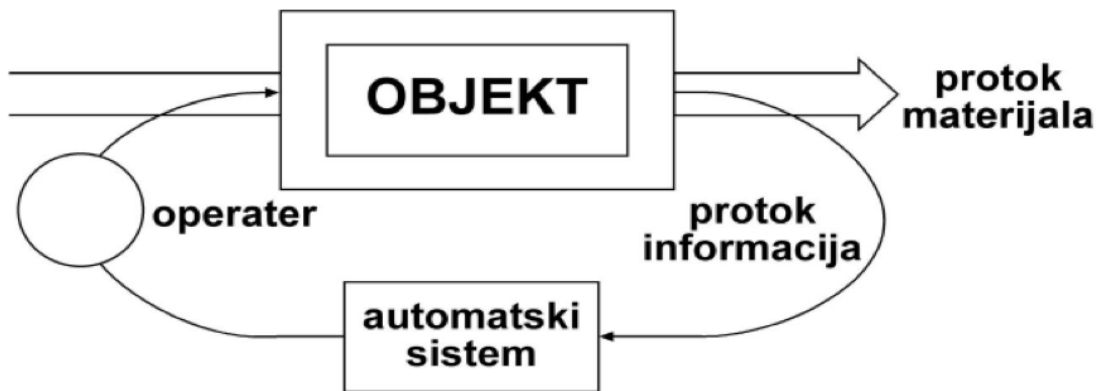


Slika 1. Regulacijski krug [5]

4. Automatizacija

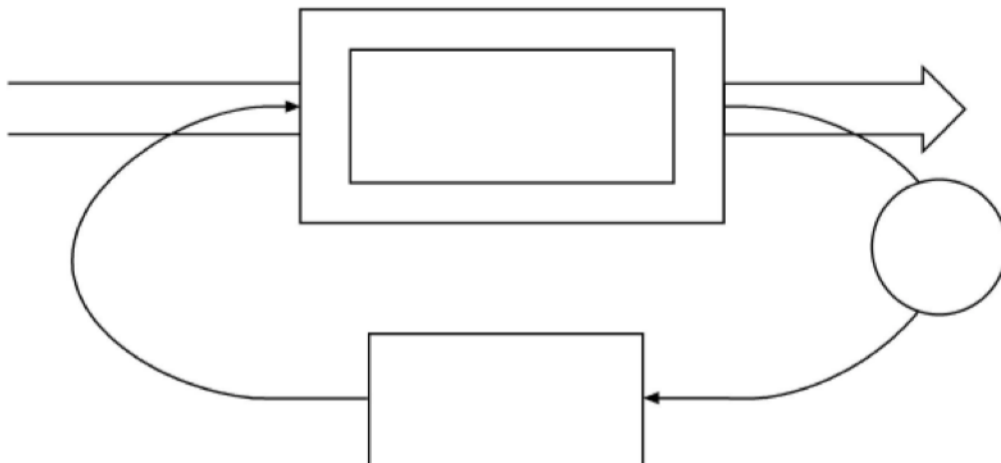
Automatizacija je počela jednostavnim oblicima, a s iskustvom je prerasla u složene sustave.

Informacijski sustav je prvi oblik automatizacije (Slika 2). Njime se dobivaju podaci o stanju objekta te njegovoj preradi. Čovjek, odnosno operater vrši neposredno upravljanje proizvodnim procesom.



Slika 2. Informatički sustav [5]

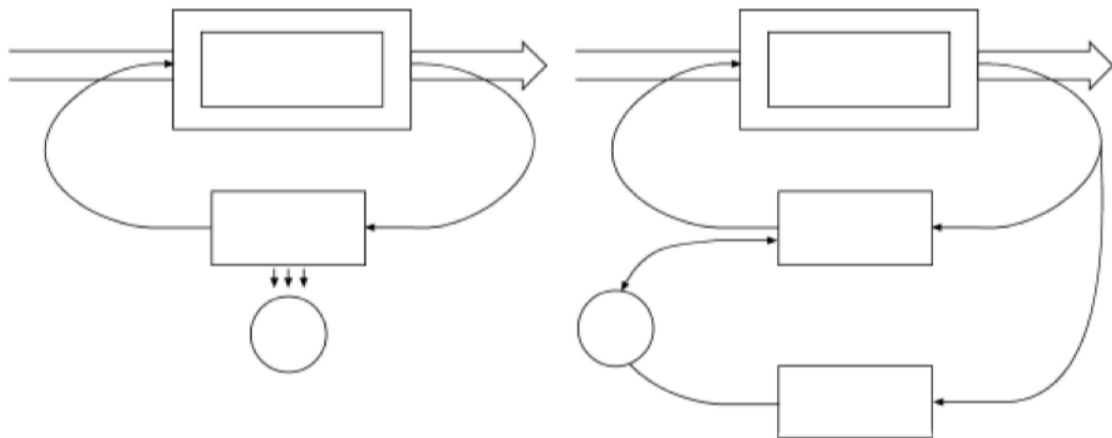
Za upravljanje objektom prema programu kojega je zadao operater koristi se upravljački sustav (Slika 3).



Slika 2. Upravljački sustav [5]

Informacijsko-upravljački sustavi predstavljaju treći oblik automatizacije (Slika 3). Bitna karakteristika je prerada automatskih informacija o stanju objekta za izdavanje naredbe izvršiocu ili izvršnom mehanizmu.

Čovjek je potpuno oslobođen od operacije upravljanja te mu ostaje samo kontrola nad procesom i eventualne intervencije u slučaju pojave nepredviđenih okolnosti.



Slika 3. Informacijsko-upravljački sustavi [5]

Dalje postoje sustavi koji se razlikuju po obuhvatnosti proizvodnih objekata, pa su tako definirani lokalni i kompleksni oblici automatizacije.

Lokalni predstavljaju automatsko upravljanje i regulaciju odijeljenim objektima i lokalnim uređajima. [5]

Kompleksni oblik automatizacije koristi se kod automatskog upravljanja složenih sustava.

Lokalni sustav predstavlja decentraliziran sustav upravljanja, dok pod centraliziranim se smatra sustav kod kojega se sve informacije vode u jedan centar gdje se obrađuju i iz kojeg se upravlja.

Kombinacijom lokalnog i centraliziranog sustava nastaje hijerarhijski sustav koji može sadržavati nekoliko razina upravljanja.

4.1. Automatizacija radnog procesa

Uz bok samom komunikacijskom sustavu ide automatizacija radnog procesa kao logični slijed razvoja tehnologije.

Protok informacija se uvelike olakšava što je nezaobilazno u dobu kada je sve bazirano na informacijama.

Automatizacija omogućava potpuno nove mogućnosti te ubrzava cijeli proces izrade grafičkih proizvoda. Kako bi se lakše integrirala u samu industriju potreban je neki oblik standardizacije.

U tiskarstvu se većina dobavljača dogovorila usvojiti standard povezivanja i slanja informacija između sustava. Koristiti će se jedan jezik, organizacija iza tog jezika je CIP4 (cip4.org), a jezik je JDF. On je jedinstven jer omogućava jednostavno povezivanje svih komponenata sustava, a zanimljiva je mogućnost povezivanja komponenata različitih dobavljača.

Integracijom informatike otvaraju se mogućnosti za daljinsko upravljanje kontrolnim procesima te stvaranja baza informacija.

Modularne jedinice koje služe za prijenos informacija se konstantno razvijaju te otvaraju mogućnosti nadogradnje.

Cijela JDF komunikacija bazirana je na XML-u koji predstavlja kodni sustav komunikacije te integrira aplikacijske sustave i standardizira tiskarske varijable.

Najveći naglasak kod razvoja stavlja se na internetsku komunikaciju.

Potrebno je ustrojiti proizvodne standarde i formirati vlastiti tiskarski rječnik kao osnovu za komunikaciju.

Za pravilno funkcioniranje sustava te kvalitetnu komunikaciju trebaju se postaviti proizvodni standardi te formirati tiskarski rječnik.

Sustav je definiran po horizontalnoj i vertikalnoj osi. Horizontalna predstavlja komunikaciju radnih operacija po fazama, a vertikalna prikazuje poveznicu uprave i administracije sa proizvodnjom te njihovu komunikaciju (Slika 4).



<http://www.ziljak.hr/tiskarstvo/tiskarstvo07/PojedinacniRadovi07/Miljkovict07clanak.html>

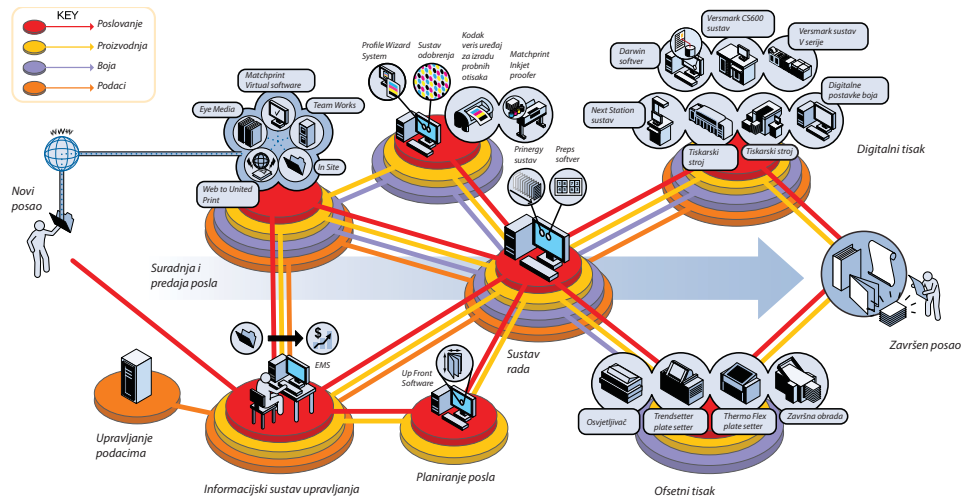
Slika 4. Vertikalna i horizontalna komunikacija [7]

Komunikacija se u svakom trenutku izvodi u svim smjerovima te osigurava kvalitetnu izmjenu informacija.

Procesna automatizacija i standardizacija grafičkih izvršnih tijekova što uključuje i protok digitalnih zapisa, iskazani su prema fazama: [7]

- omogućen je protok informacija od naručitelja prema doradi i isporuci;
- osiguran je kompletni nadzor nad svim proizvodnim procesima;
- omogućene su naknadne korekcije ili izmjene u samom procesu tijekom proizvodnje;
- izbjegnute su uske prohodnosti u radnim fazama odnosno preusmjeravanjem informacija prema novom planiranju proizvodnje kao i individualizacija svakog radnog mjesta s točnim opisom radnih zadataka;
- omogućeno je stvaranje baza i modela proizvodnje kao polazište za njenu automatizaciju.

Primjer automatiziranog radnog procesa (Slika 5):



<http://www.ambalaza.hr/hr/casopis/2009/12/ujedinjeni-tijek-rada-za-optimizaciju-poslovanja,110,3669.html>

Slika 5. Kodak *Unified Workflow* [6]

5. Indeksiranje

Pojavom računalne industrije dolazi do stvaranja baza podataka. Baza podataka skup je međusobno povezanih podataka pohranjenih na disku. Baze postaju sve veće i kompliciranije pa se uporabom indeksiranja znatno olakšava te ubrzava rad sa njima. Indeksiranje se definira kao proces stvaranja reprezentacije dokumenta.

Pojam indeksiranja dokumenata predstavlja postupak u kojem informacijski sustav pregledava dokumente na disku i katalogizira ih kako bi se omogućilo njihovo pretraživanje. U tome postupku, računalo pregledava dokument, utvrđuje sadržaj, format te određuje druge podatke kako bi korisnik kasnije mogao pretražiti sustav za tim dokumentom, slikom ili nekom drugom vrstom datoteke. Pošto je engleski najrašireniji jezik, većina metoda za indeksiranje je optimizirana upravo za taj jezik.

Razlikuje se više vrsta indeksiranja, ručno, poluautomatsko i automatsko indeksiranje.

5.1. Ručno indeksiranje

Ručno indeksiranje vrše grupe dokumentalista koje čitaju dokumente i prema njihovom sadržaju određuju prikladne ključne riječi ili ih svrstavaju u ranije definirane klase. Ovaj pristup može oduzimati mnogo vremena, kompliciran je i skup, ali ima i određene prednosti. Jezik kojim je dokument pisan, terminologija ili rječnik ne igraju ovdje nikakvu ulogu. Indeksiranje ne ovisi o tehnologiji, a lakše se prilagođava različitim korisnicima. Porast broja baza punog teksta i mogućnosti pretraživanja slobodnog teksta (*free text searching*) dala je poticaj razvoju računalnih programa koji su zamišljeni kako bi obrađivale dokumente jednako dobro, pa čak i bolje nego čovjek.

Zbog ogromne količine dokumenata i općenito podataka u današnje vrijeme, računala su nezaobilazan alat, a ručno indeksiranje postaje gotovo nemoguće.

5.2. Poluautomatsko indeksiranje

Poluautomatsko indeksiranje uključuje primjenu različitih automatskih metoda predlaganja ključnih riječi ili pojmova iz tezaurusa. Na taj način indekseri dobivaju znatno smanjen popis pojmova među kojima odabiru one bitne za dokument. Konačnu odluku donosi ponovno čovjek, ali se vrijeme i cijena indeksiranja znatno smanjuju. [10]

5.3. Automatsko indeksiranje

Automatsko indeksiranje obavlja se bez ili uz vrlo malu intervenciju čovjeka. Dobiveni rezultati su usporedivi s rezultatima ručnog indeksiranja, ali su pogreške ipak neizbježne. Ušteda vremena je velika te bez automatskog indeksiranja velika količina dokumenata nikada ne bi ni bila obrađena.

Sustav za automatsko indeksiranje moguće je prilagoditi potrebama te različitim vrstama dokumenata. Pojavljuju se problemi kao što su specifičnost terminologije, višejezičnost ili višeznačnost koji ne postoje kod ručnog indeksiranja, ali se mogu riješiti primjenom odgovarajućih metoda. Kombiniranjem ranije navedenih metoda ostvaruju se najbolji rezultati.

Kategorizacija je nadzirani proces koji se odnosi na svrstavanje dokumenata u dvije ili više klasa ili kategorija. Zahtjeva se učenje na primjerima unaprijed razvrstanih u različite kategorije. Kod indeksiranja se svakom tekstu pridodaje jedan ili nekoliko klasifikacija deskriptora. Oni se ponašaju na isti način kao kategorije kod procesa kategorizacije.

5.4. Vrste automatskog indeksiranja

Vrste automatskog indeksiranja su:

- metoda ekstrakcije
- metoda dodjeljivanja ključnih riječi
- lingvističke metode
- statističke metode

5.4.1. Metoda ekstrakcije

Metoda ekstrakcije koristi pojmove koji se već nalaze u tekstu dokumenta koji se indeksira. Ti pojmovi se odnose na samostalne riječi, ali i složenije fraze. Osnovni problem je u određivanju pojmova koji su dovoljno bitni za sadržaj dokumenta i definiranje njihove važnosti, odnosno težine. [10]

Za uspješno provođenje indeksiranja koriste se metode kao što su eliminacija stop riječi, morfološka normalizacija i formiranje fraza. Njima se eliminiraju općeniti i specifični pojmovi.

Prednost su brzina i fleksibilnost, ali postoje i mane kao što je neraspoznavanje višeznačnosti.

5.4.2. Metoda dodjeljivanja ključnih riječi

Kod metode dodjeljivanja ključnih riječi, ključne se riječi preuzimaju iz nekog izvora koji nije sam dokument. Njih može odrediti indeksirer na temelju vlastitog iskustva ili mogu biti preuzete iz nekog kontroliranog rječnika kao što je tezaurus. Okupljanje na bazi tezaurusa ili klasifikacije podrazumijeva izuzetan ljudski intelektualni napor, pa se oduvijek težilo strojnom obavljanju tog procesa.

Ovdje je riječ o konceptualnom ili tematskom tezaurusu. Dodjeljivanje ključnih riječi na temelju kontroliranog rječnika usko je povezano sa pojmom kategorizacije.

5.4.2.1. Tezaurus

Pojam tezaurusa pojavljuje se u različitim kontekstima. Postoje također različite klasifikacije i vrste tezaurusa, a onaj koji je bitan za svrhu indeksiranja naziva se konceptualni ili tematski tezaurus.

Svojstvo koje konceptualni tezaurus izdvaja kao posebnu vrstu leksikografskih priručnika je njegovo konceptualno, odnosno tematsko ustrojstvo, za razliku od abecednog koje se inače koristi. Osim toga, tezaurus karakterizira i semantička organizacija, koja se obično prikazuje kroz hijerarhijsku strukturu. Pojedini pojmovi sudjeluju u izgradnji nekog šireg i općenitijeg koncepta. Tezaurus se često pogrešno poistovjećuje s rječnikom sinonima. Razlika je u tome što on osim sinonima opisuje i mnoge druge značajno složenije leksičke relacije.

Osim opisanih svojstava, tezaurus sadrži abecedni indeks i mnoge druge informacije na nižoj razini. [10]

5.4.3. Lingvističke i statističke metode

Lingvističke metode određuju pojmove koji su bitniji za formiranje indeksa pridjeljujući im veće težine. Takvi pojmovi najčešće su imenice ili grupe imenica, a cilj lingvističkih metoda je prepoznavanje istih u tekstu. Kvaliteta indeksiranja znatno se poboljšava i prepoznavanjem fraza. Kod određivanja važnosti pojmova, lingvističke se metode kombiniraju sa statističkim metodama. Korištenjem morfološke normalizacije i eliminacije stop riječi kao odvojenih procesa prije glavne obrade, dobivaju se kvalitetniji rezultati.

6. Strojno učenje

Strojno učenje je metoda analize podataka koja automatizira analitički model izgradnje. Koristeći algoritme koji iterativno uče iz podataka, računalima je omogućeno pronaći skrivene spoznaje bez njihovog eksplicitnog programiranja gdje nešto tražiti.

Iterativni aspekt strojnog učenja je vrlo važan jer kako su modeli izloženi novim podacima, u mogućnosti su nezavisno se prilagođavati. Uče iz prijašnjih izračuna kako bi pružili pouzdane, ponovljive odluke i rezultate.

Većina algoritama strojnog učenja postoji već dugo vremena, ali za razliku od prošlosti, danas je moguće konstantno automatski vršiti kompleksne matematičke izračune za veliku količinu podataka.

Današnji primjeri strojnog učenja široke uporabe su:

- autonomna vozila
- preporuke za gledanje i slušanje na uslugama kao što su Netflix, Spotify itd.
- različite usluge na društvenim mrežama (npr. prepoznavanje lica)
- prepoznavanje prijave

Postoji nekoliko metoda strojnog učenja od kojih su dvije najčešće korištene: nadgledano učenje i nenadgledano učenje. Nadgledano učenje je daleko najkorištenije (oko 70%), dok nenadgledano učenje čini 10% do 20% sveukupnog strojnog učenja.

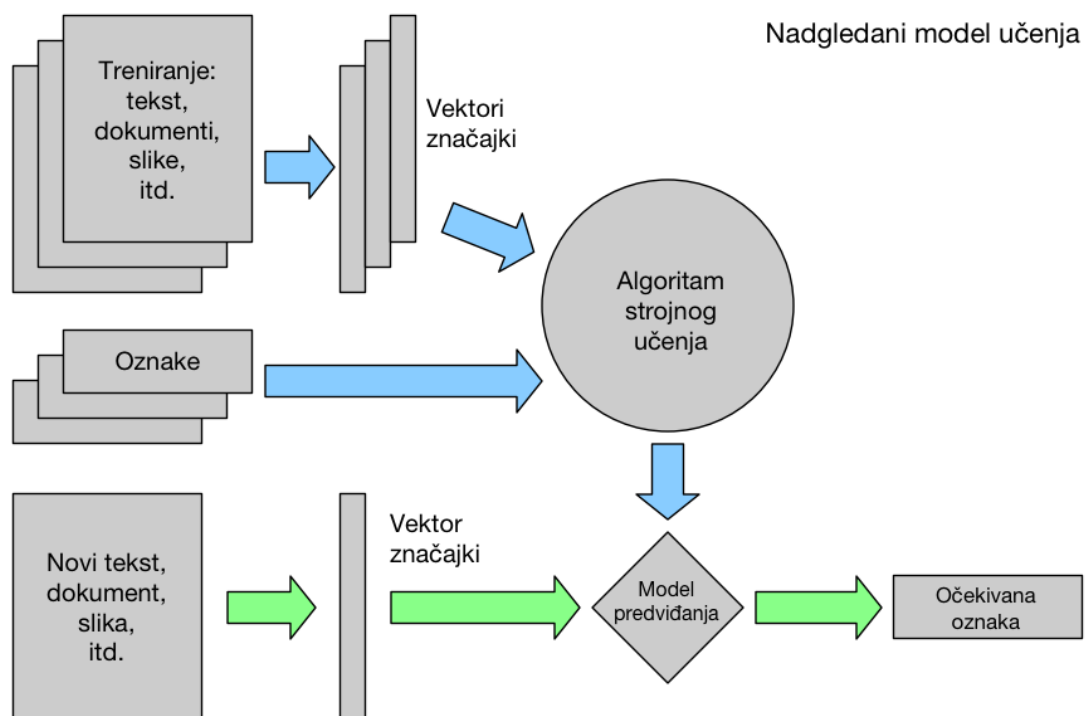
6.1. Nadgledano učenje

Algoritmi nadgledanog učenja su osposobljeni korištenjem označenih primjera, kao ulaza gdje je željeni izlaz poznat. Algoritam za učenje dobiva skup ulaza zajedno sa odgovarajućim točnim izlazima te uči na temelju usporedbe realnog izlaza sa odgovarajućim točnim izlazima (Slika 6). Prema tim podacima algoritam mijenja model kako bi rezultat bio traženi izlaz.

Pomoću metoda kao što su klasifikacija, regresija itd. kod nadgledanog učenja se koriste uzorci za predviđanje vrijednosti oznaka na dodatnim neoznačenim podacima.

Često se koristi u slučajevima gdje povijesni podaci mogu predvidjeti moguće buduće događaje.

Primjer: predviđanje kada bi moglo doći do varanja kod transakcija kreditnim karticama ili koji osiguranik bi mogao podnijeti zahtjev.



http://www.astroml.org/sklearn_tutorial/general_concepts.html, (prevedeno)

Slika 6. Model nadgledanog učenja

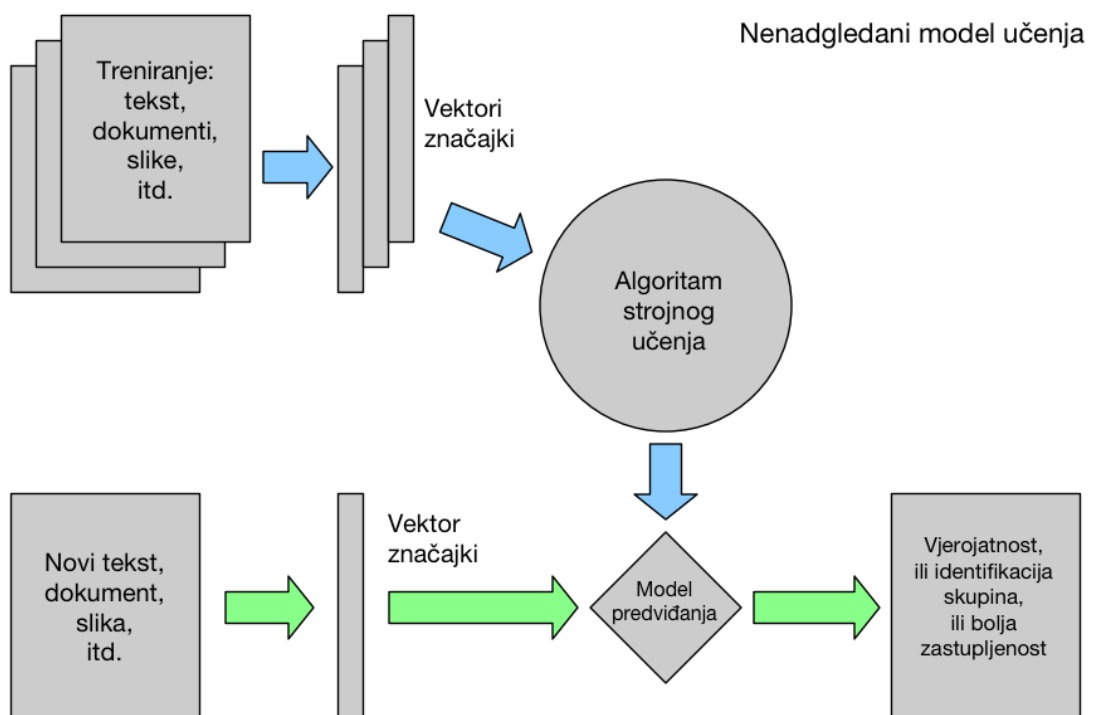
6.2. Nenadgledano učenje

Koristi se kod podataka koji nemaju povijesne oznake. Sustavu se ne daje točan izlaz sustava na temelju kojega bi mogao naučiti već algoritam mora sam zaključiti što se prikazuje (Slika 7). Cilj je istražiti podatke i pronaći strukturu.

Popularne metode su:

- samoorganizirajuće karte (*self-organizing maps*)
- mapiranje najbližim susjedom (*nearest-neighbor mapping*)
- grupiranje po k sredinama (*k-means clustering*)
- rastav singularnih vrijednosti (*singular value decomposition*)

Navedeni algoritmi se koriste i za segmentiranje tekstualnih tematika, preporuku stvari i identifikaciju podatkovnih odudaranja.



http://www.astroml.org/sklearn_tutorial/general_concepts.html, (prevedeno)

Slika 7. Model nadgledanog učenja

7. Programi

Za realizaciju sustava automatskog prepoznavanja riječi, indeksiranja i prosljeđivanja u grafičkom poduzeću, potrebo je koristiti određene softvere.

Indeksiranje se može provoditi na različite načine i različitim softverima, kao i prepoznavanje i prosljeđivanje te će neki od njih biti objašnjeni.

7.1. Gmail

Najjednostavniji primjer alata za ovu namjenu je Gmail.

Gmail je Google-ov alat za elektroničku poštu koji je postao najrašireniji alat ovoga tipa. Osnovna namjena je komuniciranje putem elektroničke pošte, ali njime je moguće učiniti i mnogo više od toga.

Jedna vrlo moćna opcija su filteri.

Filteri omogućuju automatiziranje različitih akcija za dolaznu poštu. Moguće je postaviti kriterije za automatsku primjenu pravila za određene adrese, poruke s određenim predmetom (*subject*), poruke sa određenim ključnim riječima i mnogo drugog.

Zanimljiva opcija je korištenje ključnih riječi koje se definiraju kao filteri. Odabere se opcija naprednog pretraživanja, upišu se željene ključne riječi te odabere kreiranje filtera s tim pretraživanjem (Slika 8).

Search **All Mail** x

From

To

Subject

Has the words **1**
Grf

Doesn't have

Has attachment

Don't include chats

Size **greater than** MB

Date within **1 day** of

2

Slika 8. Gmail filteri

Nakon što su definirane ključne riječi po kojima će se provoditi filter, moguće je odabrati opciju prosljeđivanja pristigle poruke od strane klijenta na neku drugu adresu kako bi poruka pristigla nadležnom stručnjaku ili odjelu grafičkog poduzeća zaduženom za traženi posao definiran u poruci.

To se čini dodavanjem internetske adrese na koju će se prosljeđivati poruke određene kreiranim filterom (Slika 9).

[« back to search options](#) x

When a message arrives that matches this search:

Skip the Inbox (Archive it)

Mark as read

Star it

Apply the label: Choose label... **3.**

Forward it [add forwarding address](#)

Delete it

Never send it to Spam

Always mark it as important

Never mark it as important

Categorize as: Choose category...

4. Also apply filter to **26** matching conversations.

Slika 9. Gmail prosljeđivanje pošte

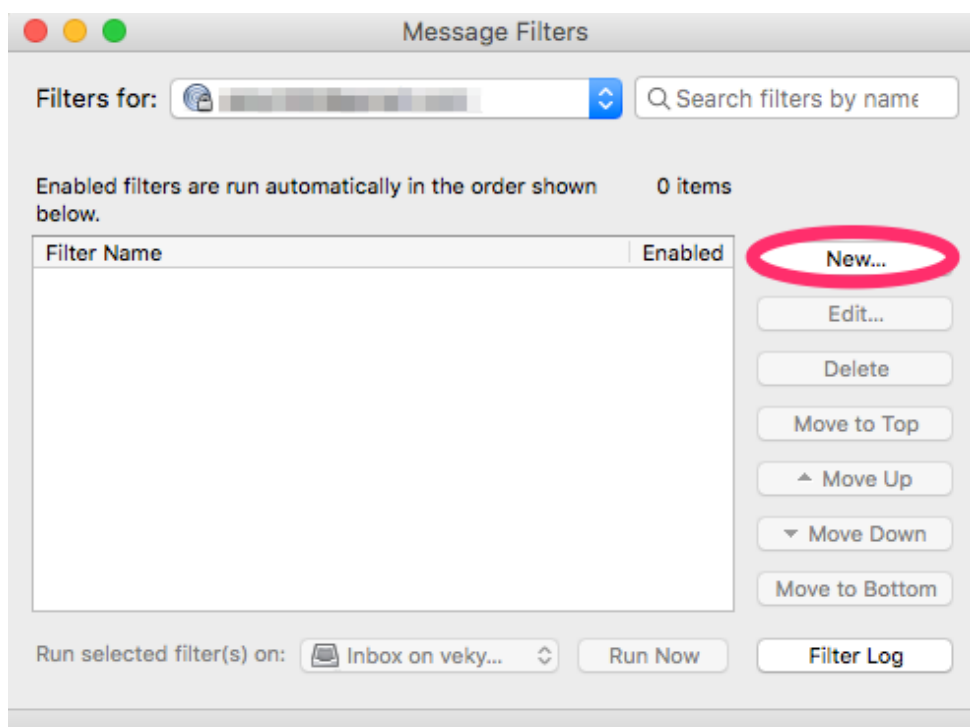
Postupak se ponavlja za svaki posao grafičkog poduzeća koji obavlja određeni stručnjak ili odjel te je tako na jednostavan način automatiziran proces za prosljeđivanje. Ovaj način nije potpuno automatski te je relativno osnovan pošto svaku ključnu riječ mora unijeti čovjek te ju pridodati određenom poslu.

6.2. E-mail klijent (Thunderbird)

Nešto naprednija opcija za korištenje filtera i prosljeđivanja elektroničke pošte je uporabom klijenta za elektroničku poštu.

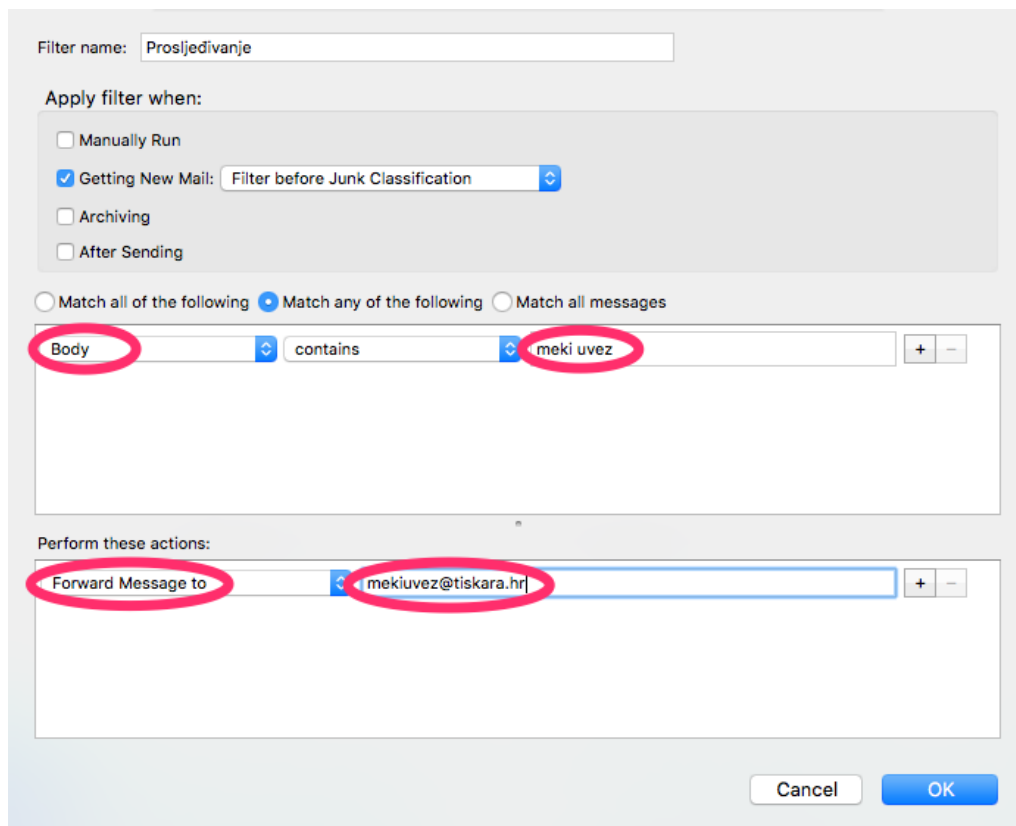
Jedan od besplatnih programa je Thunderbird.

S ugrađenom podrškom za filtere, Thunderbird nudi mnogo mogućnosti i vrsta filtriranja. Za ove potrebe koristi se filtriranje po sadržaju poruke (*body-contains*) te se unose ključne riječi po kojima će se pretraživati (Slika 10).



Slika 10. Thunderbird filteri

Zatim još preostaje unijeti elektroničku poštu na koju će poruka biti prosljeđena (Slika 11).



Slika 11. Thunderbird filteri

Kao i kod korištenja Gmail-a, ovaj način je potpuno ručni. Čovjek mora definirati ključne riječi i unositi ih u filtere po potrebi. Taj proces oduzima određeno vrijeme, ali pošto je cijeli proces pod kontrolom čovjeka, mala je vjerojatnost pojave krivih ključnih riječi za krive poslove koju grafičko poduzeće obavlja.

7. Strojno prevođenje

Ako podrazumijevamo da je grafičko poduzeće kvalitetno i prati ga dobar poslovni uspjeh, ono se širi, odnosno širi svoje poslovanje. Tako se poslovanje može proširiti van državnih granica što podrazumijeva i drugi jezik, odnosno zahtjevi za poslove dolaze najčešće na engleskom jeziku pa postoje sustavi koji mogu prevođenje olakšati.

Postoji nekoliko vrsta programa za strojno prevođenje koji se najčešće klasificiraju unutar dvije osnovne kategorije: način uporabe i princip rada.

Prva kategorija sadrži tri grupacije: "strojno prevođenje-uz-potporu-prevoditelja" (*Human-aided-Machine-Translation*, HAMT), "prevođenje-uz-potporu-programa" (*Machine-aided-Human-Translation*, MAHT) te "potpuno automatizirano strojno prevođenje" (*Fully-Automatic-High-Quality-Translation*, FAHQQT).

Prve dvije grupacije često se zajednički nazivaju računalno-potpomognuto-prevođenje (*Computer-aided Translation*, CAT) jer su međusobno isprepletene. Treća je za sada u praksi neostvariva, osim u pojedinim slučajevima jako ograničenog unosa i rezultata te je cilj programera strojnog prevođenja.

7.1. Princip rada

Gledajući klasifikaciju prema principu rada, postoji nekoliko glavnih vrsta koncepata programa za strojno prevođenje.

Najstarija i najosnovnija vrsta je bazirana na principu rječnika (*dictionary-based*).

Kod ove vrste riječi se prevode jedna po jedna, prema definiciji rječnika.

Rezultat je vrlo niska kvaliteta, osim u uskim područjima primjene kao što su inventorne liste.

Kod vrste bazirane na principu primjera funkcioniranje ovisi o usporedbi baza podataka s paralelnim zapisima na dva jezika, a nazivaju se *bilingual corpora*.

Ako direktno u *bilingual corpora* ne postoji tekst unosa, traži se što sličniji primjer. Najkorisniji je tzv. minimalni par (*minimal-pair*) gdje se primjer razlikuje samo po jednoj riječi.

Primjer:

Polazni tekst unosa: Vaš proizvod je preskup.

Bilingual corpora minimalni par: Vaša usluga je preskupa. *Your service is too expensive.*

Rezultatni tekst (korištenje minimalnog para + prijevoda samo 1 riječi)

Your product is too expensive.

Ove prve vrste pristupa pokazale su se nezadovoljavajuće kod prevođenja tekstova, ponajviše zbog složenosti međuzavisnosti unutar strukture rečenice.

Zbog toga su se pojavili koncepti koji se ne zasnivaju samo na pretraživanju i usporedbi podataka već i na gramatičkim pravilima (*rule-based*).

Zbog svoje kompleksnosti i raznih načina pristupanja problemima prevođenja, ova posljednja vrsta razvila se u više smjerova.

Transfer je jedan od njih, gdje za svaki jezik postoji potprogram koji uneseni tekst svodi na osnovno značenje koje je pogodno za jezik na koji se vrši prijevod. Takav input unosi se u potprogram drugog jezika, prevodi na osnovno značenje te se stvara konačni tekst prijevoda.

Primjer:

unos (hrvatski) → modul hrvatsko-engleski prijevod → modul englesko-hrvatski prijevod → rezultat (prijevod na engleskom)

Usljedila je i ideja općeg jezika, Interlingue. Taj se proces bazira na prijevodu polaznog teksta na "međujezik" te iz njega u određeni jezik.

Primjer:

unos (hrvatski) → modul hrvatsko-Interlingua prijevod → modul Interlingua-engleski prijevod → rezultat (prijevod na engleskom)

Zbog pojednostavljenja procesa prevođenja (jednostavna struktura i pravila) i zbog smanjenja broja potrebnih potprogramskih modula, princip Interlingua je vrlo pogodan. U slučaju da program radi sa pet jezika, svaki zasebno mora imati po osam modula za kontakte s ostalima (4 ulazna, 4 izlazna). Uvođenjem Interlingua broj potrebnih modula smanjuje se na dva (1. Jezik → Interlingua i Interlingua → 2. jezik).

Posrednički pristup (*pivot-type*) je hibridna varijanta prethodna dva pristupa. Umjesto Interlingue koristi se bolji modul jezika srodnog izvornom ili rezultatnom (najčešće neki od "svjetskih jezika") te se na taj način postiže bolji rezultat.

Primjer:

engleski → ruski → ukrajinski

Ubrzanim razvojem umjetne inteligencije i hardverske podrške obrade informacija koju provodi umjetna inteligencija te razvojem interneta kao globalne društvene komunikacijske mreže pojavili su se internetski portali za strojno prevođenje. Oni koriste spoj *rule-based* i *example-based* pristupa koji se naziva "statistički stroj" (*statistical-machine*) gdje se kod analize izvornog i sintezi rezultatnog teksta, uz gramatička pravila, koristi i ogromna baza podataka prevedenih tekstova. Metodom statističke usporedbe pronalazi se odgovarajuća riječ ili čitava rečenica s najvećom vjerojatnošću ispravnosti.

Rezultati su se proporcionalno poboljšavali kako je napredovala umjetna inteligencija te količina prevedenih tekstova.

Danas je glavni predstavnik toga tipa Google Translate koji koristi *brute-force* algoritme iz područja umjetne inteligencije te bazu podataka prijevoda baziranu na UN-ovoj dokumentaciji koja sadrži preko 200 milijardi zapisa.

7.2. Problemi

Najčešći i najteži problem prevođenja je višeznačnost. Može imati više oblika, a najpoznatiji su polisemija (više mogućih značenja iste riječi - sinonimi) i homografija (riječi koje se jednako pišu, ali imaju različito značenje).

Kod nekih drugih jezika izraženija je i kategorijska višeznačnost. U engleskom, mnogo imenica i glagola ima isti oblik (*cut* = rez (imenica), rezati (glagol)), pa je teže odrediti gramatičku kategoriju.

Višeznačnost može proizaći i iz raznih prenesenih značenja iste riječi. Ona se pojavljuje kod prevođenja skupova riječi ili fraza.

U rečeničnoj strukturi najznačajniji je problem analiza sintakse (određivanja subjekta, objekta itd.).

U hrvatskom je rečenični poredak slobodniji, ali se uloga riječi može odrediti padežima ili sufiksima gramatičkih kategorija.

Budući da su u engleskom jeziku padeži i gramatičke kategorije manje odvojene, u sintaksi se mora koristiti puno složeniji rečenični poredak.

Kod prijevoda složenih rečenica pojavljuje se problem anafore, odnosno referenciranja na drugi dio teksta. Anafora se najčešće primjećuje kroz uporabu zamjenica te je posebno zahtjevna za strojno prevođenje zbog ograničenja današnjih algoritama i procesorske snage računala, pa se samo strojno prevođenje većinom odvija rečenicu po rečenicu.

Za programe strojnog prevođenja često je jako veliki problem povezati zamjenicu s imenicom navedenom nekoliko rečenica prije.

To su samo neki od najznačajnijih problema u strojnom prevođenju od mnogo njih za značenje riječi i gramatičku točnost.

8. Eksperimentalni dio

Ovaj dio rada prikazuje kako se na jednostavan način mogu postići pojednostavljeni, odnosno automatizirani ulazni procesi u grafičkom poduzeću. Postoje i mnogo napredniji sustavi, ali primjer koji slijedi je dobar prikaz jednostavnosti i funkcionalnosti.

Kao što je već poznato, elektronička pošta ima glavnu ulogu kao ulazni proces pomoću kojega grafičko poduzeće zaprima zahtjeve za posao.

Kao primjer korišten će biti besplatni program Thunderbird koji je prethodno ukratko objašnjen.

Na isti će način i ovdje biti definirane varijable, odnosno filteri te elektroničke adrese na koje se vrši prosljeđivanje.

8.1. Definiranje problema

Kao primjer grafičkog poduzeća uzeta je zamišljena knjigovežnica čiji raspon posla obuhvaća sve vrste i forme uveza te uključuje veću svakodnevnu potražnju. Svaku dolaznu elektroničku poštu ugrađeni filter će obraditi, tj. pretražiti i pronaći ključne riječi. Zatim će odraditi definiranu akciju.

8.2. Konfiguracija programa

Konfiguracija se vrši na već prethodno opisan način.

Definiramo ključne riječi za meki, tvrdi i mehanički uvez te adrese elektroničke pošte za te odjele knjigovežnice.

Primjer ključnih riječi za meki uvez su: bilježnica, beletristika, meke korice, završni rad, strip, časopis, udžbenik, priručnik. Iste su unese u filter programa (Slika 12).

Kao adresa elektroničke pošte definirana je: mekiuvez@gmail.com.

Match all of the following	Match any of the following	Match all messages	
<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	
Body	contains	bilježnica	+ -
Body	contains	beletristika	+ -
Body	contains	meke korice	+ -
Body	contains	završni rad	+ -
Body	contains	strip	+ -
Body	contains	časopis	+ -
Body	contains	udžbenik	+ -
Body	contains	priručnik	+ -

Perform these actions:

Forward Message to	mekiuvez@gmail.com	+ -
--------------------	--------------------	-----

Slika 12. Filtri za meki uvez

Primjer ključnih riječi za tvrdi uvez: tvrde korice, monografija, slikovnica, knjiga leksikon (Slika 13).

Definirana adresa: uveztvrdi@gmail.com

The screenshot shows an email filter configuration window. At the top, there are three radio buttons: 'Match all of the following' (unselected), 'Match any of the following' (selected), and 'Match all messages' (unselected). Below this, there are six rows of filter criteria. Each row consists of a dropdown menu set to 'Body', a dropdown menu set to 'contains', and a text input field with a specific keyword. To the right of each row are '+' and '-' buttons. The keywords are: 'tvrde korice', 'monografija', 'slikovnica', 'knjiga', 'leksikon', and 'konac'. Below the filter criteria, there is a section titled 'Perform these actions:'. It contains a dropdown menu set to 'Forward Message to' and a text input field containing the email address 'uveztvrdi@gmail.com'. To the right of this section are '+' and '-' buttons.

Slika 13. Filtri za tvrdi uvez

Primjer definiranih ključnih riječi za mehanički uvez: spirala, matica, vijak, zakovica, kalendar (Slika 14).

Definirana adresa: mehanickiuvez@gmail.com

The screenshot shows an email filter configuration window. At the top, there are three radio buttons: 'Match all of the following' (unselected), 'Match any of the following' (selected), and 'Match all messages' (unselected). Below this, there are five rows of filter criteria. Each row consists of a dropdown menu set to 'Body', a dropdown menu set to 'contains', and a text input field with a specific keyword. To the right of each row are '+' and '-' buttons. The keywords are: 'spirala', 'matica', 'vijak', 'zakovica', and 'kalendar'. Below the filter criteria, there is a section titled 'Perform these actions:'. It contains a dropdown menu set to 'Forward Message to' and a text input field containing the email address 'mehanickiuvez@gmail.com'. To the right of this section are '+' and '-' buttons.

Slika 14. Filtri za mehanički uvez

8.3. Rezultati i rasprava

Provođenjem eksperimenta može se zaključiti da sustav funkcionira u ograničenom smislu. Filteri su prepoznali riječi ako su napisane u identičnom obliku kao što je definirano te u slučaju ako je izmijenjen padež kao u primjeru: časopis – časopisa.

Veća odudaranja oblika riječi od definiranog te korištenje riječi bez dijakritičkih znakova rezultiralo je neuspješnim prepoznavanjem.

Time se utvrđuje da je za ispravno funkcioniranje ovakvog sustava potreban određeni vremenski period kako bi se kroz iskustvo i prikupljanjem određene količine primjera mogla proširiti baza ključnih riječi.

Mana ovog sustava je ta što je potreban čovjek koji bi u tom vremenskom periodu popunjavao bazu te tijekom cijelog vremena funkcioniranja periodički provjeravao postoji li moguća nadopuna baze na temelju pristiglih elektroničkih poruka.

Ipak je sveukupna ušteda vremena veća od vremena koje je normalno potrebno čovjeku za čitanje i podjelu poslova odjelima.

Pojavljuje se i problem preklapanja pojedinih područja. Kao primjer može poslužiti bilježnica koja može pripadati i mekoj, i tvrdoj, i mehaničkoj vrsti uveza. Taj problem djelomično se može riješiti korištenjem više filtera koji bi prosljeđivali poruke samo u slučaju ako postoje dvije ili više ključnih riječi koje se u istom trenutku podudaraju. Primjer bi bio da je uvjet za prosljeđivanje prepoznavanje riječi bilježnica, ali i riječi spirala u istoj poruci te da se tek tada pokreće prosljeđivanje na adresu za mehanički uvez.

Postoji i mnogo slučajeva gdje je teško pronaći zadovoljavajuće rješenje, a glavni uzročnik je što algoritmi još uvijek nisu dovoljno napredni kako bi kontekstualno razumjeli ljudski način komuniciranja.

Napreci u strojnom učenju te umjetnoj inteligenciji bi s vremenom i taj problem trebali riješiti.

9. Zaključak

Grafička industrija kao grana sa svojim velikim opsegom proizvoda i uključenošću u mnoge druge industrijske grane ostavlja prostor za stalno unapređivanje i moderniziranje, pa tako i automatizaciju.

Automatizaciji se teži već dugo vremena, a današnje vrijeme je vrijeme velikih i brzih napredaka. Kako bi se potpuna automatizacija omogućila, mnogo zahtjeva mora biti zadovoljeno.

Kao što je navedeno na početku rada, mora postojati komunikacijski sustav te se uvoditi jedno od rješenja automatiziranog radnog procesa koje će postaviti dobre temelje za veće stupnjeve automatizacije.

Kako bi se pojednostavili ulazni procesi bilo kojeg sustava i poduzeća, pa tako i grafičkog poduzeća, potrebno je koristiti sustave za pojednostavljenje tih procesa, koji mogu biti manje ili više napredni. U ovom radu prikazan je jednostavniji sustav koji se može dovesti do iznimno visoke razine kroz određeni period korištenja i optimizacije.

Napredniji sustavi koriste strojno učenje te oblike umjetne inteligencije koja svakim danom postaje sve bolja. Njenim razvojem, čovjek neće biti potreban u mnogim područjima koje danas obavlja te će takvi sustavi moći zaživjeti u svom punom smislu.

Svaki od tih sustava bazira se na danas već naprednim algoritmima indeksiranja. Oni su međusobno u korelaciji, pa se tako, paralelno, i razvijaju.

U cijeloj toj priči nedostaje jedna važna komponenta, a to je jednostavna i efikasna komunikacija između čovjeka i stroja. Strojnim učenjem ona polako postaje dostižna.

Međuljudska komunikacija također se znatno olakšava. U svijetu s mnogo različitih jezika postoji određena udaljenost između ljudi različitih govornih područja. Strojno prevođenje je grana razvoja koja te udaljenosti umanjuje. Razvojem algoritama moguće je, vjerojatno u skorijoj budućnosti, doživjeti govorno ujedinjeni svijet. Time se otvaraju mnoge mogućnosti, a u poslovanju

grafičkog poduzeća omogućuje se jednostavno poslovanje bilo gdje, odnosno širenje izvan granica države u kojoj poduzeće posluje.

Sav taj razvoj strojnog učenja, prevođenja, umjetne inteligencije, ima još veću težinu znajući da grafička industrija prelazi iz industrije fizičkih, opipljivih proizvoda u industriju digitalnih proizvoda, čije je globalno širenje neusporedivo jednostavnije.

10. Literatura

1. <http://www.enciklopedija.hr/natuknica.aspx?id=23031>, Enciklopedija, Grafička industrija, 17.5.2016.
2. <http://www.enciklopedija.hr/natuknica.aspx?ID=23033>, Enciklopedija, Grafička tehnologija, 17.5.2016.
3. Čerepinko, D. (2012.): Komunikologija: Kratki pregled najvažnijih teorija, pojmova i principa; Veleučilište u Varaždinu, Varaždin
4. http://gogss.hr/wp-content/uploads/2012/10/graficka_tehnologija_prvi_razred.pdf, Škola za dizajn, grafiku i održivu gradnju – Split, Uvod u grafičku industriju, 19.5.2016.
5. Nedjeljko Perić, Ivan Petrović, Mario Vašak (2013). Procesna automatizacija, dostupno na: https://www.fer.unizg.hr/download/repository/Skripta_Proaut_veljaca_2013.pdf, 20.5.2016.
6. http://www.ambalaza.hr/hr/croprint/2007/1/automatizacija-radnog-tijeka-bit-uspjesnoga-tiskarskog-posla-_workflow-automatisation-the-essence-of-successful-printing,80,2680.html, Institut za ambalažu i tiskarstvo Tectus, Automatizacija radnog tijeka- bit uspješnoga tiskarskog posla / Workflow Automatisation – The Essence of Successful Printing, 21.5.2016.

7. [http://www.ziljak.hr/tiskarstvo/tiskarstvo07/PojedinacniRadovi07/Miljkovic t07clanak.html](http://www.ziljak.hr/tiskarstvo/tiskarstvo07/PojedinacniRadovi07/Miljkovic%20t07clanak.html), Automatsko upravljanje grafičkom proizvodnjom u digitalnom okruženju, 28.5.2016.
8. <http://www.judygordon.biz/what-is-indexing>, What is indexing?, 26.6.2016.
9. Automatic vs. manual indexing, European Library Automation Group, 2001, <http://www.stk.cz/elag2001/Workshop/ws4.doc>, 21.5.2016.
10. Ana Cvitaš, (2005). Automatsko indeksiranje dokumenata u modelu vektorskog prostora, diplomski rad, Fakultet elektrotehnike i računarstva
11. http://www.sas.com/en_id/insights/analytics/machine-learning.html, SAS, Machine learning, 28.6.2016.
12. Šimić, J., Uglarik, D. (2010). Mogućnosti i ograničenja strojnog prevođenja, Praktični menadžment: stručni časopis za teoriju i praksu menadžmenta, 1(1), 81-85. Preuzeto s: <http://hrcak.srce.hr/67847>, 25.6.2016.