

[Open in app ↗](#)

Search

69



Attention Economy

Article on Attention Economy in Marathi Newspaper, Sakal



Yogesh Haribhau Kulkarni (PhD)

Published in Desi Stack

5 min read · Just now

 [Share](#) [More](#)

'लक्ष' द्यावे उमजून

भाष्य



डॉ. योगेश कुलकर्णी

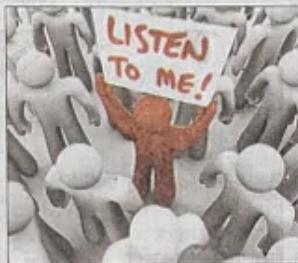
सध्या मोबाईल आणि समाजमाध्यमांमुळे एक नवीनच अर्थव्यवस्था प्रभावी होत आहे...
त्याला आपण 'अटेन्शन इकॉनॉमी' म्हणजेच 'लक्ष अर्थव्यवस्था' म्हणू शकतो.
सध्या माहितीची वानवा नसून तुमच्या 'अवधानाची आहे. तुमचे लक्ष कुठे आहे. यावर डिजिटल युगातील कंपन्यांचे लक्ष आहे.
त्यातून मिळाला अमाप नाही हे त्यांचे 'लक्ष' आहे.

■ लेखक 'कृष्ण बुद्धिमत्ता' विषयावे सत्तागम आहेत.

फा

र पूर्वीची अर्थव्यवस्था शेतीवर आधारित होती. औद्योगिक क्रांतीनंतर उद्योग व सेवाक्षेत्र अर्थव्यवस्थेचे महत्वाचे भाग झाले. संगणक आणि इंटरनेट क्रांतीमुळे माहितीची अर्थव्यवस्था बढून घर लागले. सध्या, मोबाईल व समाजमाध्यमांमुळे एक नवीनच अर्थव्यवस्था प्रभावी होत आहे. त्याला आपण 'अटेन्शन इकॉनॉमी' म्हणजेच 'लक्ष अर्थव्यवस्था' म्हणू शकतो. १९७५च्या 'अटेन्शन शासी' या लेखात शावक मानवकल गोळवडवर यांनी 'लक्ष आधारित अर्थव्यवस्थेची परिवर्त्यादा वाच्यता केली. त्याच्या म्हणण्यानुसारा, सध्या स्पृशी ही फक्त माहिती पुरवण्यात राहिलेलीच नाही, तर तुमचे लक्ष राखण्याकाऱ्ये वाललेली आहे. एकांदी वास्तव घायलीची असेहा तर इंटरनेटर त्यांनी माहिती असेही विकासातील निझो. जर एखादी साईट जास्त खिळवूऱ्या ठेवारी असेहा तर तुम्ही तेथे जास्त बेळ थांबाल; तर तुम्हाला जास्त वाचवाता येतो. मग त्या सांख्यिक्याची कंपनीला त्याचा जास्त कायदा होईल, नाही का? सध्या माहितीची वानवा नसून तुमच्या 'लक्ष'ची आहे. डिजिटल युगातील कंपन्यांचे तुमच्या लक्षावर त्यांनी आहे; त्यातून मिळाला अमाप नाही हे त्यांचे 'लक्ष' आहे.

बन्याच जणांना वाढो वी, हे सारे मस्त, फुकट मिळाला आहे, त्यांना जाण निझीतर आहो. यापाल करमणक आहे तांना एक विकासातील असे लक्ष येईल वी, हे विलकून फुकट नसून तुम्ही (म्हणजे तुमचे 'लक्ष') विकले जात आहो, जेव्हे जास्त तुमचे लक्ष ते खेचतील, गोळा करतो, ठेवाडा जास्त जाहितीचा मुहालू. लक्ष अर्थव्यवस्था' मोबाईलस्टोर किंवा स्ट्रीमिंग वज्रवृऱ्या ठेवण्यावर भर देतो. एकांदा खिळिंगो अपारा रोल पाहातानाच दुर्घे संदोर येते, अपारा मुख्यले याते. फोटोचे-चायाचिंहांचे ही तेच. आपारा एकांदा एक बघायला लागलो वी, तांसात कसे गेले हे कठवाच नाही. स्लिंग काहीतो नव्यान, आवाक्षणक संगोर येते. 'डोयामाई' मिळतो, मग अजून अजून बघायेसे यातो. तुम्हार काही संदर्भ (मेसेज-नोटिफिकेशन) आवाक्षणक कीला वाचावता. तर नाहीतर किमान ईंगींती तात्काळ टाकावीची वाचावता. बांसच्या इ-मेसेज तर पुढच्या मिनिटालाच उत्तर गेले पाहिजे याची घडपट, नियम टाटाटीचा नस्तानासुदा. हे अप्सनांनेतोकडे नेप्पासरखेच आहे. काही करमणी मोबाईल, समाजमाध्यम वास्यापास नाही मिळतो, जवळ नसेल तर अर्थव्यवस्था वाहयाव रहेते. विकास, त्रांग तर काहीच नाही, वैष्णव्याकडे वाटावर मुक्क होते. मोबाईल तुम्हाराची विषयावेस आठवा! तेहा आण गेझी लक्ष देतानां कायचो. मुख्यला याचेच, नाटक सिनेमा पालने, त्याकुमे लक्ष हे स्वतःहून काही वाई नाही. पण सतत बदलावते लक्ष याचाक आहे. त्याला 'कॉटेनेस्ट सिलिंग' म्हणजेच 'लक्ष' वाचावच्या गोष्टीकडे वारंवार बदलते



राहणे' असे म्हणातात. एक सुरुचील मेसेज विभिन्नांनी लोग दुसरा प्रश्न करत नेव्ही आहे. सध्याकाळीन डिक्टीपापांनी पाहिले तर बन्याच्या लेखांनी तसेही मोबाईलस्टोरमध्ये आपल्यांनी यांनी विचारांची नंदी लाभून दृश्य असल्यात वाचावता कोणी विचारांची नंदी लाभून दृश्य असल्यात वाचावता कोणी का? विचारांची नंदी लाभून दृश्य असल्यात वाचावता कोणी का?

कॅलिकोनिंग-अटेन्शन विचारांचील न्होरिया मार्फ याच्या संशोधनासु, एकांदा व्यवस्थापाने फुस्त पूर्वीच्याव संशोधन विचार विचारांची याच्या संसाधारणापाने लेवेस मिळीले लाभातात. एक सध्या लोग दुसर्या मिनिटाला अजून एक नोटिफिकेशन-व्यापय आसेलच असतो. आपली संखोल विचारालांमा दृश्यावृही दैनंदीन तुम्ही होत आहे. तुम्हालाच नस्त याप्ये व इतरांनीही जाळते आहे का, की काही दृश्य नेतृ व सोळ विचार करण्याच्या कामांना आता जास्त बेळ लागतो, किंवा जातव नाही. औंडोगिक खालेल बोलायेचे तर आपली 'उदादाकां' काही होत आहे. तुम्ही कायदी स्लिंग एकांदे पुढतक (मासिक-कर्तमन्त्रक नव्ये) वाचू शकता का? तर नाही असेही तर, विचार गेझी आहे असे समाजा. मोबाईलस्टोर 'डिजिटल केल वीझ' सारखी अंप बघा. त्यात तुम्ही दिवसभार कोळण्या अंवर विचार येव्हे यालवाता, हे दिसेतो. हे आकडे घासावर आणारे ठारीतले आण्यांग असावाईनेतेवर.

या सध्याची संसाधारणाची वापर काळात वेळू, दुर्घटनाम कर्ते टाळता येतील, हे आवाया हातात आहे. यांची गेझी-नोटिच मोबाईल-समाजमाध्यम आणि काम अनिटालैन राहण्यावर असेहीच आहे. त्याना काही पांच्य नाही; एगं यांची तरी तिक्की नाही त्यांच्यासाठी काही उपयोग सुचवता येतील. वैहील उपयोग म्हणजे 'डिविटल मिनिमलिज़म'. कॅलीत कमी वापर, आवाक्षणक अंम, अक्टेन्शन्स, स्वॅच्छिक्षण, डिक्टीपापास काळू टाक्यो, 'डिजिटल फासिंग' म्हणजे वाचावच्याम, मोबाईल न वापरणे. अस 'उदादाक' अनिटालैन एक-दोन दिसम खालीली गेझी मुलांच्या हातात फक्त काळत येतो. त्यांनी असे बोलतो, असे बोलेतो यांनी तुम्ही, नवीकर्त्त्वाची काळत येती 'बोअ' होये किंवा काही न करता चमू राहणे यांनी बोलत बघाव गोष्टी साध्य होतात. मन शात बघायला लागतो, बाहू-उत्तरावेसी विचार वरता येतो, आकुम्हाच्या नवीन कर्त्त्वाना सुचावता, एखाद लाभावला वेळ मिळतो, नवीन कर्त्त्वाना सुचावता, एखाद तास काही न करता, एकूण वसून बघाव.

वसाहतवादाचे आधुनिक रूप

मोठांना समाजाच्यांना कंपन्या खाली वर नसेल एक पॉर्टफोली (यंग) उपलब्ध कराव देतात. तुम्हाच्यावर काही गेझी (मिस, विचारां, रील्स, वैडक्सलस्टर इ.) बनवातात आणि आपण ते पाहाव वसतो. या 'प्रामाणको'ना (असली मोबाईलच्या आणि खुण जगतीप्रमाण असणाऱ्या) वांगले-टॉक्टोकांक पैसेही मिळतात. त्यामुळे माणीची निर्माण-उपभोग, अजून माणपांच-अजून निर्माण हे चक्र अव्याहत चालूच रहेत. फायदा प्रामुख्याने आंतराग्रौप बलाडव कंपन्यांना

(Source: Sakal)

'लक्ष' द्यावे उमजून

मोबाईल न वापरता राहण्याचा एक दिवस प्रयत्न करा. काही न करता बसून राहण्याने बन्याच गोष्टी साध्य होतात. मन शांत होतो. आयुष्याच्या जमा-खर्चाचा मेळ लावता येतो, नवनवीन कल्पना सुचतात आणि बरेच काही साधते.

फार पूर्वीची अर्थव्यवस्था शेतीवर आधारित होती. औद्योगिक क्रांतीनंतर उद्योग व सेवा-क्षेत्र हे अर्थव्यवस्थेचा महत्वाचा भाग होऊ लागले. संगणक आणि इंटरनेट क्रांतीमुळे माहितीची अर्थव्यवस्था बढून घर लागले. सध्या, मोबाईल व समाजमाध्यमांमुळे एक नवीनच अर्थव्यवस्था प्रभावी होत आहे. त्याला आपण 'अटेन्शन इकॉनॉमी' म्हणजेच 'लक्ष अर्थव्यवस्था' म्हणू शकतो. १९७५च्या 'अटेन्शन शासी' या लेखात शावक मानवकल गोळवडवर यांनी 'लक्ष आधारित अर्थव्यवस्थेची परिवर्त्यादा वाच्यता केली. त्याच्या म्हणण्यानुसारा, सध्या स्पृशी ही फक्त माहिती पुरवण्यात राहिलेलीच नाही, तर तुमचे लक्ष राखण्याकाऱ्ये वाललेली आहे. एकांदी वास्तव घायलीची असेहा तर तुमच्या लक्षावर त्यांनी आहे. त्यातून मिळाला अमाप नाही हे त्यांचे 'लक्ष' आहे.

१९९७ सालच्या 'अटेन्शन शॉपर्स' या लेखात शास्त्रज्ञ मायकल गोल्डहाबर यांनी 'लक्ष' आधारित अर्थव्यवस्थेची पहिल्यांदा वाच्यता केली. त्यांच्या म्हणण्यानुसार सध्या स्पर्धा ही फक्त माहिती पुरवण्यात राहिलेलीच नाही तर तुमचे लक्ष राखण्याकडे वळलेली आहे. एखादी वस्तू घ्यायची असेल तर इंटरनेटवर त्याची माहिती असंख्य ठिकाणी मिळू शकते, पण जर एखादी साईट जास्त खिळवून ठेवणारी असेल तर तुम्ही तेथे जास्त वेळ थांबाल तर तुम्हाला जास्त वेळ जाहिराती दाखवता येतील मग त्या साईट-च्या कंपनीला त्याचा जास्त फायदा होईल, नाही का? सध्या माहितीची वानवा-कमतरता नसून तुमच्या 'लक्ष'ची आहे. सध्याच्या डिजिटल युगातील कंपन्यांचे तुमच्या लक्षावर लक्ष आहे आणि त्यातून मिळणारा अमाप नफा हे त्यांचे 'लक्ष्य' आहे.

बन्याच जणांना वाटते की हे सारे मला मस्त फुकट वापरायला मिळत आहे, छान छान क्हिडीओ बघायला मिळत आहे, धमाल करमणूक होत आहे, पण नीट विचार केलात तर असे लक्षात येर्इल की हे बिलकुल फुकट नसून तुम्ही (म्हणजे तुमचे 'लक्ष') विकले जात आहात. जेवढे जास्त तुमचे लक्ष ते खेचतील, गोळा करतील, तेवढा जास्त जाहिरातींचा महसूल. 'लक्ष अर्थव्यवस्था' तुम्हाला मोबाईलला किंवा स्क्रीनला जखडून ठेवण्याकडे भर देते. एखादा क्हिडीओ अथवा रील पाहात असतानाच दुसरे समोर येते अथवा सुचवले जाते. फोटोंचे-छायाचित्रांचे पण तसेच. आपण एका मागून एक बघायला लागलो की तासंतास कसे गेले हे कळतच नाही. सतत काहीतरी नवीन, आकर्षक समोर येत राहते, 'डोपामाईन' मिळते, मग अजून अजून बघावेसे वाटते. ग्रुपवर काही संदेश (मेसेज-नोटिफिकेशन) आला की लगेच वाचावेसे वाटते, उत्तर नाहीतर किमान ईमोजी ताळ्काळ टाकावीशी वाटते. बॉसच्या इमेलला तर उत्तर पुढच्या मिनिटालाच गेलेच पाहिजे याची धडपड, अगदी विषय खूप तातडीचा नसताना सुद्धा. हे सर्व अगदी व्यसनाधीनतेकडे नेण्यासारखेच आहे. काही कारणांनी मोबाईल-समाज-माध्यम वापरायला नाही मिळाले, जवळ नसेल तर अस्वस्थ व्हायला होते, चिडचिड-त्रागा तर काहीच नाही, वैफल्याकडे वाटचाल सुरु होते.

मोबाईल युगाच्या आधीचे दिवस आठवा! तेंव्हा पण आपण गोष्टी लक्ष देऊनच करायचो. पुस्तक वाचणे, नाटक-सिनेमा पाहणे. त्यामुळे 'लक्ष' हे स्वतःहून काही वाईट नाही. पण सतत बदलणारे लक्ष घातक आहे. त्याला कॉन्टेक्स्ट स्विचिंग म्हणजेच 'लक्ष वेगवेगळ्या गोष्टींकडे वारंवार बदलत राहणे' असे म्हणतात. लक्ष सतत विचलित होत राहते. एका ग्रुप वरील मेसेज बघितला की लगेच दुसरा ग्रुपवर नजर गेलीच समजा. पहिल्या रील नंतर दुसरे. ब्राऊझर मध्ये तर असंख्य टॅब्स उघडलेले. याने विचारांची स्थिरता होत नाही. काही नीट विचार करावा लागेल असे मग काही बघावेसेच वाटत नाही. जरा मोठा लेख आला तर तो टाळून आपण पुढे जातो, नाही का? कॅलिफोर्निया-अरवाईन विद्यापीठातील ग्लोरिया मार्क यांच्या संशोधनानुसार एखाद्या व्यत्ययानंतर पुन्हा पूर्वीच्याच सखोल विचार स्थितीत यायला सर्वसाधारणपणे २३ मिनिटे लागतात, पण सध्या लगेच दुसऱ्या मिनिटाला अजून एक नोटिफिकेशन-व्यत्यय आलेलाच असतो. यामुळे आपली सखोल-विचार क्षमता हळूहळू क्षीण होत चालली आहे. तुम्हालापण स्वतःमध्ये व इतरांमध्येही जाणवले आहे का, की काही खूप नीट व खोल विचार करण्याच्या कामांना आता जास्त वेळ लागतो, किंवा जमतच नाही. औद्योगिक भाषेत बोलायचे झाले तर आपली 'उत्पादकता' कमी होत आहे.

उथळपणाकडे आणि सनसनाटी गोष्टींकडे कल जास्त वाढत चालला आहे. याची स्वतःवर एक चाचणी घेऊ शकता. तुम्ही सध्या कधीही (किंवा अगदी सुट्टीच्या दिवशी) सलग एक तास, शांत चित्ताने एखादे पुस्तक (मासिक-वर्तमानपत्र नव्हे) वाचू शकता का? उत्तर नाही असेल तर, विषय गंभीर आहे असे समजा. अजून एका प्रकारे 'स्व' ची ओळख करून घेण्यासाठी तुमच्या मोबाईल मधील 'डिजिटल वेल बीइंग' सारखी एप बघा. त्यात तुम्ही दिवसभरात कोठल्या एप वर किती वेळ घालवता हे दिसते. हे आकडे डोळे उघडणारे ठरू शकतात.

वसाहतवादाचे आधुनिक रूप

मोठ्या समाज माध्यम कंपन्या खरंतर फक्त एक प्लॅटफॉर्म (मंच) उपलब्ध करून देत असतात. तुमच्यातलेच काही (इन्पलुएन्सर्स, प्रभावाक) काही गोष्टी (मिम्स, पिक्चर्स, रील्स, पॉडकास्ट्स इ.) बनवतात आणि आपण सर्व ते पाहत बसतो. नाही म्हणायला या प्रभावकांना (अगदी मोजक्याच आणि खूप जनसंग्रह असणाऱ्या) चांगले-ठीकठाक पैसेही मिळतात. त्यामुळे मागणी-निर्माण-उपभोग-अजून मागणी-अजून निर्माण हे चक्र अव्याहत चालूच राहते. फायदा प्रामुख्याने आंतरराष्ट्रीय बलाढ्य कंपन्यांचा. काहीवेळेला या कंपन्या दादागिरीपण करतात. कुठल्या विचारधारेचा मजकूर ठेवायचा, कोणाला बढावा द्यायचा हे त्यांच्या (म्हणजे त्यांनी बनवल्येल्या प्रणाल्यांच्या, अल्गोरिदम्सच्या) हातात असल्याने ते समाजमन वळवतच नाही तर मोठ्या प्रमाणात नियंत्रित करतात. मागील निवडणुकांमध्ये त्याचा परिचय आपल्याला झालाच असेल. त्यामुळे ज्यांचा समाज माध्यमांवर ताबा त्यांचा (काही प्रमाणात का होईना) जगावर व जगाच्या समाजमनावर ताबा राहतो हे एक ढळढळीत सत्य आहे. मोठी ताकत आहे ही. वसाहतवादाचे हे आधुनिक रूप तेही तुमच्या नकळत. मोबाईल व पर्यायाने समाज माध्यमांचा वापर हा समाजाला व्यसनाधीनतेकडे नेण्यात होतो आहे.

कोठल्याही सार्वजनिक ठिकाणी पाहिले, तर बन्याचश्या लोकांचे लक्ष कोठे असते तर मोबाईल मध्ये. कोणी विचारांची तंद्री लावून शून्य आकाशात बघताना दिसतो का? तर, फारच क्षमित. याचा परिणाम वैयक्तिकच नाही, सामाजिकच नाही तर आंतरराष्ट्रीय सुरक्षेवर पण होऊ शकतो. अशी विचार-क्षीण, मोबाईल मध्ये गुरफटलेली जनता कोठल्या शत्रू राष्ट्राला नको आहे?

एवढ्या सगळ्या वाईट गोष्टी ऐकल्यावर आपल्याला वाटेल की मग मोबाईलचा आणि समाजमाध्यमांचा त्यागच करायचा का? तर, नक्कीच नाही. यांचा कितीतरी प्रभावी पद्धतीने वापर करता येतोय. त्यांनीच ज्ञानाची कवाडे सर्वसामान्यांसाठी खुली केली आहेत. जगातील कानाकोपर्यातील नातेवाईकांना, मित्रमंडळीना जवळ आणले आहे. संवाद वाढला आहे. पारदर्शिकता वाढली आहे. लोकशाही जास्त सुदृढ झाली आहे कारण आता जनतेपासून काहीही लपवणे अवघड होऊन बसले आहे. घरबसल्या काम करणे शक्य झाले आहे. आपली आर्थिक-जीवन पद्धती आमूलाग्र बदलली आहे. हे सर्व फायदे नक्कीच आहेत. आक्षेप आहे तो त्यांच्या वायफळ अति-वापराबद्दल आणि कालांतराने होणाऱ्या व्यसनाधीनतेबद्दल.

आपल्या हातात आहे की या सर्वांचा सकारात्मक वापर कसा वाढवता येईल आणि दुष्परिणाम कसे कमी करिता येतील. ज्यांची रोजी-रोटीच मोबाईल-समाज-माध्यम आणि कायम ऑनलाईन राहण्यावर

अवलंबून आहे त्यांना काही पर्याय नाही पण ज्यांची तशी परिस्थिती नाही त्यांचासाठी काही उपाय सुचवता येतील. पहिला उपाय म्हणजे 'डिजिटल मिनिमलिझम' म्हणजेच या तंत्रज्ञानाचा कमीत कमी वापर. अनावश्यक एप्स, अकाउंट्स, सबस्क्रिप्शन्स, डिक्हाइसेस काढून टाकणे, अगदी मन मारून. अजून पुण्य पाहिजे असेल तर 'डिजिटल फास्टिंग' म्हणजेच समाज-माध्यम-मोबाईलला न वापरत करण्यात येणारा उपवास, आठवड्यातून १-२ दिवस. शाळेत जाणाऱ्या मुलांच्या हातात फक्त कॉल करता येणारे डब्बा-फोन्स, असे बरेच काही. असे विचारावेसे वाटते की तुम्ही, नजीकच्या काळात कधी 'बोअर' (कंटाळा) झाला होतात? मोबाईलमुळे पुढ्यात इतक्या गोष्टी वाढून ठेवल्या आहेत की कंटाळा येणेच दुरापास्त झाले आहे. 'बोअर' होणे किंवा काही न करता बसून राहणे याने खरंतर बन्याच गोष्टी साध्य होतात. मन शांत क्हायला लागते, बाह्य-उत्तेजनेशिवाय विचार करता यायला लागतो, आयुष्याच्या जमा-खर्चाचा मेळ लावायला वेळ मिळतो, नवीन कल्पना सुचतात आणि अजून बरेच काही. गंमत म्हणून एखादा तास काही न करिता, एकटे बसून बघाच.

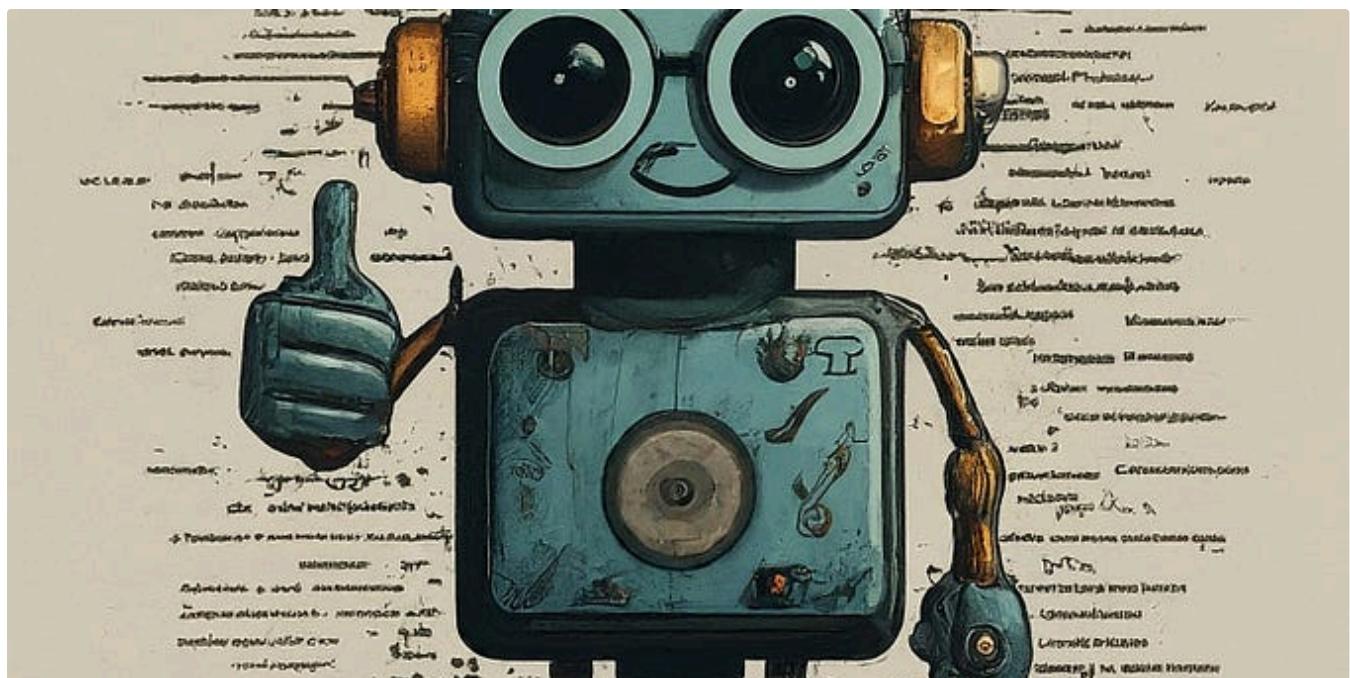
[Marathi](#)[Attention Economy](#)[Mobile](#)[Digital Minimalism](#)[India](#)[Edit profile](#)

Written by Yogesh Haribhau Kulkarni (PhD)

1.6K Followers · Editor for Desi Stack

PhD in Geometric Modeling | Google Developer Expert (Machine Learning) | Top Writer 3x (Medium) | More at <https://www.linkedin.com/in/yogeshkulkarni/>

More from Yogesh Haribhau Kulkarni (PhD) and Desi Stack



Yogesh Haribhau Kulkarni (PhD) in Analytics Vidhya

MESOP, a New UI framework

A new kid on the block from Google

Jun 19 ⚡ 37 🎧 1



Yogesh Haribhau Kulkarni (PhD) in ILLUMINATION

Unraveling the Secrets of the Mind

A Tour of Mental Models

Jun 3 ⚡ 251 🎧 5





 Yogesh Haribhau Kulkarni (PhD) in Google Developer Experts

SaaS LLM

Software-as-a-Service based on Large Language Models

Mar 22, 2023  81



...

 Yogesh Haribhau Kulkarni (PhD) in ILLUMINATION Videos and Podcasts

The Equation of True Happiness

Based on talks by Arthur C Brooks

Sep 17, 2023

179

1



...

See all from Yogesh Haribhau Kulkarni (PhD)

See all from Desi Stack

Recommended from Medium



Figure 1: (a) The comparison showcases Ferret-v2's superior referring and grounding abilities over Ferret (You et al., 2023), particularly in identifying objects and texts within small regions (we zoom in on the regions only for a clearer visualization). (b) Ferret-v2 notably exceeds Ferret's performance in tasks requiring detailed regional and global reasoning and understanding (all w/ 7B models).

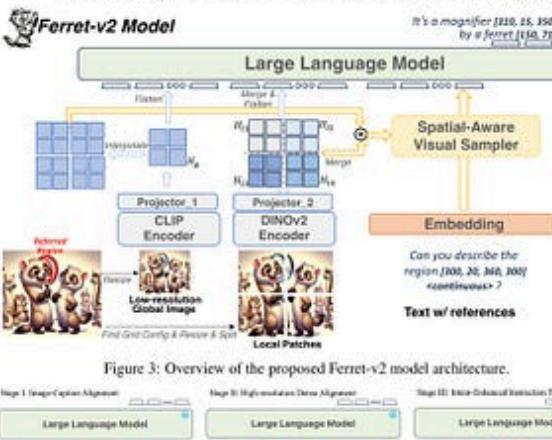


Figure 3: Overview of the proposed Ferret-v2 model architecture.

Models	val	RECOCO	RECOCO+	RECOCOg	Flickr30k Entities
	val	sent	sent	sent	val
MASNet (Yao et al., 2018)	76.28	84.95	76.28	86.00	69.87
CPT-4 (Wang et al., 2022)	79.99	87.60	79.99	87.75	74.48
UNITER (Chen et al., 2020)	81.41	87.88	74.12	75.99	81.45
ViLBERT (Lin et al., 2020)	82.59	87.48	74.84	76.17	82.54
UniTAB (Yang et al., 2022)	86.32	88.84	80.16	78.22	86.48
MDetr (Karpathy et al., 2021)	86.75	89.58	83.41	79.52	84.00
G-DINO4.0 (Liu et al., 2023c)	90.50*	93.19*	88.24*	82.75*	88.95*
Shaka-TB (Chen et al., 2023b)	87.01	90.61	81.24	81.60	87.36
MinGPT-2.7B (Chen et al., 2023d)	88.06	91.29	84.30	79.58	85.52
Qwen-7B (Hou et al., 2023)	88.55	92.27	84.51	82.82	88.59
MoCoV2-1.5B (Xie et al., 2023)	81.10	92.88	87.01	86.62	89.45
LLava-0 (Opher et al., 2023b)	99.16	-	-	91.68	-
VistaLM (Parikh et al., 2023)	88.1	91.5	83.0	82.9	89.8
Ferret-7B (You et al., 2023)	87.49	91.35	82.45	80.78	87.38
Ferret <2-1B (Ours)	82.79	94.68	88.60	87.75	92.75
Ferret <2-1B (Ours)	-	-	89.42	89.42	89.27
Ferret <2-1B (Ours)	-	-	85.52	85.52	85.53

Table 4: Comparison with SoTA methods on 10 benchmarks. Ferret-v2 achieves comparable performance with others. *The training images of the datasets are observed during training.

Models	RECOCO	RECOCO+	RECOCOg	Flickr30k Entities
	val	sent	sent	val
MASNet (Yao et al., 2018)	76.28	84.95	76.28	86.00
CPT-4 (Wang et al., 2022)	79.99	87.60	79.99	87.75
UNITER (Chen et al., 2020)	81.41	87.88	74.12	75.99
ViLBERT (Lin et al., 2020)	82.59	87.48	74.84	76.17
UniTAB (Yang et al., 2022)	86.32	88.84	80.16	78.22
MDetr (Karpathy et al., 2021)	86.75	89.58	83.41	79.52
G-DINO4.0 (Liu et al., 2023c)	90.50*	93.19*	88.24*	82.75*
Shaka-TB (Chen et al., 2023b)	87.01	90.61	81.24	81.60
MinGPT-2.7B (Chen et al., 2023d)	88.06	91.29	84.30	79.58
Qwen-7B (Hou et al., 2023)	88.55	92.27	84.51	82.82
MoCoV2-1.5B (Xie et al., 2023)	81.10	92.88	87.01	86.62
LLava-0 (Opher et al., 2023b)	99.16	-	-	91.68
VistaLM (Parikh et al., 2023)	88.1	91.5	83.0	82.9
Ferret-7B (You et al., 2023)	87.49	91.35	82.45	80.78
Ferret <2-1B (Ours)	82.79	94.68	88.60	87.75
Ferret <2-1B (Ours)	-	-	89.42	89.42
Ferret <2-1B (Ours)	-	-	85.52	85.52

Table 3: Performance comparison (Acc@0.5) on the REC (RefCOCO, RefCOCO+, RefCOCOg) and phrase grounding (Flickr30k Entities) tasks. * indicates that the method is specifically fine-tuned in the second stage.

Models	ENTS (F9)	ENTS (F9)				
	Front	Back	Front	Back	Front	Back
Random Guess	50	50	50	50	50	50
KoTter-2	X	86.25	X	83.87	X	-
Shaka-TB	37.42	65.21	54.15	56.82	38	-
GPT4-80B	X	63.26	X	55.02	X	-
CogVLM-17B	X	79.62	X	61.77	X	-
SIBRNX-2k	72.43	82.97	61.21	63.39	X	-
Ferret-7B	87.94	79.42	69.71	61.91	62.99	57.34

Models	Front	Back	Front	Back	Front	Back
LLava	41.4	31.2	28.8	34.0	-	-
KoTter-2	21.8	35.7	48.4	64.8	-	-
Shaka-TB	46.0	41.6	30.1	45.9	-	-
CogVLM-17B	67.1	67.6	51.7	62.1	-	-
Osgrep-TB	72.2	67.8	-	-	-	-
SPINHX-2k	55.6	70.2	56.6	64.0	-	-

Andrew Lukyanenko

Paper Review: Ferret-v2: An Improved Baseline for Referring and Grounding with Large Language...

Ferret with flexible image resolution handling, DINOv2 and three-stage training!

Apr 15 10



...

InstructBLIP: Towards General-purpose Vision-Language Models with Instruction Tuning

Wenliang Dai^{†1,2*} Junnan Li^{†,✉,1} Dongxu Li¹ Anthony Meng Huat Tiong^{1,3}
Junqi Zhao³ Weisheng Wang³ Boyang Li³ Pascale Fung² Steven Hoi^{✉,1}

¹Salesforce Research ²Hong Kong University of Science and Technology

³Nanyang Technological University, Singapore

<https://github.com/salesforce/LAVIS/tree/main/projects/instructblip>

[†]Equal contribution [✉]Corresponding authors: {junnan.li,shoi@salesforce.com}



Ritvik Rastogi

Papers Explained 156: InstructBLIP

This paper conducts a systematic and comprehensive study on vision-language instruction tuning based on the pretrained BLIP-2 models. 26...

Jun 28

20



...

Lists



Icon Design

36 stories · 360 saves



Visual Storytellers Playlist

60 stories · 391 saves



MODERN MARKETING

171 stories · 730 saves



Interesting Design Topics

257 stories · 649 saves



Louis Chan in Towards Data Science

Computational Thinking: The 4 Step Process for Solutionising Any LLM Use Cases

The Missing puzzle piece to LLM Enterprise Augmentation



Mar 31

141



...



Paul DelSignore in Generative AI

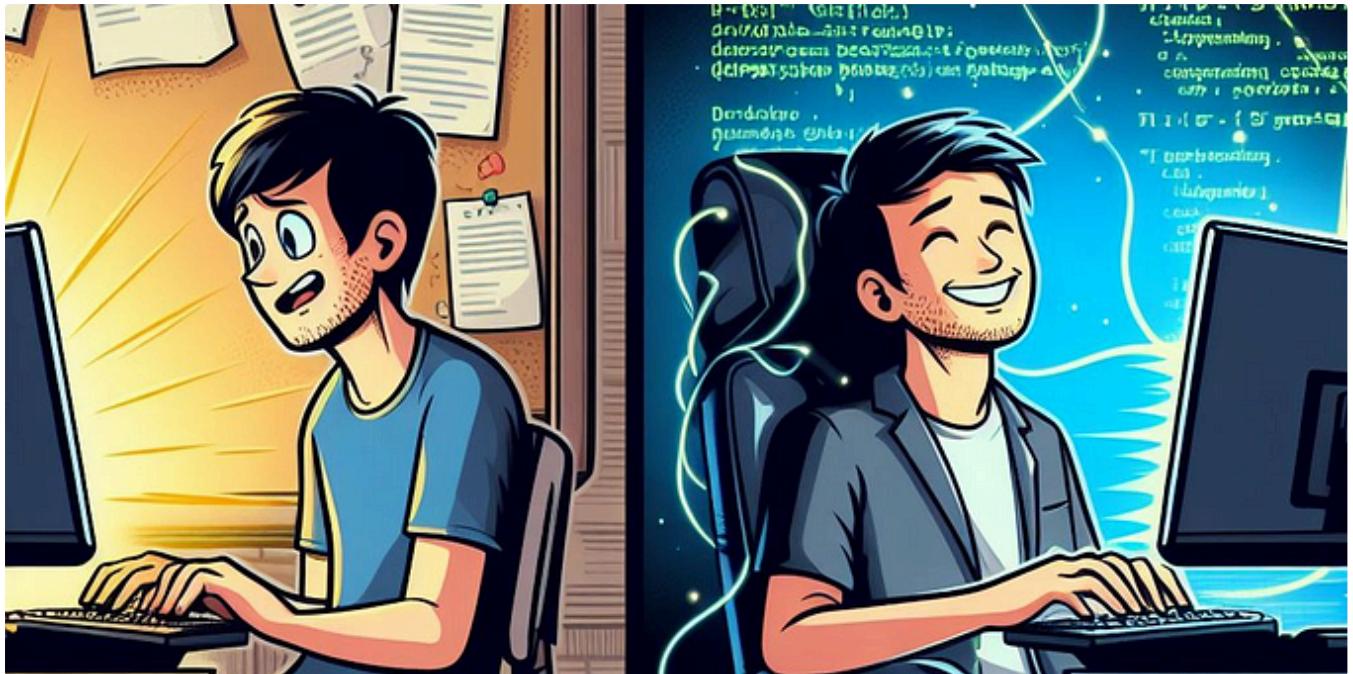
From AI Models to Products: The Shift in AI Strategy

Why Model Performance No Longer Matters

Jul 5 448 10



...



 Abhay Parashar in The Pythoneers

17 Mindblowing Python Automation Scripts I Use Everyday

Scripts That Increased My Productivity and Performance

5d ago 2.8K 20



...

 Emir Ozsaraclar

How an Influencer with 1 Million Followers Uses AI Chatbots to Clone Herself

Imagine a world where your favorite influencers are just a message away, ready to chat with you anytime...

Apr 19  127



...

See more recommendations